

Ingeniería de Software en la Web

Tesis Doctoral

Metodología Cuantitativa para la Evaluación y Comparación de la Calidad de Sitios Web

Luis Antonio OLSINA

Presentada a la Facultad de Ciencias Exactas de la UNLP como parte de los
requisitos para la obtención del título de
Doctor en Ciencias

Director de Tesis: *Dr. Gustavo Hector ROSSI*
Co-director: *Dr. Juan Manuel CUEVA LOVELLE*

La Plata, Noviembre de 1999

**Facultad de Ciencias Exactas
Universidad Nacional de La Plata - Argentina**

Agradecimientos

- ✓ Al Dr. Gustavo Rossi, mi director, por las discusiones enriquecedoras, guía y apoyo constante a mi tarea.
- ✓ Al Dr. Juan Manuel Cueva Lovelle, mi co-director, por su guía y apoyo constante aún en la distancia física que separa a Argentina de España.
- ✓ A los que leyeron la tesis, porque con sus comentarios y observaciones me permitieron enriquecerla.
- ✓ A la Facultad de Ingeniería, de la UNLPam, por el soporte económico facilitado que hicieron posible en buena parte este estudio.
- ✓ A mi esposa Elena y a mis hijas, por soportar varias veces mi ausencia y acompañar mi esfuerzo con cariño y apoyo.
- ✓ A mis colegas del grupo GIDIS, que me alentaron constantemente.
- ✓ A la gente del LIFIA por haberme facilitado las instalaciones del Laboratorio, y haber compartido momentos de trabajo y compañerismo.

Resumen – Abstract – Keywords

Resumen.

El objetivo de esta investigación es definir y discutir una metodología cuantitativa, integral, robusta y flexible para la evaluación de la calidad en aplicaciones centradas en la Web. La metodología, de ahora en más llamada, *Metodología de Evaluación de Calidad de Sitios Web* (Web-site QEM) pretende realizar un aporte ingenieril al proponer un enfoque sistemático, disciplinado y cuantitativo que se adecue a la evaluación, comparación y análisis de calidad de artefactos Web más o menos complejos. Analizamos sus fases y actividades, describimos los distribuibles producidos, y analizamos modelos, métodos, procedimientos, criterios y herramientas a aplicar en dichas actividades. Delineamos nuestro enfoque y soluciones, utilizando a lo largo de la tesis dos casos de estudio para sitios Web, y otros estudios de campo realizados. Además, presentamos un marco conceptual para la validación de métricas y discutimos la validación teórica de algunas métricas empleadas en el proceso. Finalmente, describimos detalladamente un modelo de proceso de evaluación al que Web-site QEM puede adherir. Algunos aspectos y estudios adicionales se presentan en los apéndices.

Abstract.

This research is aimed at defining and discussing a quantitative methodology, the one integral, robust and flexible, for the quality evaluation of Web-based applications. The methodology, from now on called *Web-site Quality Evaluation Method* (Web-site QEM) is intended to perform an engineering contribution proposing for such an end a systematic, disciplined and quantitative approach customized to the evaluation, comparison and analysis of the quality of more or less complex Web artifacts. We analyze their phases and activities, we describe the produced deliverables, and we present models, methods, procedures, approaches and tools to apply in these activities. We outline the approach and its solutions, using along the thesis two case studies for Web sites, and other carried out field studies. Also, we present a conceptual framework for metric validation and we discuss the theoretical validation of some utilized metrics in the Web evaluation process. Finally, we describe thoroughly an evaluation process model to which Web-site QEM can adhere. Some additional aspects and studies are presented in annexes.

Keywords: Web-site QEM, Methodology, Web-site Evaluation, Quality Models, Quantitative, Quality, Characteristic, Attributes, Metric, Process Model, Validation, and Aggregation Models.

Tabla de Contenidos

1. Introducción	1
1.1 Principales Contribuciones	3
1.2 Estructura de la Tesis	3
2. Aspectos del Desarrollo y de la Evaluación de Sistemas de Información Centrados en la Web.	6
2.1 Panorama	6
2.2 Artefactos Web como Software	8
2.3 Necesidad de Modelos de Proceso de Desarrollo de Artefactos Web	11
2.4 Necesidad de Evaluar para Comprender y Mejorar	13
3. Problemática en la Evaluación de Calidad de Artefactos de Software	15
3.1 Presentación del Problema	15
3.1.1 Subjetividad y Objetividad en los Sistemas de Evaluación Complejos	18
3.1.2 Enfoques Cuantitativos para la Evaluación de Calidad	20
3.1.2.1 Modelos de Agregación y Puntaje	21
3.2 Estado del Arte en la Evaluación de Artefactos Web	23
4. Panorama de la Metodología Cuantitativa de Evaluación y Comparación de Calidad de Sitios Web	28
4.1 Motivaciones y Aportes	28
4.1.1 Necesidad de una Metodología Integral, Robusta y Flexible para Evaluar la Calidad de Artefactos Web Complejos	28
4.1.2 El Enfoque Propuesto: Metodología de Evaluación de Calidad de Sitios Web o, en su forma abreviada, Web-site QEM	29
4.1.3 Necesidad de un Modelo de Proceso de Evaluación	30
4.2 Panorama de las principales fases, procesos y modelos de la Metodología Web-site QEM	32
4.2.1 Ejemplos a Utilizar	36
4.2.1.1 Poniendo el foco en la Fase Operativa de un Sistema de Información centrado en la Web	37
4.2.1.2 Caso de Estudio de Museos en la Web	38
4.2.1.3 Caso de Estudio de Sitios Web Académicos	40
5. Fase de Definición y Especificación de los Requerimientos de	43

Calidad

5.1 Definición y/o Selección del Dominio de la Aplicación	43
5.2 Definición de las Metas de Evaluación	45
5.2.1 Enfoques Orientados a Metas	46
5.3 Selección del Perfil de Usuario	47
5.3.1 Perfiles de Usuario Prescritos por el Estándar ISO/IEC 9126	47
5.3.2 Perfiles de Usuario considerando las Metas y el Dominio de Aplicación en la Web	48
5.4 Representación de Características y Atributos de Calidad: Modelos de Calidad	49
5.4.1 Los Primeros Modelos de Calidad de Producto	50
5.4.1.1 Otros Modelos o Marcos Conceptuales de Descomposición	52
5.4.2 Modelos de Calidad de Producto Prescritos en los Estándares ISO 9126, e IEEE 1061	54
5.4.2.1 El Modelo de Calidad Actualizado en el Estándar ISO 9126-1	58
5.4.3 Seleccionando el Enfoque de Modelo de Calidad	60
5.4.4 Seleccionando Características conforme al Perfil de Usuario	61
5.5 Hacia un Arbol de Requerimientos de Calidad General para dominios Web	62
5.6 Características, Subcaracterísticas y Atributos en la Web: Un Ejemplo de Arbol de Requerimientos de Calidad.	63
5.7 Plantilla de Referencia de Variables y Parámetros	66
5.8 Consideraciones Finales	70
6. Fase de Definición e Implementación de la Evaluación Elemental	72
6.1 Criterio de Evaluación Elemental para Atributos	72
6.1.1 Criterios de Preferencia de Calidad Elemental	73
6.2 Representación Notacional de los Criterios	76
6.3 Tipos de Criterios de Preferencia de Calidad Elemental	77
6.3.1 Criterios Elementales Absolutos con Variable Continua	78
6.3.1.1 Criterio de Variable Unica	78
6.3.1.2 Criterio de Variable Normalizada	79
6.3.1.3 Criterio de Multi-variables Continuas	79
6.3.1.4 Criterio de Preferencia de Calidad Directa	79
6.3.2 Criterios Elementales Absolutos con Variable Discreta	80
6.3.2.1 Criterio Binario	80
6.3.2.2 Criterio de Multi-nivel	81
6.3.2.3 Criterio de Multi-nivel definido como Subconjunto	81
6.3.2.4 Criterio de Multi-variables Discretas	81
6.3.3 Otros Criterios Elementales	82
6.4 Un Marco Descriptivo para Representar a las Características y Atributos de Calidad en consideración de sus Criterios de Preferencia de Calidad:	82

Ejemplos.	
6.4.1 Especificación de Algunas Características de Calidad	83
6.4.2 Especificación de Algunas Subcaracterísticas de Calidad	85
6.4.3 Especificación de Atributos	86
6.5 Implementación de las Mediciones Elementales	94
6.5.1 Herramientas de Soporte a la Automatización en el Proceso de Recolección de Datos	94
6.6 Generación de Informes	96
6.7 Análisis de las Preferencias Elementales	97
7. Fase de Definición e Implementación de la Evaluación Global	99
7.1 Fase de la Evaluación Global	99
7.2 El Modelo de Agregación Lógica de Preferencias	101
7.2.1 Principios del Modelo LSP	101
7.2.1.1 Grados de Conjunción y Disyunción.	103
7.2.1.2 Algunas Propiedades Deseables de la Función CDG	105
7.2.2 Empleo de los Operadores de LSP para modelar Relaciones Lógicas entre Características y Atributos	107
7.2.2.1 Tipos de Funciones de Agregación.	108
7.3 Agregación de las Preferencias Elementales para Producir un Indicador de Calidad Global	109
7.3.1 Consideraciones sobre las Actividades en el Proceso de Agregación	110
7.3.2 Cómputo de las Preferencias Parciales y Globales	114
7.4 Procedimientos y Herramientas de Soporte	115
7.4.1 Algunos Detalles del Procedimiento para el Cómputo de las Preferencias de Calidad	115
7.4.2 Ambiente integrado WebQEM_Tool para el soporte de cálculos y documentación	116
8. Fase de Análisis de Resultados, Recomendaciones y Documentación	118
8.1 Empleo de Tablas, Gráficos, y Herramientas	118
8.2 Análisis y Comparación de Resultados Parciales y Globales	119
8.2.1 Análisis para Usabilidad	121
8.2.2 Análisis para Funcionalidad	122
8.2.3 Análisis para Confiabilidad y Eficiencia	125
8.3 Consideraciones finales sobre el caso de estudio	127
9. Ejemplos: Empleo de Web-site QEM	129
9.1 Estudios de Campo Realizados	129
9.2 Caso de Estudio sobre Museos Típicos en la Web	130
9.2.1 Arbol de Características y Atributos	130

9.2.2 Evaluación Elemental	134
9.2.2.1 Mediciones Elementales	139
9.2.3 Evaluación Global	139
9.2.4 Análisis y Comparación de la Calidad de los Sitios Evaluados	141
9.3 Caso de Estudio en Preparación: Sitios Web de Comercio Electrónico	147
10. Aspectos de Validación de Métricas para la Web	150
10.1 Introducción	150
10.1.1 Sistema Relacional Empírico y Formal	153
10.1.2 Magnitud, Tipos de Escala y Unidades	154
10.2 Un Marco Conceptual para Validar Métricas	158
10.2.1 Estructura y Modelo Conceptual para Métricas	158
10.2.1.1 Entes, Atributos y sus Relaciones	159
10.2.1.2 Unidades, Tipos de Escalas y sus Relaciones	160
10.2.1.3 Valor	160
10.2.1.4 Instrumento de Medición	160
10.2.1.5 Métricas Indirectas	160
10.2.2 Implicaciones del Modelo para Validación de Métricas	161
10.2.3 Otros Modelos para Validación de Métricas	162
10.2.4 Estrategias para Validación de Métricas	163
11. Modelo de Proceso de Evaluación de Artefactos de Software	167
11.1 Modelo de Proceso de Evaluación definido en el Estándar ISO 9126	168
11.2 Proceso para Evaluadores definido en el Estándar ISO/IEC 14598	170
11.2.1 Desarrollo del Proceso de Evaluación	172
11.3 Consideraciones Finales	178
12. Conclusiones y Líneas de Trabajo Futuro	180
12.1 Futuros Avances	182
12.1.1 Evaluación basada en Preferencias de Calidad y Costo	182
12.1.2 Hacia Requerimientos de Calidad Generales para Dominios Web	183
12.1.3 Patrones de Diseño para Dominios Web	184
12.1.4 Validación de Métricas Predictivas para proyectos Web	184
12.1.5 Integración del Ambiente WebQEM_Tool	185
13. Glosario de Términos	186
13.1 Aclaraciones	186
13.2 Definiciones	186
14. Referencias	200
APENDICE A. Información Adicional de los Estudios de Campo Empleados	210

A.1 Información Adicional para el Caso de Estudio de Sitios Académicos	210
A.1.1 Plantillas de Características y Atributos de Calidad	210
A.1.2 Resultados de las Preferencias Elementales para los Sitios Universitarios	213
A.2 Información Adicional para el Caso de Estudio de Museos en la Web	215
A.3 Un Survey sobre 24 Museos	216
A.3.1 Preparación del Survey	216
A.3.2 Análisis de los Datos y Tendencias	218
A.3.2.1 Análisis para Usabilidad	218
A.3.2.2 Análisis para Funcionalidad	220
A.3.2.3 Análisis para Eficiencia	223
A.3.3 Observaciones Finales	223
APENDICE B. Una Estructura Conceptual para un Modelo de Proceso Genérico	228
B.1 Conceptos sobre Procesos de Software y Modelado de Procesos y Productos	228
B.2 Definición de un Modelo Conceptual	236
B.2.1 Descripción de las Clases y sus Relaciones Fundamentales	238
B.2.2 Comentarios finales	240
APENDICE C. Empleo de un Modelo Orientado a Metas (Enfoque Goal-Question-Metric en el Proyecto <i>Facultad de Ingeniería</i>)	242
C.1 Características deseables de Artefactos, Procesos, y Recursos en Proyectos de Hipermedia	242
C.2 Goal-Question-Metric como Enfoque Orientado a Metas para Mediciones en Proyectos de Software	246
C.3 Midiendo el Nivel de Interconexión en el Contexto “Ubicación” del CD- ROM “ <i>Facultad de Ingeniería</i> ”	249
C.4 Consideraciones Finales	253
APENDICE D. Aspectos del Modelo LSP	255
D.1 Función para el Cálculo de las Preferencias Parciales y Globales	255

Capítulo 1

Introducción.

Los desarrollos Web son cada vez más complejos y, además, están creciendo rápidamente, entre ellos las aplicaciones de software centrados en la Web. Este tipo de artefactos (sitios) Web puede ser un sistema de publicación de catálogos con lógica de comercio electrónico, o un sistema de planificación y programación de proyectos de desarrollo colaborativo, entre otros, proveyendo funcionalidad que está más cercana a una implementación de software cliente/servidor tradicional que a un sitio Web estático orientado a la documentación. Sin embargo, modelos de proceso de producto bien definidos que promuevan un ciclo de vida de desarrollo repetible y eficiente, y modelos de proceso y metodologías de evaluación que promuevan la comprensión y la mejora de la calidad de artefactos Web, no están acompañando este rápido crecimiento observado. Más bien las prácticas actuales para desarrollar sitios Web son ad hoc, y el aseguramiento y control de la calidad es, por lo general, un proceso a cumplimentar en el futuro [Deshpande et al 99; Lowe et al 99, Olsina 98c, Powell et al 98]. Por lo tanto, en este estado de situación, la utilización sistemática y disciplinada de métodos, modelos, y técnicas de Ingeniería de Software para el desarrollo, el mantenimiento, y la evaluación de la calidad de los sitios Web debiera ser un requerimiento obligatorio, principalmente en los proyectos de mediana o gran escala.

Una de las metas principales de la evaluación y comparación de calidad de artefactos Web, radica en medir, analizar y comprender el grado de cumplimiento de un conjunto de características y atributos con respecto a los requerimientos de calidad establecidos, para un perfil de usuario y dominio de aplicación dados.

Por una parte, los desarrollos centrados en la Web en los más diversos dominios de aplicación como comercio electrónico, sistemas académicos, financieros, entre otros, se están tornando cada vez más en sistemas complejos. Por lo tanto, desde el punto de vista de la valoración de productos, un proceso de evaluación cuantitativo e integral, que considere la mayoría de las características y atributos relevantes para cierto perfil de usuario, se torna también en una tarea compleja. La complejidad en la evaluación es producto de la gran cantidad de características y atributos que pueden intervenir en los requerimientos de calidad y en las varias relaciones existentes entre los atributos subcaracterísticas y características, entre otros aspectos.

Por otra parte, en términos generales podemos clasificar a las metodologías, métodos y técnicas de evaluación en dos amplias categorías: cuantitativas y cualitativas. Aún si la evaluación de sistemas informáticos y, particularmente de programas de software, acredita más de tres décadas como disciplina [Gilb 69, McCall et al 77, entre otros], la

evaluación cuantitativa y sistemática de aplicaciones de Hipermedia y, específicamente, de sitios Web es más bien una preocupación reciente y, a veces, poco considerada. De hecho, es importante destacar que Garzotto et al. [Garzotto et al 95] han introducido algunos criterios de evaluación como riqueza, facilidad, consistencia, legibilidad, reuso, etc., para evaluar de un modo cualitativo a sistemas o aplicaciones de Hipermedia; no obstante, este enfoque es conveniente cuando el producto a evaluar es más bien simple e intuitivo. En el caso en que se desee expresar a la aplicación con mayor cantidad de factores y atributos de más bajo nivel se vuelve difícil una evaluación justificable y precisa, en donde se dificulta identificar menores diferencias entre valores similares de atributos de sistemas comparativos.

Además, a mediados de la década de los noventa surgieron guías de estilo, diseño y publicación para la Web [IEEE WPG, Rosenfeld et al 98], con el objeto de asistir a los participantes en el proceso desarrollo. En esta misma dirección, surgido de organizaciones reconocidas como W3C [W3C 99], se han prescrito listas de buenas prácticas, con asignación de prioridades, para que el desarrollador las tenga en cuenta a la hora de diseñar sitios de un modo más usable, navegable, eficiente y accesible. Por lo tanto, estas guías y principios son muy útiles, debido a que documentan características, atributos y criterios de calidad a tener en cuenta en el proceso de evaluación. El empleo apropiado de esas guías y criterios, debe ayudar a mejorar el proceso de diseño y autoría en la Web pero, obviamente, no constituyen de por sí una metodología de evaluación de artefactos Web.

Finalmente, estudios y surveys de usabilidad y de interfaces son muy bien conocidos en la comunidad de Hipermedia e Interfaces de Usuario [Bevan 95 y 97, Nielsen 93, Schneiderman 87]. Particularmente, Nielsen ha realizado muy interesantes surveys sobre características y atributos en la Web, a partir de 1994, que están en línea en su columna electrónica denominada Alertbox [Nielsen_Alert]. Asimismo, están emergiendo actualmente estudios y evaluaciones cuantitativas en dominios específicos de la Web [Kirakowski et al 98, Lohse et al 98]. Por ejemplo, Lohse et al. identificaron y midieron 32 atributos para 28 negocios conectados a Internet e identificaron características y atributos de diseño que influyen al tráfico y a las ventas. Sin embargo, estos estudios estadísticos y descriptivos (ver además los surveys de [GVU 99], aunque reveladores para marcar el estado del arte de la calidad de los atributos y marcar tendencias, no constituyen una metodología amplia para evaluar características, atributos y sus relaciones, de sitios Web como un todo.

Por lo tanto, podemos afirmar que en el proceso de evaluación de requerimientos de calidad de artefactos Web complejos, se observa la necesidad de contar con una metodología cuantitativa, integrada, flexible y robusta, que se apoye en principios y prácticas de Ingeniería de Software para la evaluación y comparación de características y atributos, con el fin de obtener resultados objetivos y justificables.

1.1 Principales contribuciones

Las principales contribuciones de esta tesis para el área de proyectos de evaluación y comparación de calidad de aplicaciones centradas en la Web son, a saber:

- ✓ *Un enfoque integral, sistemático, disciplinado y cuantitativo, para evaluar y comparar artefactos Web, tanto en la fase operativa del ciclo de vida como en la fase de desarrollo.* Se discuten las fases, principales actividades y artefactos (documentos) producidos; esta clara división en fases y actividades favorece la visibilidad de un proyecto de evaluación, ayudando a la planificación, programación, ejecución y control de las mismas.
- ✓ *Un conjunto de modelos, métodos, procedimientos, criterios y herramientas a ser utilizado en las diferentes fases y actividades.* Los mismos pueden ser empleados en el marco conceptual de un modelo de proceso de evaluación estandarizado.
- ✓ *Contribuir potencialmente a comprender y mejorar la calidad en el diseño y rediseño de aplicaciones centradas en la Web.* Esta declaración está motivada en que la metodología de evaluación de calidad propuesta propende a un uso riguroso y sistemático de principios y prácticas establecidas de la Ingeniería de Software aplicados al dominio de aplicaciones Web.
- ✓ *El empleo de un marco conceptual para validar métricas en la Web.* Se validan teóricamente varias métricas empleadas (orientadas a la valoración antes que a sistemas de predicción) .
- ✓ *Definir y especificar características, subcaracterísticas y atributos que contribuyen a la calidad de artefactos Web, en consideración de un perfil de usuario y dominio dados.* (Estas investigaciones nos permitirán generalizar subcaracterísticas y atributos para conjuntos de dominios: así, patrones de reuso se pueden catalogar denominados más precisamente en la literatura como *patrones de diseño*).

1.2 Estructura de la Tesis

En esta investigación presentamos a la *Metodología de Evaluación de Calidad de Sitios Web* (Web-site Quality Evaluation Method, o Web-site QEM), cuyo objetivo es realizar un aporte ingenieril que se adecue a la evaluación, comparación y análisis de calidad de sistemas de información centrados en la Web más o menos complejos. Presentamos sus fases y actividades, y analizamos modelos, métodos, procedimientos, criterios y herramientas a aplicar en dichas actividades. Específicamente utilizaremos a lo largo de la tesis, dos casos de estudio realizados [Olsina 99, Olsina et al 99 a, c] para

ejemplificar sus distintos aspectos.

En la primera parte de la tesis (capítulos 2 y 3), nos dedicamos a discutir aspectos generales relacionados a los artefactos Web, la necesidad de usar modelos de proceso de desarrollo y modelos de proceso de evaluación. Además, nos concentramos en la problemática en la evaluación de software y el estado del arte.

Particularmente, en el capítulo 2, realizamos una breve discusión de las peculiaridades de los artefactos Web y las semejanzas y diferencias con los productos de software tradicionales; presentamos la necesidad de contar con modelos de ciclo de vida de desarrollo adaptados a la Web, y de contar con modelos de proceso y metodologías de evaluación. En el capítulo 3, analizamos la problemática en la evaluación de calidad de artefactos de software en general, y nos detenemos en desarrollar un conjunto de preguntas y temas claves. Además, discutimos el estado del arte en la evaluación de artefactos Web y trabajos relacionados.

En la segunda parte de la tesis (capítulos 4 al 9), presentamos y desarrollamos detalladamente a la *Metodología de Evaluación de Calidad de Sitios Web*.

En el capítulo 4 describimos, en la primera sección, las motivaciones de contar con una metodología integral, robusta y flexible para evaluar la calidad de artefactos Web más o menos complejos; en la segunda sección, presentamos un panorama de las principales fases, procesos y modelos de la Metodología Web-site QEM. Finalmente, introducimos los ejemplos que se irán utilizando a lo largo de esta tesis.

La fase de definición y especificación de los requerimientos de calidad se discute en el capítulo 5. Esta fase trata con actividades y procedimientos para la elicitación, modelado y especificación de los requerimientos de calidad. Particularmente, se discuten distintos modelos de calidad y se propone un enfoque de modelo mixto. Este enfoque se emplea para representar el árbol de requerimientos de calidad de artefactos Web, para un perfil de usuario y dominio dados. En el capítulo 6, analizamos la fase de definición e implementación de la evaluación elemental. La misma trata con actividades, modelos, herramientas y heurísticas para determinar criterios de evaluación y métricas para cada atributo cuantificable. En el capítulo 7, discutimos la fase de definición e implementación de la evaluación global; se presenta actividades, modelos, procedimientos y herramientas para determinar los criterios de agregación de las preferencias de calidad elemental (obtenidas en la fase anterior, a partir del árbol de requerimientos), para producir la preferencia global para cada artefacto Web interviniente.

En la fase de análisis de resultados, recomendaciones y documentación (en el capítulo 8), se realizan las actividades de análisis y comparación de las preferencias de calidad elementales, parciales y globales, y, asimismo, la justificación de los resultados. A partir

de las metas establecidas y del punto de vista de usuario a evaluar, el proceso culmina con las conclusiones y recomendaciones del caso. Se discuten los resultados del estudio de sitios académicos en la Web (el cual se fue desarrollando en los capítulos 5, 6 y 7). Finalmente, en el capítulo 9, presentamos otros estudios de campo realizados por medio del uso de Web-site QEM.

En la tercera parte de la tesis, presentamos un marco conceptual para la validación de métricas (capítulo 10) y discutimos la validación teórica de algunas métricas empleadas en el proceso. Por otro lado, en el capítulo 11, describimos detalladamente un modelo de proceso de evaluación al que Web-site QEM puede adherir.

Por último, exponemos las conclusiones y futuros avances en distintas direcciones.

El lector encontrará a lo largo del trabajo referencias bibliográficas, que se encuentran al final, como así también un glosario de palabras y frases claves que son de importancia para la tesis. Además, cuando sea oportuno, se remite al lector a los distintos capítulos y secciones para facilitar la comprensión de este documento.

Adicionalmente, hemos incorporado cuatro apéndices con información complementaria. En el apéndice A, se vuelca información adicional de los estudios de campo empleados. Es de interés los resultados arrojados por el survey realizado sobre el dominio de museos en la Web (sobre 24 sitios) empleando unas pocas actividades de Web-site QEM. En el apéndice B, se define una estructura o modelo conceptual útil para comprender y soportar a la modelización de procesos en general. En el apéndice C, se describe una experiencia de evaluación que realizamos sobre una aplicación editada en CD-ROM (en 1995), empleando el enfoque GQM; y, por último, en el apéndice D, se presenta un algoritmo codificado en C++, que es una parte del programa para obtener las preferencias parciales y globales en el proceso de evaluación.

Capítulo 2

Aspectos del Desarrollo y de la Evaluación de Sistemas de Información Centrados en la Web.

2.1 Panorama

El crecimiento de Internet (y las así llamadas Intranets y Extranets), y, en general, el crecimiento de la Web (World Wide Web) está implicando un fuerte impacto y cambio paradigmático en el manejo de la información en tan diversos sectores como comerciales, educativos, industriales, financieros, de entretenimiento, gubernamentales, y porqué no decirlo también, poco a poco está cambiando en nuestras vidas personales, el modo de ver y relacionarlos con el mundo (en un modelo globalizado). Particularmente, en nuestros días asistimos a un crecimiento paulatino de sitios Web con funcionalidad de comercio electrónico, y hasta vale mencionar, que se han realizado experimentos de cómo personas “aisladas” pueden proveerse de bienes y servicios, e intentan valerse y sobrevivir (con algún éxito), ¡tan sólo con una computadora conectada a Internet ante sus manos! Lo cierto es que una amplia gama de aplicaciones distribuidas cada vez más complejas está emergiendo en el entorno de la Web. El indiscutido incremento de popularidad y aceptación de las aplicaciones centradas en la Web tiene su origen en sus propias características y naturaleza: provee un modelo de información que soporta enlazar todo tipo de contenido y medios; permite a los usuarios un fácil acceso a información a partir de una representación ubicua de la misma; a diferencia de un documento en papel, tiene el valor agregado de la funcionalidad de software; y, entre otros aspectos, permite a los usuarios (no necesariamente profesionales del área de informática) crear y enlazar fácilmente sus propios documentos por medio de herramientas ampliamente disponibles.

Sin embargo, como indicábamos en la introducción del capítulo 1, esta explosión de interés en publicar en la Web y en adicionarle contenido y funcionalidad de aplicación a los sitios Web ha generado sus propios desafíos a la Ingeniería de Software. En general se observa que modelos de proceso de producto bien definidos que promuevan un ciclo de vida de desarrollo repetible y eficiente, y modelos de proceso de evaluación que promuevan la comprensión, el control y la mejora de la calidad de artefactos Web, no están acompañando este rápido crecimiento. Más bien, las prácticas para desarrollar sitios Web son ad hoc en donde el proceso típico consiste en iteraciones de subprocesos como implementar (y prototipar), probar un poco y distribuir (a semejanza del proceso “code and fix” o “just-do-it” para desarrollar de un modo ad hoc, software tradicional). Es cierto que tal estrategia puede ser adecuada para pequeños proyectos de desarrollo Web, en donde el futuro mantenimiento es previsiblemente mínimo.

Con todo, actualmente existe en la comunidad científica del área una legítima y creciente preocupación tanto en el modo en cómo se desarrollan los sistemas centrados en la Web como en la integridad y calidad de los productos a mediano y largo plazo. Por ejemplo, en ausencia de procesos disciplinados para producir artefactos Web, se está enfrentando con serios problemas en el efectivo desarrollo y ulterior evolución, principalmente en proyectos de mediana o gran envergadura. Desarrollos de aplicaciones centrados en la Web realizados mediante estrategias ad hoc tienen alta probabilidad de fracaso en fases de mantenimiento y evolución, y esto se acentúa más, en la medida en que crece la complejidad de los mismos. Con el fin de evitar una posible crisis en los desarrollos en la Web [Gibbs 94, Powell et al 98] y alcanzar mayor efectividad en este contexto de creciente diversidad y complejidad de requerimientos funcionales y no-funcionales [IEEE 830], es necesario emplear enfoques disciplinados y sistemáticos, herramientas y metodologías robustas y flexibles para construir, evolucionar y evaluar artefactos Web. Es oportuno resaltar que tales estrategias, metodologías y herramientas deban tener en cuenta aspectos específicos de este nuevo medio como: 1) el nivel de orientación a la documentación versus el nivel de orientación a la funcionalidad de la aplicación de software; 2) la multiplicidad de los perfiles de usuarios y sus distintos comportamientos; 3) características y atributos de calidad que respondan a las necesidades de las diferentes audiencias en consideración de las peculiaridades del medio Web; 4) diferentes tipos de habilidades y conocimientos de los participantes en un proyecto centrado en la Web; 5) procesos de desarrollo de rápida generación de productos pero flexibles y robustos en cuanto a la evolución tanto de la estructura y contenido como de la funcionalidad asociada.

Principalmente, se está observando la necesidad de contar con un enfoque ingenieril; esto es, el empleo disciplinado, cuantificable y sistemático de principios y prácticas reconocidas de Ingeniería de Software para la creación, evaluación, mantenimiento y evolución de aplicaciones centradas en la Web¹.

Por lo tanto, en este capítulo introduciremos en qué medida los artefactos y desarrollos Web se asemejan a los artefactos y desarrollos de software “tradicional” y en qué

¹ En los últimos dos años, se observa una creciente preocupación de la comunidad científica de Ingeniería de Software (y otras comunidades relacionadas) por los temas de modelos de proceso de desarrollo y, en menor grado, de evaluación de aplicaciones centradas en la Web, de modo que recientemente se ha dado en llamar a la disciplina, Ingeniería de Software en la Web, o brevemente, Ingeniería en la Web. Por ejemplo, en 1998, se celebró el primer workshop sobre Ingeniería en la Web, en conjunción con la 7ma Conferencia World Wide Web (WWW), en Australia. Asimismo, durante este año (1999) se celebraron dos workshops de Ingeniería en la Web: uno en el congreso WWW8, en Toronto, Canadá, y el otro en el 21 Congreso Internacional de Ingeniería de Software (ICSE), en Los Angeles. En cuanto a publicación de libros, en 1998, se editó el libro de Powell et al denominado “*Web Site Engineering: beyond Web Page Design*” [Powell et al 98], y recientemente, apareció el libro de Lowe & Hall denominado “*Hypermedia and the Web: an Engineering Approach*”. En cuanto a revistas científicas tratamientos del tema se encuentran en IEEE Software (Sep/Oct 98); en CACM (Jul 98); en WebNet Journal (Nº 1, 1999); y en NRHM Journal (por aparecer la edición de 1999).

medida se diferencian. Esto nos ayudará a comprender características y atributos semejantes y distintivos de los artefactos y sistemas Web, útiles a tener en cuenta en un proceso de desarrollo y evaluación de calidad. Por otra parte, discutiremos la necesidad de contar con procesos y metodologías bien definidas para el desarrollo y la evaluación. Nosotros consideramos que la utilización sistemática de métodos, modelos, y técnicas de Ingeniería de Software para la evaluación, control y mejoramiento de la calidad de los productos Web debiera ser un requerimiento obligatorio en todo proyecto de mediana o gran escala. En este sentido, el objetivo principal de este trabajo consiste en desarrollar una metodología cuantitativa, flexible y robusta que cubra la mayor parte de las actividades en el proceso de evaluación, comparación, y selección de artefactos Web (capítulos 4 al 9). En el capítulo 11, discutiremos el estado del arte en modelos de proceso para evaluadores de productos software y su aplicabilidad en la Web.

El empleo sistemático de estrategias y metodologías basadas en principios y prácticas ingenieriles no nos proveerá la bala de plata (“the silver bullet” [Brooks 87]), pero ayudará a minimizar los problemas y riesgos, y proveerá un marco conceptual y pragmático para administrar eficientemente proyectos de desarrollo, su evolución y la valoración de sus distribuibles.

2.2 Artefactos Web como Software

Hay muchos objetivos por los cuales se puede construir un sitio Web; estos pueden incluir promoción de una organización, comercio electrónico, entretenimiento, soporte técnico, relaciones entre inversores, reclutamiento de personal, presentación de expresiones artísticas y culturales, satisfacción personal, expresiones del dominio académico y de la investigación, reservaciones, por citar algunos. Por lo tanto, los artefactos Web se están tornando cada vez más semejantes a los productos de software tradicionales. Un único sitio Web, por ejemplo, debe satisfacer necesidades tan diversas que pueden ir desde la distribución y presentación de documentos multimediales hiperenlazados hasta la automatización de procesos de negocios.

Según la definición dada por [Sommerville 92], “*software es un programa de computadora, o un conjunto de programas de computadora, además del material relacionado como documentación, el cual se usa para realizar alguna tarea*”. Ahora bien, la pregunta que cabe formularnos es la siguiente:

¿Puede ser considerado un sitio Web como un artefacto de software?

Según nuestra visión, al igual que la de varios autores [Lowe et al 99, Powell et al 98, Pressman 98] un sitio Web puede en muchos casos ser considerado un artefacto de software (y hasta con complejidad añadidas). Un simple sitio Web puede ser no más que un conjunto de páginas estáticas (documentos hiperenlazados) recuperados por un

usuario, a partir de un navegador, desde un servidor remoto. No obstante, en la medida que el sitio provee capacidad de recolectar información desde formularios, provee recuperación de información desde bases de datos por medio de consultas (y acceso a sistemas legados), permite generar páginas dinámicamente (consultas, búsquedas, generación de páginas personalizadas, etc.), permite realizar cálculos por medio de applets o scripts, etc. la funcionalidad del artefacto Web es comparable a la funcionalidad de programas de software.

Sin embargo, dada la naturaleza de orientación a documentos y contenido, no todos los componentes de un sitio Web pueden ser vistos como con funcionalidad de software. Mientras que la publicación de un sitio con componentes estáticos orientados a la documentación debe lograr un adecuado compromiso de calidad entre organización, presentación, contenido y navegación, los componentes orientados a la aplicación (programación) deben proveer la adecuada integración y funcionalidad de la lógica agregada. Con todo, características de calidad del artefacto como confiabilidad y eficiencia, puede ser afectada por ambos componentes. Por ejemplo, un sitio estático pobremente testeado puede contener varios enlaces rotos o inválidos, lo cual degrada la confiabilidad del mismo desde el punto de vista del usuario, o, por otra parte, una falta de testeo de funcionalidad de applets o scripts, puede provocar deficiencias en la presentación dependiendo del tipo de navegador. Del mismo modo, un diseño descuidado de las páginas en consideración de la suma de los tamaños de todos sus componentes, puede degradar la performance del sitio (y en definitiva la eficiencia).

La figura 2.1, muestra una representación de aplicaciones Web considerando rangos de complejidad (estáticas versus dinámicas), y grados de orientación (orientado a documentos versus orientado a aplicaciones). Según esta figura los sitios pueden ser categorizados en [Powell et al 98]:

- ✓ *Sólo Estático*
- ✓ *Sitio Estático con Formularios de Entrada*
- ✓ *Sitio con Acceso de Datos Dinámicos*
- ✓ *Sitio Creado Dinámicamente*
- ✓ *Aplicación de Software basada en la Web*

Sitio Sólo Estático. En su forma más simple, un sitio Web es una colección de páginas estáticas (documentos, o información editada y publicada en formato HTML). Desde el punto de vista de la funcionalidad, ésta es dada básicamente por los enlaces que permiten navegación, ya sean enlaces provistos por los controles principales y navegacionales de la interface, o por los enlaces estructurales (mecanismos de acceso como un índice) y semánticos. En un sitio de este tipo, el énfasis de diseño está puesto en la organización de la estructura y el contenido, en la estética de la presentación, y en las facilidades de navegación (no obstante hay otros desafíos de usabilidad, además de eficiencia, confiabilidad y mantenibilidad).

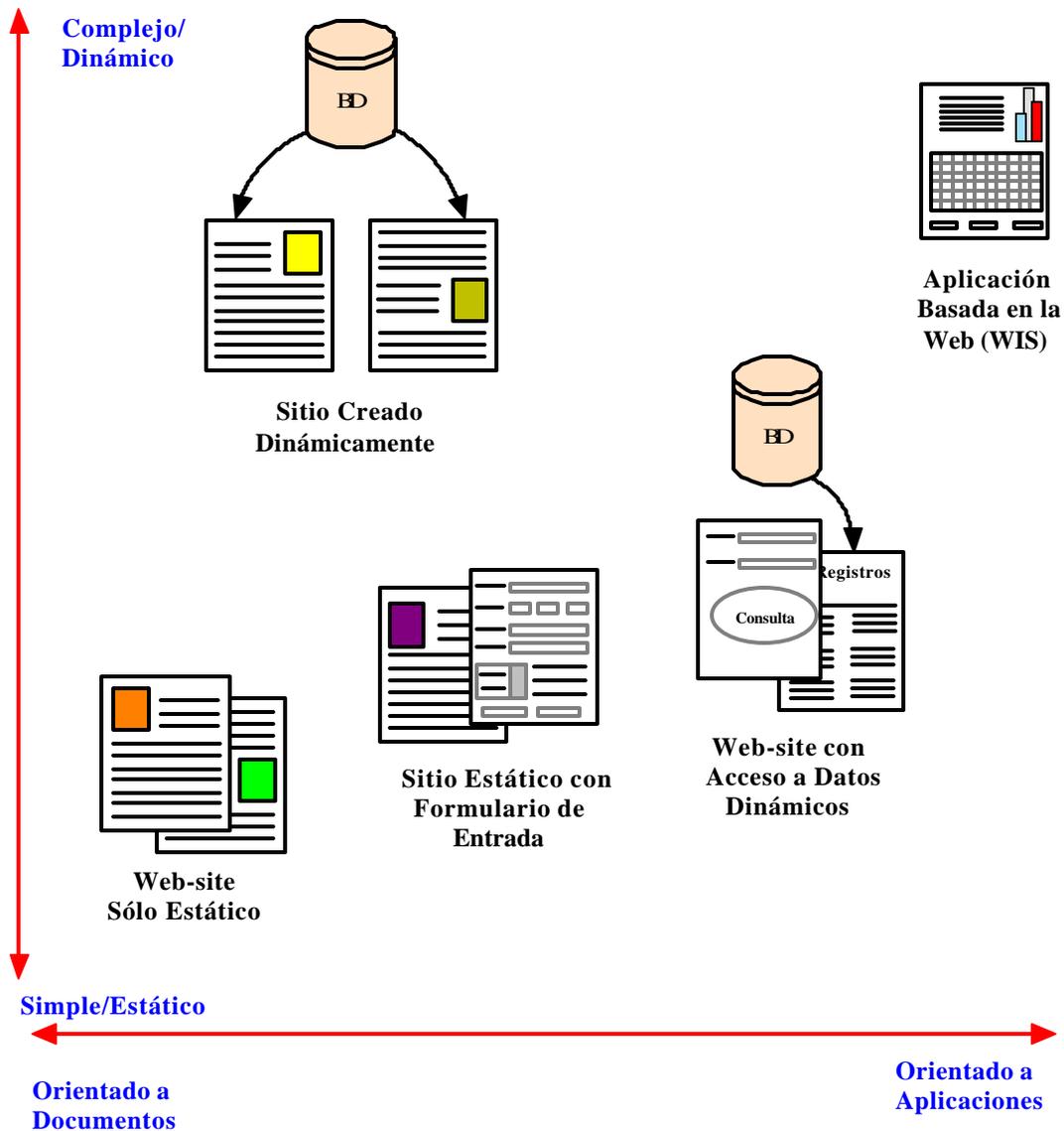


Figura 2.1 Rangos de complejidad y niveles de orientación de las aplicaciones Web [Powell et al 98]

Sitio Estático con Formularios de Entrada. En un sitio de este tipo, el énfasis de diseño está puesto en la organización de la estructura y el contenido, en la estética de la presentación, en las facilidades de navegación, etc., proveyendo además al usuario un nivel de interacción básico implementado por medio de formularios de entrada. Esto favorece la usabilidad del sitio al permitir mecanismos de retroalimentación por parte del usuario. Atributos como *Cuestionarios*, *Libros de Invitados*, o *Comentarios y Sugerencias*, son fácilmente implementables favoreciendo a la comunicación en línea (estos atributos, entre otros, están representados en el árbol de requerimientos de calidad, como se observa en el código 1.2.5 de la figura 5.7).

Sitio con Acceso a Datos Generados Dinámicamente. En un sitio de este tipo, además de las características previamente comentadas, el usuario puede acceder por medio de

las páginas (del lado del cliente) a datos almacenados en bases de datos remotas, por medio de consultas y búsquedas. Los datos retornados a partir de la interacción iniciada por el usuario son generados dinámicamente (ASP, etc.), y presentados en formato de documentos HTML (DHTML o XML).

Sitio Creado Dinámicamente: surge la necesidad de diseñar este tipo de sitios Web, en el caso en que el mismo deba proveer requerimientos semejantes aunque personalizados en consideración del contenido de las páginas para cada instancia de usuario; o cuando por aspectos de compatibilidad tecnológica, se requiere construir dinámicamente a sitios en conformancia con el entorno del navegador del usuario. Para ello, los documentos estáticos deben ser mudados dinámicamente aunque en el lado del cliente, no provean interactividad alguna.

Aplicación de Software basada en la Web: este tipo de artefacto puede ser el más complejo y con mayor orientación a la aplicación, como se aprecia en la figura 2.1. Este tipo de sitio Web (a veces implementado como Intranet o Extranet) puede ser un sistema de control y seguimiento de inventarios, o un sistema educación a distancia, etc. etc., proveyendo funcionalidad que está más cercana a una implementación cliente/servidor tradicional que a un sitio Web estático. Sin embargo, todas las características potenciales de los tipos antes discutidos pueden estar incorporadas a la aplicación.

Como se observa de lo anterior, los artefactos Web se pueden tornar muy complejos. Y al igual que el desarrollo de productos de software tradicional, el desarrollo y evolución de aplicaciones centradas en la Web puede ser un proceso con muchos desafíos, máxime si se emplean estrategias ad hoc.

2.3 Necesidad de Modelos de Proceso de Desarrollo de Artefactos Web

Como indicamos previamente, la mayoría de las estrategias de procesos de desarrollo en la Web son ad-hoc, y las estrategias de evaluación y control de calidad están más bien ausentes o son más bien intuitivas. Por lo tanto, modelos de proceso de producto bien definidos que promuevan un ciclo de vida de desarrollo eficiente, y modelos de proceso de evaluación con estrategias integrales y cuantitativas que promuevan la comprensión y la mejora de la calidad de artefactos, deben ser utilizados en la práctica cotidiana de proyectos Web de mediana y gran escala, dado el rápido crecimiento e integración de sistemas de documentación, publicación y de funcionalidad de aplicación Web.

Por una parte, es importante tomar conciencia de la necesidad de estrategias más disciplinadas, sistemáticas y justificables para el proceso de desarrollo de aplicaciones Web, y comprender que dichos desarrollos no son sólo un problema de diseño gráfico, interface de usuario, y organización de estructura y contenido, sino que existe la

necesidad de integrar una creciente y, en varios casos, compleja funcionalidad de software. Así, los proyectos de desarrollo centrados en la Web se están tornando más semejantes a los proyectos de software antes que a un trabajo de autoría y arte.

Por lo tanto, se tiene la necesidad de contar con un modelo de proceso de software adaptado a los requerimientos de desarrollo de artefactos Web. Un modelo de proceso para las aplicaciones Web deberá considerar básicamente las fases y actividades; los métodos, herramientas y recursos humanos; guiar a los desarrolladores a especificar las complejidades del sitio Web; deberá minimizar los riesgos de desarrollo y ulterior mantenimiento (tratar con la probabilidad de cambio); deberá permitir ciclos rápidos de generación de artefactos y proveer retroalimentación al administrador del proyecto, entre otros aspectos. Si bien no es el objetivo de esta tesis postular y discutir un modelo de proceso de desarrollo de artefactos Web, en la literatura relacionada [Lowe et al 98, 99; Olsina 98a,b; Powell et al 98] se sugiere distintas alternativas para distintos tipos de proyectos y grados de madurez organizacional para proyectos de Hipermedia y Web. Por ejemplo, según los autores, se podrían adaptar modelos de proceso como el de cascada modificado (con introducción de concurrencia e iteración en las actividades de definición del problema y análisis/especificación de requerimientos), el modelo de proceso en espiral [Boehm 88, Lowe et al 98], o el modelo de proceso flexible de Hipermedia adaptado a la Web, con ciclos de prototipación rápida-funcional y evolutiva [Olsina 97b].

Por otra parte, es necesario tomar conciencia que si bien la Ingeniería en la Web involucra a la modelización y desarrollo de funcionalidad de software y adopta (y deberá adoptar) principios y prácticas establecidas de la Ingeniería de Software, hay características propias en los desarrollos y artefactos Web que deben ser debidamente puntualizados. Varios autores afirman que, en tanto los sitios Web pueden ser considerados software (según vimos en la sección 2.2), los desafíos que abarca la Ingeniería en la Web no es solamente los que enfrenta la Ingeniería de Software [Deshpande et al 99; Lowe et al 99, Powell et al 98]. Los sitios Web y su desarrollo tienen algunos aspectos que lo hacen y harán distintos del software tradicional, a saber:

- ✓ La mayoría de los sitios Web seguirán teniendo la característica de orientación a la documentación: en lo previsible no hay razón para pensar que se abandone esta característica básica surgida desde los inicios de la Web. Las páginas se generan estática o dinámicamente.
- ✓ Las aplicaciones Web continuarán focalizadas en la interface de usuario y en su apariencia y estética (look & feel), favoreciendo la creatividad visual y la incorporación de multimedia. Muchos sitios están dirigidos por objetivos de presentación, promoción de productos y servicios, y metas de negocios; esto parece ir en incremento.
- ✓ La mayoría de los sitios Web seguirán siendo orientados al contenido. Esta característica se observa en muchas aplicaciones de software que ofrecen

documentación y ayuda en línea; sin embargo el alcance de la incorporación de contenido en un sitio Web es mucho más abarcativo. Además, el proceso de autoría es generalmente propio de desarrollos Web o CD-ROMs multimediales.

- ✓ Un sitio Web puede atender una multiplicidad de perfiles de usuario y sus distintos requerimientos. El construir un sitio orientado a audiencias no siempre es una tarea sencilla.
- ✓ El medio donde corren las aplicaciones basadas en la Web es generalmente más impredecible que el medio donde se ejecutan las aplicaciones de software tradicionales. Esto puede afectar la percepción y aceptabilidad que tiene el usuario del artefacto.
- ✓ Un sitio Web implica un mayor acercamiento entre arte y ciencia que lo que puede implicar un artefacto de software
- ✓ En general un proyecto Web de mediana y gran escala requiere mayor multiplicidad de tipos y niveles de habilidades (y roles) de los participantes que un proyecto de software.
- ✓ La Web representa un nuevo paradigma en la distribución de documentación y software.
- ✓ La mayoría de los desarrollos en la Web requieren estar operativos en un corto lapso de tiempo. Esto puede dificultar la aplicación del mismo nivel de rigor formal o semiformal de varias actividades del proceso que en un proyecto tradicional

Lo anterior representa las principales características distintivas de un proyecto de desarrollo en la Web.

2.4 Necesidad de Evaluar para Comprender y Mejorar

Muchas características y atributos de calidad de artefactos Web tales como usabilidad, navegabilidad, seguridad, características de búsquedas y recuperación de información, mantenibilidad, performance, accesibilidad, no-deficiencia, y aspectos de funcionalidades específicas del dominio, por citar algunas, deben ser sistemáticamente tenidas en cuenta durante el proceso de desarrollo de modo de garantizar la calidad en consideración de la satisfacción de los requerimientos de los diferentes perfiles de usuario.

Por lo tanto, para poder obtener artefactos de calidad a la misma se la debe planificar, programar y controlar. Así, la calidad no podrá ser agregada a los artefactos Web al final del proceso de desarrollo sino que, por el contrario, se necesitará considerarla durante todo el ciclo de vida de desarrollo: *la calidad debe ser planificada*. La calidad es un resultado del proceso, y debe ser un factor directriz del mismo.

Una de las metas principales en el desarrollo de aplicaciones Web (como en cualquier otro tipo de desarrollo de software) es producir artefactos de calidad, los que deben

estar regidos por un conjunto de características y atributos deseados, observables y cuantificables. Debemos asegurar los mecanismos por medio de los cuales podamos construir productos que cumplimenten tales características y atributos a partir del planteo de un conjunto de requerimientos de calidad. Por medio de la evaluación de la calidad de artefactos Web, podemos comprender el grado de cumplimiento de un conjunto de características, subcaracterísticas y atributos con respecto a los requerimientos de calidad establecidos.

Para tal fin, se debe prever dentro del modelo de ciclo de vida de desarrollo de artefactos Web, actividades, metodologías y procedimientos para el aseguramiento y control de calidad. Las actividades y métodos a su vez deben estar comprendidas en el marco de un modelo de proceso de evaluación, como veremos en el capítulo 11. Estos modelos de proceso de evaluación que discutiremos no prescriben ni recomiendan metodologías, métodos ni procedimientos específicos para realizar las actividades sino que representan un marco conceptual (y normativo) genérico, en donde los distintos métodos, procedimientos y herramientas se puedan aplicar.

Pero la meta de un proyecto de evaluación, no está solamente ligada a un proceso de desarrollo de artefactos Web. El objetivo podría consistir en evaluar el estado actual de un artefacto en la fase operativa, o comparar la calidad de dos o más artefactos semejantes. Por ejemplo, si la meta de evaluación consistiera en la determinación de la calidad actual de un artefacto Web en la fase operativa, el resultado del proceso podría arrojar recomendaciones para introducir mejoras, y la decisión final podría consistir en lanzar una nueva versión del sitio o en modificar algunos pocos atributos ausentes o pobremente diseñados.

No obstante, como se extrae de lo discutido, para especificar el árbol de requerimientos de características y atributos de calidad, para determinar los criterios de medición elementales e implementarlos, para realizar la agregación apropiada de manera de producir indicadores parciales y globales, se requiere una metodología integral. Es decir, un conjunto bien definido y cooperante de estrategias, métodos, modelos y herramientas que, aplicados sistemáticamente a las distintas actividades del proceso, produzca los indicadores o resultados finales. Por lo tanto, se necesita una metodología que cubra las principales actividades del modelo de proceso de evaluación. Este es el tema central del presente trabajo.

Finalmente, es oportuno recordar estos pares de máximas o principios:

- ✓ *“No se puede medir lo que no se puede comprender”*
- ✓ *“Si no se sabe dónde se está parado, un mapa no ayuda”*
- ✓ *“Lo que no se puede medir no se puede controlar sistemáticamente”*
- ✓ *“No podemos mejorar algo a menos que podamos evaluarlo”*

Capítulo 3

Problemática en la Evaluación de la Calidad de Artefactos de Software (y Web).

3.1 Presentación del Problema

Antes de pasar a discutir aspectos claves en la evaluación de la calidad de artefactos, particularmente para el dominio de aplicaciones centrados en la Web, veamos cuál es la problemática general para este tipo de actividades y enfoques, expresadas primero en algunas preguntas, y luego, con comentarios, haciendo referencia en algunos casos a investigadores y a organizaciones internacionalmente reconocidas.

En pocas palabras, se puede formular las siguientes cuestiones centrales:

- ✓ ¿La evaluación de la calidad del producto, está más bien basada en métricas y modelos cuantitativos de calidad, o está librada a la intuición y subjetividad del evaluador?
- ✓ ¿Las decisiones y procedimientos fundamentales de la evaluación, comparación y ranquin, están centradas en la intuición de los evaluadores, o están cimentadas en modelos de estructuración de características y atributos, y, además, en modelos y métodos cuantitativos de cálculo de puntaje?
- ✓ ¿Las decisiones de la evaluación están centradas y dirigidas por el juicio del usuario, o más bien guiadas e integradas por el juicio de expertos humanos?
- ✓ ¿El sistema o artefacto puede ser considerado simple o complejo, desde el punto de vista de la evaluación?

Y algunas afirmaciones se pueden indicar, como una interesante combinación de las preguntas anteriores, a saber:

- ✓ La evaluación está basada en modelos y métricas de calidad (a partir de características y atributos) pero las decisiones están libradas a la intuición antes que a modelos y métodos cuantitativos de puntaje.
- ✓ La evaluación está basada en métricas y modelos de calidad y las decisiones y procedimientos están centrados en modelos y métodos cuantitativos.
- ✓ El proceso de evaluación se basa en modelos y métodos reconocidos y se centra en el juicio surgido de la experticia humana.
- ✓ El artefacto consiste, desde el punto de vista de la evaluación, de un buen número de subsistemas o componentes, los que a su vez se descomponen en varios elementos. Además, existen diferentes relaciones entre elementos y subsistemas.

Y aquí surge esta otra pregunta:

- ✓ Para evaluar la calidad global de un artefacto, para sistemas complejos como la Web, ¿ es conveniente basarse en un modelo de agregación de atributos y de puntaje meramente lineal y aditivo, o se requiere algo más robusto y flexible?

Pero además de las preguntas centrales antes formuladas, hay otras no menos importantes para el éxito de un proyecto de evaluación, a saber:

- ✓ Para que el proceso de evaluación sea claro y efectivo, ¿no debe focalizarse en metas específicas, en el contexto de una organización?
- ✓ ¿No debe establecerse claramente cual será el perfil del usuario y, consecuentemente, sus requerimientos?
- ✓ Además, ¿cuál es el rol y el beneficio de aplicar un modelo de proceso bien definido para la evaluación de artefactos centrados en la Web?

Para entrar en contexto, utilizaremos algunas declaraciones de investigadores y de estándares de organizaciones autorizadas, acerca de aspectos antes mencionados, a saber:

“La evaluación de un sistema es un proceso de decisión lógica ... la meta básica de un sistema de evaluación está en determinar el grado en que un sistema cumplimenta un conjunto específico de requerimientos” [Dujmovic et al 82]

“Definir a la calidad de software para un sistema es equivalente a definir una lista de atributos de calidad de software requeridos por ese sistema” [IEEE Std 1061].

“Para medir atributos de calidad de software se debe identificar un conjunto apropiado de métricas” [IEEE Std 1061].

“Las métricas de software miden atributos específicos de un producto de software o un proceso de desarrollo de software” [Grady 94], o también, atributos de un recurso.

Por otra parte, la Organización de Estándares Internacionales ISO, en trabajo conjunto con IEC, definen seis características de muy alto nivel que describen, con un mínimo de solapamiento, a la calidad del producto y que son, a saber: usabilidad, funcionalidad, confiabilidad, eficiencia, portabilidad, y mantenibilidad. *“Esas características proveen una línea base para ulteriores refinamientos y descripciones de la calidad del software” [ISO/IEC 9126].*

De manera que, *“muchos atributos de interés en Ingeniería de Software no son directamente mensurables. Esta situación nos fuerza a usar vectores de medidas con reglas para combinar los elementos del vector en una medida más grande e indirecta”*

[Fenton et al 97]

Esta misma situación es revelada en la publicación reciente del estándar del Proceso para Evaluadores de la ISO [ISO/IEC 14598-5], que declara: *“La calidad de los productos de software puede ser descripta en términos de características de calidad como las definidas en ISO/IEC 9126. No obstante, el estado del arte en mediciones de software es tal que, en general, no son prácticas las mediciones directas de esas características. Lo que sí es posible, es valorar esas características sobre la base de medidas de atributos del producto de menor nivel de abstracción”*.

Y luego, agrega: *“En este contexto, el/la evaluador/a puede usar su experiencia en Ingeniería de Software para realizar la valoración”* [ISO/IEC 14598-5].

Como idea central podemos extraer que, el estándar de características de calidad de productos, provee un marco conceptual muy importante al prescribir dichas características a un alto nivel de abstracción, pero si bien sugiere subcaracterísticas de un modo informativo, *“... no provee [prescribe] subcaracterísticas y métricas ni métodos para la medición del puntaje y evaluación”* [ISO/IEC 9126]¹. Es primeramente responsabilidad del evaluador y del usuario solicitante acordar todos los requerimientos (esto es, expresarlo en una lista de características y atributos cuantificables).

Por otra parte, como indicamos al inicio, la evaluación de sistemas complejos es primeramente un problema de decisión. Consecuentemente, las estrategias y técnicas para su solución deben contener componentes subjetivos. Esto se afirma en las siguientes sentencias: *“El uso de métricas de software reduce la subjetividad en la evaluación de la calidad de software al proveer una base cuantitativa para tomar decisiones acerca de la calidad del software”* [IEEE Std 1061].

“No obstante, el uso de métricas de software no elimina la necesidad del juicio humano en la evaluación del software” [IEEE Std 1061].

Por ejemplo, para especificar un conjunto de metas a ser alcanzado por el proceso de evaluación, y, a partir de éstas, especificar requerimientos de calidad, los componentes de subjetividad en la toma de decisión son partes necesarias para iniciar el proceso. Sin embargo, en un sentido positivo, subjetividad significa basado en la experiencia humana. *“ Si la necesaria experticia está ausente, las metas y requerimientos pueden ser ambiguos, ausentes, o aún peor, y no se puede esperar que el proceso de evaluación genere resultados correctos y útiles”* [Dujmovic et al 82] (por más que el proceso, los modelos y los métodos empleados sean los correctos).

¹ Es importante indicar que el nuevo estándar ISO/IEC 9126-1, que está en modo borrador al momento de esta escritura, prescribe subcaracterísticas al segundo nivel de descomposición e introduce un conjunto sugerido de métricas internas, externas y de calidad de uso, como veremos en la sección 5.4.2.1

Por otra parte, y como observaciones finales, podemos citar el principio de Metas Borrosas de Gilb, que dice que: *“Los proyectos sin metas claras no alcanzarán sus metas claramente”* [Gilb 88], y, por otra parte, *“De acuerdo a muchos estudios realizados sobre la aplicación de métricas y modelos en ambientes industriales, las métricas para ser efectivas deben estar: 1) Focalizadas en metas específicas; 2) Aplicarlas a todo el ciclo de vida de productos, procesos y recursos; 3) Interpretarlas sobre la base de las características y comprensión del contexto organizacional, entorno y metas”* [Basili et al. 94]

Finalmente, es importante tener en cuenta que en lo que respecta a productos de software, su utilización y su “consumo” se están tornando cada vez más vitales para la industria y los servicios actuales, y, asociado a esto los muy mentados conceptos de satisfacción del usuario, calidad total, entre otros. Por lo tanto, es necesario evaluar la calidad de los productos de software, para comprender, controlar, mejorar y predecir. En este contexto, la industria del software en el dominio de la Web, es bastante reciente, y su aceptabilidad e integración a sistemas legados, es cada vez mayor y necesario. Además, la Web plantea sus propios desafíos a la Ingeniería de Software desde diversos ángulos: del artefacto, del tipo de estrategia de desarrollo, de la distribución y aceptación, del reuso etc., y desafíos a otras ciencias con sus aspectos legales, de estrategias de comercio electrónico, etc. Consecuentemente, urge saber dónde estamos y hacia dónde vamos. Solamente de este modo podremos planificar para permanecer, cambiar y mejorar.

Si bien la evaluación de sistemas de software tiene más de tres décadas, la mayoría de las técnicas y prácticas actuales para desarrollar software y evaluar en la Web son ad-hoc (como discutimos en el capítulo 2). Quizá, el mayor énfasis puesto en la evaluación y monitoreo del uso de artefactos y recursos en la Web por parte de los usuarios, surge del estudio de registros de entrada (log files) y de seguimiento de enlaces ([Zeff et al 97], Cap. 4); estudios de demografía (para predecir tendencias y comportamientos) [GVU 99]; carga y empleo de recursos en servidores, etc. No obstante, se requiere una visión más amplia en la concepción de la calidad de un artefacto Web, conforme a las necesidades y deseos explícitos e implícitos de los usuarios. Para ello aportaremos nuestras ideas e integraremos varias de las expresadas anteriormente, a la metodología propuesta, como discutiremos en las próximas secciones y a lo largo de esta tesis.

3.1.1 Subjetividad y Objetividad en los Sistemas de Evaluación Complejos

Este es un punto que puede suscitar discusiones, y es importante tomar conciencia de algunos componentes que pueden intervenir en los argumentos de la discusión. Tal es el objetivo de esta sección.

Partamos de la meta que lo que se quiere es evaluar y comparar la calidad global de un

artefacto Web, como lo es un sitio. Este artefacto (o producto) no sólo posee las propiedades de estar centrado en la documentación como sitio de publicación y presentación sino que también brinda funcionalidad de aplicación de software (acceso a bases de datos, motores de búsqueda y recuperación, procedimientos de cálculos por medio de applets, etc.). Por lo tanto, consideramos que el sistema en estudio es complejo desde el punto de vista de la evaluación, ya que a la calidad se la expresa en función de varias características (en consideración de un perfil de usuario), y que las mismas se descomponen recursivamente en subcaracterísticas y atributos.

Entonces, ¿cuáles son al menos las principales fases o módulos del proceso de evaluación de un artefacto Web? Se puede argumentar los siguientes: (1) Definición y especificación de requerimientos de calidad; (2) Definición e implementación de la evaluación elemental; (3) Definición e implementación de la evaluación global; (4) Análisis y conclusiones (las fases y principales actividades serán discutidas en la sección 4.2 y en los siguientes capítulos)

Por ejemplo, ¿qué podemos decir de la subjetividad u objetividad en la definición y especificación de requerimientos de calidad?; o ¿qué implica objetividad o subjetividad en la definición e implementación de la evaluación elemental?, ¿y de la global?

En principio, dos conceptos extremos (aunque no excluyentes en la práctica) entran en juego: objetividad y subjetividad; y que, en el contexto de la evaluación (y sin entrar en detalles epistemológicos) estamos queriendo significar al proceso de decisión basado en modelos y métodos cuantitativos, o al proceso de decisión basado en la intuición². Además, cuando hablamos de intuición, no lo decimos en un sentido peyorativo sino que por el contrario queremos significar al mecanismo mental, o conocimiento claro, íntimo o instantáneo, surgido de la experticia. Con todo, es importante resaltar que la evaluación intuitiva es un ejercicio mental del día a día que todo ser humano realiza, ya sea más o menos experto en el dominio del objeto que evalúa. (Y de hecho, la evaluación intuitiva representa la base de muchas de las decisiones tomadas por el ser humano). Por otra parte, más de un investigador ha afirmado que la métrica de un atributo tiene que estar bien definida, esto es, basada en la observación empírica, consistente con la *intuición* y que represente una apropiada correspondencia entre el dominio empírico y el sistema numérico [Fenton 91, Goldberg et al 95].

Además, e intentando responder a la pregunta antes formulada, el proceso de evaluación siempre se basa en un conjunto de requerimientos que el artefacto (en nuestro caso el sitio Web), debe satisfacer. Estos requerimientos generalmente se derivan de un conjunto de metas dadas y/o elicitadas para el objeto de evaluación. Obviamente, el conjunto de metas sólo puede ser elicitado y definido por los tomadores

² “Percepción o conocimiento claro, íntimo o instantáneo de una idea o verdad sin el auxilio de la razón”, Diccionario Sopena Arg.; “Conocimiento inmediato de una cosa, idea o verdad, sin el concurso de razonamientos” Enciclopedia Microsoft™ Encarta 98

de decisión (esto es, seres humanos). Por lo tanto, ya las metas como los requerimientos correspondientes se especifican más bien de un modo subjetivo; según el pensamiento de varios investigadores no hay otro modo “más objetivo o racional” para iniciar un proceso de evaluación, comparación y selección (interpretando al concepto “subjetivo” en un sentido positivo, como mencionamos anteriormente). De manera que los componentes subjetivos no pueden ser obviados, pero deben ser minimizados y/o usados apropiadamente.

Por otra parte, e intentando responder a la pregunta previa: ¿qué implica objetividad o subjetividad en la definición e implementación de la evaluación elemental?, Podemos decir, por ejemplo, que a partir de la especificación de requerimientos, muchos de los atributos podrán medirse elementalmente de un modo objetivo, para determinar el grado o porcentaje de satisfacción de cada requerimiento. Este es el caso para saber si un artefacto Web tiene mapa de sitio, o para calcular el porcentaje de páginas que no sobrepasan un tamaño de página umbral (y que en consecuencia, la performance para cargarla es satisfactoria); o el nivel de enlaces rotos, etc. En el primer caso no sólo se puede especificar y consensuar un criterio fácilmente observable (el mapa está o no está presente), sino que en los restantes, se puede objetivamente automatizar la recolección de datos y modelar el criterio mediante una fórmula matemática. No obstante, dentro de los requerimientos algunos atributos sólo podrán comprobarse de un modo subjetivo³, a partir del juicio de evaluadores expertos (y en lo posible más de dos). Por ejemplo, para el atributo “*Uniformidad en el Estilo Global*” del sitio, puede ser conveniente medirlo mediante un criterio de preferencia directa y subjetiva. Realizarlo de otro modo, podría implicar mayores costos dada la relativa baja prioridad de ese atributo o la alta complejidad para descomponerlo. Sin embargo, dicho atributo es un elemento necesario en la composición para determinar a la calidad global del artefacto.

Finalmente, podemos agregar que una metodología de evaluación exitosa debe proveer una correcta y equilibrada agregación de componentes objetivos y subjetivos. La metodología propuesta que discutiremos en el presente trabajo está basada en métricas y modelos de calidad, derivados de las características de alto nivel prescriptas en el estándar ISO/IEC 9126 y personalizadas al dominio de aplicaciones Web, en tanto que las decisiones y procedimientos de evaluación elemental y global están centrados en modelos y métodos de puntaje cuantitativos.

3.1.2 Enfoques Cuantitativos para la Evaluación de Calidad

En este punto es importante resaltar que, frecuentemente, se utilizan dos conceptos

³ La afirmación “*sólo podrán comprobarse de un modo subjetivo*” debe interpretarse no de un modo categórico, sino relativo al contexto, costos, relevancia, precisión y criticidad del atributo. Generalmente, un atributo medido con un criterio de preferencia directa y subjetivo, se lo podría descomponer, en otros objetivos, más alguno subjetivo. Piense el lector, por ejemplo, en evaluar, comparar y seleccionar una plataforma o sistema operativo, en donde el atributo “calidad de la documentación” conforma uno entre otros quinientos dentro del árbol de requerimientos [Dujmovic et al 97]

relacionados a la categoría o tipo de evaluación, en consideración de sus estrategias, métodos y técnicas. Estamos hablando de la *evaluación cualitativa* o de la *evaluación cuantitativa*.

Como indica Dujmovic [Dujmovic 96], entre otros autores, los métodos y técnicas de evaluación cualitativa, se basan generalmente en una lista de características a ser analizadas para un producto o productos competitivos. La lista puede contener características de distinto tipo (relacionadas al ente a evaluar como producto, proceso o recurso), tales como características técnicas, de costo, etc. Luego del proceso de evaluación, en la etapa de análisis y conclusiones, los tomadores de decisión frecuentemente crean para cada sistema (a comparar) una lista de ventajas y desventajas. Finalmente, mediante un mecanismo intuitivo, comparando las ventajas y desventajas para cada característica de cada sistema, arriban a un ranquin final. Este enfoque es obviamente conveniente y atractivo cuando el objeto de la evaluación y el proceso de decisión es suficientemente simple. Sin embargo, en un proceso de evaluación, comparación y selección en donde intervienen, por ejemplo, más de cuarenta características y atributos para cada sistema seleccionado, y en donde se identifican distintas relaciones entre los mismos, un enfoque como el anterior carece de las propiedades de precisión y justificación objetiva necesarias.

Esta dificultad puede minimizarse mediante el uso de un enfoque cuantitativo. La meta de una metodología, método o técnica cuantitativa es proveer un proceso de evaluación flexible, bien estructurado, suficientemente preciso, y basado en principios ingenieriles de manera de proveer indicadores cuantitativos elementales, parciales y globales los cuales son usados como base y justificación de las decisiones más óptimas.

3.1.2.1 Modelos de Agregación y Puntaje. Al inicio de este capítulo, cuando mediante preguntas y comentarios, presentamos la problemática general para el proceso de evaluación, dijimos:

Las decisiones y procedimientos fundamentales de la evaluación, comparación y ranquin, ¿están centradas en la intuición de los evaluadores, o están cimentadas en modelos de estructuración de características y atributos, y además, en modelos y métodos cuantitativos de cálculo de puntaje?

En esta sección, presentaremos dos enfoques cuantitativos de utilidad en la actividades de estructuración de atributos y determinación de puntajes. Particularmente, discutiremos modelos de puntaje aditivos y lineales versus modelos de decisión multiatributos no-lineales.

Respecto de la idea que sustenta a los modelos de puntaje aditivos y lineales es bastante simple. Para cada característica o sistema a evaluar y comparar se identifican n atributos necesarios, cuya preferencia o indicador elemental (IE) se debe computar.

Supongamos que los valores individuales de $IE_1 \dots IE_n$ están normalizados de manera que: $0 \leq IE_i \leq 1$; o, en la escala de porcentaje $0 \% \leq IE_i \leq 100\%$.

En el caso que todos los atributos intervinientes sean equi-pesados, podríamos expresar el indicador o preferencia global (IG) mediante el uso de una sumatoria. Pero si los elementos no tienen la misma importancia en el modelo de estructuración, debemos introducir la idea de pesos positivos y normalizados, de manera que: $0 \leq P_i \leq 1$, para $i = 1 \dots n$ y, $P_1 + \dots + P_n = 1$. Por lo tanto, el puntaje o indicador global puede ser expresado mediante la siguiente expresión:

$$IG = P_1 IE_1 + \dots + P_n IE_n \quad \text{para: } 0 \leq IE_i \leq 1$$

o, indicado de un modo más conversacional:

$$\text{Puntaje Global} = \sum (\text{componente Peso} \times \text{componente Preferencia o Indicador Elemental})$$

Este enfoque puede ser adecuado para casos más simples, en donde la cantidad de atributos es suficientemente baja de modo que los pesos tengan mayor relevancia. Considere el lector, que si la cantidad de atributos es 100 el peso más bajo es casi irrelevante, en tanto que el componente peso en promedio es de 0,01 o 1% para cada indicador elemental. Esto le quita sensibilidad, flexibilidad y robustez al modelo, principalmente para evaluar sistemas de mediana o alta complejidad.

Dujmovic en sus investigaciones [Dujmovic 96] identificó al menos siete inconvenientes en aplicar modelos de puntaje aditivos [Gilb 76]. Es importante enunciar dichas desventajas (algunas contenidas en dicho trabajo), para considerarlas dentro de la problemática en la evaluación de calidad de artefactos Web complejos, a saber:

- 1- *Número limitado de componentes para la evaluación.* Si la evaluación incluye n componentes entonces el peso promedio de cada componente es $100/n \%$. Esto limita el número de componentes a ser evaluados debido a la sensibilidad de los pesos.
- 2- *Malgasto en el esfuerzo de la evaluación.* El esfuerzo total de la evaluación depende de la cantidad de componentes y de la complejidad de sus relaciones. La evaluación de características o atributos que tienen un efecto ínfimo en el resultado final implica malgastar recursos.
- 3- *Imposibilidad de modelar requerimientos obligatorios.* Los modelos aditivos no soportan requerimientos obligatorios dado que la ausencia de una característica necesaria no disminuye el puntaje parcial o global más que por el valor relativo del peso de la característica. Sin embargo, muchas veces esa ausencia no puede ser compensada.
- 4- *Imposibilidad de modelar requerimientos simultáneos.* Los modelos aditivos lineales no son apropiados para expresar relaciones de simultaneidad entre

características y/o atributos. La aditividad asume que la presencia insuficiente de un atributo se puede siempre compensar por la suficiente presencia de *cualquier* otro atributo.

- 5- *Imposibilidad de modelar requerimientos suficientes.* Los modelos aditivos lineales no pueden modelar reemplazabilidad de atributos.
- 6- *Imposibilidad de modelar relaciones lógicas asimétricas .* Dichos modelos usan los pesos como el único medio de diferenciar la relativa importancia de las entradas. Esto no es suficiente para expresar relaciones lógicas asimétricas tales como combinaciones de características obligatorias, deseables y opcionales; o, combinación de relaciones necesarias y suficientes.
- 7- *Desequilibrio en el esfuerzo de evaluación.* La evaluación de sistemas complejos implica un esfuerzo considerable. Un nivel alto de experticia se puede necesitar en este proceso. Por lo tanto, no es razonable producir costosas y precisas entradas a un modelo de decisión con baja sensibilidad

Por ejemplo, resolver el problema de los requerimientos obligatorios no se puede satisfacer por medio del mero uso de una media geométrica. Por ello, el modelo de lógica de preferencia de puntajes (LSP), como una extensión del modelo aditivo y lineal, viene a resolver dichos inconvenientes, argumenta Dujmovic. Se puede definir al modelo LSP como uno de agregación lógica de preferencias centrado en medias de potencia pesada. Permite modelar relaciones de reemplazabilidad, neutralidad, y simultaneidad entre atributos y características. Se puede utilizar operadores de preferencia lógica para modelar distintos nivel de intensidad de polarización “y/o”, entre otros aspectos. En el capítulo 7, introduciremos y utilizaremos este modelo para ejemplificar la obtención de la preferencia de calidad global de los sitios académicos seleccionados para el caso de estudio.

3.2 Estado del Arte en la Evaluación de Artefactos Web

Finalmente, ¿cuál es la situación actual en las estrategias de evaluación, principalmente de artefactos de información centrados en la Web e Hipermedia? ¿Se utiliza enfoques ingenieriles o estrategias ad hoc? ¿Se utiliza heurísticas y criterios cualitativos o se emplea criterios, modelos y métodos cuantitativos? ¿La evaluación está centrada en cuestionarios llenados a partir de la intuición y experticia de los usuarios; está centrada en modelos de diseño y guiada por expertos; o está centrada en modelos de calidad de productos, en modelos de agregación cuantitativos y centradas en expertos humanos?

Como mencionamos en la introducción del capítulo 1, si bien la evaluación de sistemas de software tiene más de tres décadas [Gilb 69, McCall et al 77, Miller 77, White et al 63], la mayoría de las técnicas y prácticas actuales para la evaluación, comparación y ranquin de artefactos Web complejos son ad-hoc o están emergiendo. No existe hasta el momento (inicio del 1999) una metodología amplia y reconocida, cuantitativa y

objetiva, para evaluar artefactos y sistemas Web complejos. Como dijimos anteriormente, tal vez el mayor énfasis se ha puesto en la evaluación y monitoreo del uso de los visitantes a los sitios Web. Para ello se examinan registros de entrada (log files), número de visitas (hits), registros de seguimiento de enlaces (para predecir tendencias y comportamientos), carga de recursos en servidores, etc. (ver, por ejemplo, una experiencia en [Yu et al 98]), o los así llamados estudios demográficos (ver [GVU 99] y ediciones previas). Sin embargo, nosotros propendemos a una visión de la calidad de un artefacto Web, más amplia e integral, conforme a las necesidades y deseos explícitos e implícitos de un perfil de usuario dado [Olsina 99, Olsina et al 99c].

Por una parte, a mediados de la década de los noventa, Garzotto et al. emplearon métodos y heurísticas de evaluación cualitativa en el campo de Hipermedia. Además, recientemente, Paolini y Garzotto han estado avanzando en el proyecto SUE (Systematic Usability Evaluation) en el Laboratorio de Hipermedia del Politécnico de Milán. En uno de los primeros artículos de sus investigaciones describen un método de evaluación orientado al diseño [Garzotto et al 95]. Básicamente, se podría describir al mismo como una evaluación heurística, cualitativa, centrada en el análisis del artefacto de hipermedia a partir de un modelo de diseño apropiado (como HDM [Garzotto et al 91, 93]) empleando, además, criterios generales de usabilidad, aunque no métricas cuantificables. Los autores identifican cinco dimensiones para evaluar aplicaciones de Hypermedia, a saber:

- ✓ *Contenido*, relacionado a los objetos de información incluidos en la aplicación
- ✓ *Estructura*, la organización del contenido
- ✓ *Presentación*, cómo se muestra a los lectores el contenido y las funciones de la aplicación
- ✓ *Dinámica*, cómo los usuarios se mueven e interactúan con los elementos y componentes de información
- ✓ *Interacción*, cómo se usan los componentes dinámicos

Entonces, ellos argumentan que para un análisis sistemático de la aplicación de hipermedia, se necesita un modelo de diseño con el cual se pueda describir a la misma (recordar que por ese tiempo además del modelo HDM, otros modelos y métodos como OOHDM [Rossi 96, Schwabe et al 95a y b], RMM [Isakowitz et al 95] ya eran reconocidos en la comunidad de Hipermedia). Una vez descrita la aplicación en base a un modelo, los autores le aplican varios criterios de evaluación o heurísticas, como por ejemplo:

- ✓ *Riqueza*, que representa la abundancia de ítems de información y el modo de alcanzarlos
- ✓ *Facilidad*, que evalúa la accesibilidad a la información y cuán fácil de operar y comprender son las operaciones
- ✓ *Consistencia*, que podría resumirse en (según los autores) elementos

conceptualmente similares (dentro de la aplicación) son tratados de modo similar y elementos conceptualmente diferentes son tratados diferentemente.

- ✓ *Auto-evidencia*, representa cuán bien los usuarios se arman la idea mental del propósito de lo que se les está siendo presentado
- ✓ *Predictibilidad*, expresa cuán bien los usuarios se anticipan al resultado de una operación
- ✓ *Legibilidad*, expresa un sentimiento de la validez de toda la aplicación. (Y dicen: “*readability depends upon all factors mentioned*” [Garzotto et al 95] en pp.75)
- ✓ *Reuso*, considera el uso de objetos y operaciones en diferentes contextos y propósitos.

Sin embargo, este enfoque es sólo conveniente cuando el problema de evaluación es más bien simple e intuitivo. En el caso en que se desee expresar a la aplicación con mayor cantidad de factores y atributos de más bajo nivel se vuelve difícil una evaluación justificable y precisa, en donde se dificulta identificar menores diferencias entre valores similares de atributos de sistemas comparativos. Además, aunque no comentaremos el proyecto SUE (para no desviarnos del foco), este trabajo avanzó en líneas de investigación a partir de estas bases; no obstante, hasta el presente no se ha convertido en una metodología reconocida para evaluar sitios Web.

Por otra parte, y relacionado al dominio de la Web, a mediados de la década de los noventa comenzaron a aparecer guías de estilo y de diseño para la Web para asistir a los desarrolladores en el proceso [IEEE WPG, Rosenfeld et al 98]. Asimismo, surgido de organizaciones reconocidas como W3C, se han prescripto listas de buenas prácticas, con asignación de prioridades, para que el autor o desarrollador las tenga en cuenta a la hora de diseñar sitios de un modo más usable, navegable, eficiente y accesible [W3C 99]. Además, en [Thüring et al 95] los autores exponen un conjunto de criterios y principios que el diseñador de aplicaciones de autoría debe tener en cuenta de modo que ayuden al usuario en el proceso de construcción de su modelo mental, fortaleciendo a la *coherencia* entre los factores positivos y debilitando a la *demora cognitiva* como influencia negativa. Estudios empíricos han demostrado que la capacidad de los lectores para comprender y recordar un texto depende del grado de coherencia del mismo. Investigaciones psicolingüísticas enfatizan la relación existente entre coherencia y procesamiento de información: un documento es coherente si el lector puede construir a partir del mismo un modelo mental que se corresponde a hechos y relaciones del posible mundo. Para construir el modelo se debe tener en cuenta tanto la coherencia a nivel local (a nivel de nodo), como la coherencia a nivel global (en la red de nodos). Para incrementar la coherencia se debe limitar la fragmentación resultante de la segmentación de la información en nodos disjuntos que se muestran en ventanas separadas. Una de las consecuencias de la fragmentación puede resultar en la falta de un contexto interpretativo y entonces parecer que el hiperdocumento es una agregación de piezas de información débilmente acopladas antes que un todo cohesivo.

Por lo tanto, esas guías, principios y criterios son muy útiles, debido a que documentan características, atributos y reglas prácticas a tener en cuenta en el proceso de evaluación de calidad. El empleo apropiado de esas guías y criterios, debe ayudar a mejorar el proceso de diseño y autoría en la Web pero, obviamente, no constituyen por sí mismas un modelo y método de evaluación de artefactos Web.

Finalmente, estudios y surveys de usabilidad y de interfaces son muy bien conocidos en la comunidad de Hipermedia e Interfaces de Usuario [Bevan 95 y 97, Nielsen 93, Schneiderman 98]. Particularmente, Nielsen ha realizado muy interesantes surveys sobre características y atributos en la Web, a partir de 1994, que están en línea en su columna electrónica denominada Alertbox, y que con frecuencia citaremos en la sección 6.4 [Nielsen_Alert].

Con todo, surveys y evaluaciones cuantitativas en dominios específicos de la Web, como comercio electrónico [Lohse et al 98, Tilson et al 98], museos y dominios de sitios académicos están recién emergiendo. En esta dirección, Lohse and Spiller identificaron y midieron 32 atributos para 28 negocios on-line (1996) e identificaron características y atributos de diseño que influyen al tráfico y a las ventas. Los autores identifican seis categorías para un comercio en la Web (Mercancías, Servicios, Promoción, Conveniencia, Confirmación de Venta (checkout), y Navegación) y a partir de esto, describen los atributos que influyen al tráfico y a las ventas. Pero estos estudios estadísticos y descriptivos, aunque reveladores para marcar el estado del arte de la calidad de los atributos y marcar tendencias, no constituyen una metodología amplia para evaluar características, atributos y sus relaciones de sitios Web como un todo.

Finalmente, es importante detenernos un poco y referenciar a las investigaciones de Kirakowski et al [Kirakowski et al 94, 98] quienes han realizado interesantes experimentos para evaluar la usabilidad en función de la satisfacción del usuario. Los autores llaman a su enfoque “subjective user-based evaluation”, o, que se puede traducir como evaluaciones subjetivas basadas en el usuario (y dirigidas por el usuario). La tesis sostenida por Kirakowski et al es que evaluar la satisfacción de los usuarios por medio de cuestionarios administrados en línea en la WWW (o en su versión en papel), puede ser una fuente de información útil no sólo para medir la usabilidad alcanzada en el diseño del sitio, sino también dar información de diagnóstico acerca de cómo mejorar al sitio bajo evaluación.

Para elaborar el cuestionario de “satisfacción del usuario” (1996/97), se basaron en un cuestionario de estudios previos sobre aplicaciones de escritorio (cuestionario SUMI), y consultaron opiniones de diseñadores, usuarios, webmasters, respecto de experiencias positivas y negativas encontradas por ellos cuando visitaban y usaban sitios Web. Dichas declaraciones las analizaron por contenido y las incluyeron en un modelo de cinco factores y sesenta ítems o preguntas. Cada factor está compuesto de doce ítems evaluados en una escala de uno a siete. Al nuevo cuestionario lo denominaron WAMMI

(Web Analysis and Measurement Inventory). Los factores los citamos a seguir:

- ✓ *Atracción*, el grado que a los usuarios les gusta el sitio si lo hallan agradable de usar
- ✓ *Control*, el grado en que los usuarios se sienten comprometidos (“in charge”), ya que el sitio les permite navegar con facilidad, o que el sitio “se comunica” con el usuario acerca de lo que está haciendo.
- ✓ *Eficiencia*, el grado con que los usuarios sienten que el sitio tiene información que están buscando, si funciona a velocidad razonable, y si se adapta al navegador.
- ✓ *Ayuda*, el grado que los usuarios sienten que el sitio les posibilita resolver sus problemas con respecto a encontrar información y explorarla
- ✓ *Aprendizaje*, el grado que los usuarios sienten que se acostumbran al sitio cuando ingresan por primera vez, y el grado que los usuarios sienten que pueden aprender otras facilidades o acceder a otra información una vez que han comenzado a usar el sitio

Los autores convierten la escala de uno a siete en valores numéricos cuyo máximo es 70 y el punto crítico fue colocado artificialmente en 50 de modo que los factores de los sitios están por encima o debajo del valor esperado (en consideración de un desvío estándar de 10).

Si bien este método de evaluación es un avance en el estado del arte (además, está disponible comercialmente), es importante resaltar algunos inconvenientes. El trabajo informado en [Kirakowski et al 98] es un interesante experimento en un “ambiente naturalísticamente controlado”, en donde los usuarios a llenar los cuestionarios (recolectar los datos) fueron especialmente “reclutados”. No obstante, en un ambiente natural, los usuarios (evaluadores) sentados en el sitio Web, antes de emitir su juicio, debieran conocer el sitio y pasar por el cuestionario de 60 preguntas. Esto tiene sus inconvenientes, aún reconocido por los autores: menos del 5% del número de visitantes del sitio llenaban el cuestionario; de estos, muchos debían desecharse por incompletos. Además, reconocen que un usuario promedio no podría completarlo en menos de 20 minutos. (Y estudios de usabilidad realizados por Nielsen demuestran que el usuario es generalmente resistente a llenar cuestionarios, y cuando los llenan estos no deben estar compuestos de muchos ítems y sub-ítems). Evidentemente, la evaluación centrada (y guiada) por los usuarios introduce el ruido o impedancia propio de los enfoques no controlados y guiados por evaluadores expertos. Sin embargo, tiene de atractivo que es efectivo en costos.

Finalmente, podemos indicar los estudios e investigaciones realizados (a partir de 1998), sobre diversos dominios empleando la metodología propuesta para la evaluación cuantitativa de sitios Web. En el próximo capítulo describiremos las motivaciones y los aportes de nuestra metodología al estado del arte en las evaluaciones cuantitativas de la calidad de artefactos Web, centradas en expertos humanos. Asimismo, presentaremos los casos de estudio realizados que utilizaremos para ejemplificar a lo largo del trabajo.

Capítulo 4

Panorama de la Metodología Cuantitativa de Evaluación y Comparación de Calidad de Sitios Web.

4.1 Motivaciones y Aportes

En el capítulo 2 discutimos de un modo introductorio, la problemática en el desarrollo de sistemas de información centrados en la Web, los nuevos desafíos que introduce la Web como disciplina a la Ingeniería de Software, y la necesidad de aplicar modelos de ciclo de vida de desarrollo bien definidos y la de evaluar los artefactos para comprenderlos y mejorarlos. En el capítulo siguiente (capítulo 3), nos concentramos en la problemática de la evaluación; discutimos distintas estrategias, y enfoques para evaluar la calidad de artefactos en Hipermedia y en la Web. Además, presentamos el estado del arte de la evaluación en este campo y trabajos relacionados. En las secciones siguientes, y a partir de ahora, mostraremos nuestra propuesta y pasaremos a desarrollarla.

4.1.1 Necesidad de una Metodología Integral, Robusta y Flexible para Evaluar la Calidad de Artefactos Web Complejos

Hasta ahora hemos observado que los enfoques utilizados para la evaluación, no se han concentrado en medir a la calidad como una característica de muy alto nivel que contiene y describe a otras características como *usabilidad*, *funcionalidad*, *eficiencia*, *confiabilidad*, u otras como *mantenibilidad* y *portabilidad*. Antes bien, se especializan en evaluar usabilidad y satisfacción del usuario para cierto perfil, sin que siempre se tenga muy en cuenta esa visión integral de la calidad.

Esta visión de la calidad de artefacto como característica de alto nivel (entre otras características, como costo y productividad), ha sido bien modelada desde hace más de treinta años, y plasmada, por ejemplo, en los estándares de ISO e IEEE. Estos estándares, a partir de las seis características generales (arriba remarcadas) describen, con mínimo solapamiento, a la calidad del software. “*Esas características proveen una línea base para ulteriores refinamientos y descripciones de la calidad del software*” [ISO/IEC 9126].

Si bien en muchos casos prácticos, y a partir de una meta, requerimientos y usuarios específicos, se puede necesitar medir sólo un par de atributos o subcaracterísticas de calidad y sacar conclusiones de ese aspecto del artefacto (como por ejemplo, sus

mecanismos de búsqueda y flexibilidad en la recuperación de la información, o aspectos de navegación, etc.), nos parece importante explorar esa visión integral de la calidad. Consecuentemente, para especificar el árbol de requerimientos de características y atributos de calidad, para determinar los criterios de medición elementales e implementarlos, para realizar la agregación apropiada de manera de producir un indicador global, se requiere una metodología integral. Es decir, un conjunto bien definido y cooperante de estrategias, métodos, modelos y heurísticas que, aplicados sistemáticamente a las distintas actividades del proceso, produzca el indicador esperado. Se requiere una metodología que cubra las principales actividades del proceso de evaluación.

Pero además, la metodología debe ser robusta: esto es, debe permitir procesos con resultados repetibles, reproducibles y objetivos. En cuanto a la característica de repetitibilidad de la evaluación de un mismo producto y para la misma especificación de requerimientos y por los mismos evaluadores implica que debe producir resultados aceptados como idénticos. En cuanto a reproducibilidad de la evaluación de un mismo producto y para la misma especificación de requerimientos y realizada por evaluadores diferentes implica que debe producir resultados aceptados como idénticos (ver [ISO/IEC 14598-5]). En cuanto a objetividad, como vimos en la sección 3.1.1, los resultados deben ser factuales, observables y cuantitativos, no desviados por la opinión parcial o intencional de los evaluadores.

Finalmente, la metodología debe ser flexible. Primero, porque debe permitir agregar o sacar características, subcaracterísticas y atributos de un modo modular (con la idea de cohesión y mínimo solapamiento), conforme a las necesidades específicas del perfil o perfiles de usuario, del dominio Web, y del contexto organizacional. Segundo, porque dependiendo del dominio, criticidad del artefacto y necesidad de precisión, debe permitir ajustar los criterios de evaluación y/o procedimientos de escrutinio. Tercero, debe ser flexible en el sentido que dependiendo de la complejidad en la cantidad de componentes y/o relaciones, se pueda utilizar modelos más sencillos como los meramente aditivos y lineales o más sofisticados como los de criterios multi-atributo y no-lineales (presentados en 3.1.2.1). Finalmente, la metodología debe ser flexible para integrar componentes subjetivos y objetivos de evaluación, guiados por expertos en el dominio. Los componentes subjetivos, en lo posible, deben ser minimizados.

4.1.2 El Enfoque Propuesto: Metodología de Evaluación de Calidad de Sitios Web (o, en su forma abreviada, Web-site QEM)

El enfoque propuesto, es esencialmente integral, flexible y robusto, y cubre la mayor parte de las actividades en el proceso de evaluación, comparación, y selección de artefactos Web.

La estrategia propuesta, denominada Metodología de Evaluación de Calidad de Sitios Web (o, en inglés, Web-site Quality Evaluation Method, o, metodología Web-site QEM), pretende realizar un aporte ingenieril al proponer un enfoque sistemático, disciplinado y cuantitativo que se adecue a la evaluación, comparación y análisis de calidad de sistemas de información centrados en la Web (más o menos complejos).

Como hemos indicado en el capítulo 1 y 2, las aplicaciones de software centradas en la Web son cada vez más complejas y están creciendo rápidamente. Por lo tanto, este rápido crecimiento genera nuevos desafíos, como por ejemplo: cómo diseñar y producir artefactos Web para diferentes dominios, necesidades y comportamientos de usuarios, teniendo a la calidad como objetivo esencial (en consideración de costos); o, desde otro punto de vista, cómo medir, evaluar y potencialmente mejorar la calidad de los sitios Web. (“*Un mapa no sirve si no se sabe dónde se está parado*”)

Una estrategia efectiva para encarar estos desafíos es el modelado de procesos y productos, usando enfoques prescriptivos y descriptivos [Olsina 98d, 98e]. Particularmente, el modelado de productos nos permite comprender, analizar, optimizar, y predecir la calidad de los mismos. Es desde esta perspectiva, la de modelado de procesos y, particularmente, de productos (artefactos Web), que se propone construir y especificar a la metodología.

Si bien la metodología puede ser empleada en cualquier fase del ciclo de vida de los productos (a saber, fase de exploración, fase de desarrollo y fase operativa), las etapas en este estudio se focalizarán primeramente en la fase operativa, esto es, en la evaluación de artefactos Web ya existentes u operativos. No obstante, es también objetivo de esta investigación, emplear sistemáticamente a Web-site QEM, en la evaluación de nuevos proyectos de desarrollo.

Web-site QEM, incluye un conjunto de fases, actividades, productos, modelos y constructores de proceso que introduciremos seguidamente, en la sección 4.2. Una de las metas principales de la evaluación y comparación de calidad de artefactos Web, radica en comprender el grado de cumplimiento de un conjunto de características y subcaracterísticas con respecto a los requerimientos de calidad establecidos. De este modo, otro aporte interesante consiste en la definición de características, subcaracterísticas y atributos cuantificables considerando dominios de aplicaciones Web particulares. Estas investigaciones nos permitirán generalizar subcaracterísticas y atributos para conjuntos de dominios: así, patrones de reuso se pueden catalogar a semejanza de los patrones de diseño [Rossi et al 97; 99]

4.1.3 Necesidad de un Modelo de Proceso de Evaluación

Como indicamos previamente, modelos de proceso de producto bien definidos que

promuevan un ciclo de vida de desarrollo eficiente, y modelos de proceso de evaluación que promuevan la comprensión y la mejora de la calidad de artefactos, no están siendo utilizados en la práctica cotidiana pese al rápido crecimiento de sistemas de publicación, documentación, y de funcionalidad Web. La mayoría de las estrategias son ad hoc.

Si bien en investigaciones previas, hemos avanzado en modelos de proceso de productos de hipermedia [Olsina 98a, 98b], (el lector se puede remitir al Apéndice B, para introducirse en los conceptos de procesos y en un marco conceptual genérico), los modelos de proceso de evaluación de artefactos presentan sus propios desafíos. Estos modelos se pueden aplicar fuera de un proyecto de desarrollo (por ejemplo, evaluar la calidad de un sitio en la fase operativa, para determinar el nivel de calidad actual), como así también se pueden aplicar de un modo integrado al ciclo de vida de desarrollo de productos (en actividades de control y aseguramiento de calidad).

A pesar de que en esta tesis no desarrollaremos y postularemos a un modelo de proceso de evaluación de artefactos Web específico, le dedicaremos un capítulo de este trabajo (el capítulo 11), dada su relevancia para el éxito de cualquier proyecto. Con todo, continuamente usaremos la terminología de modelado de procesos, esto es, conceptos de meta y objetivo, proceso, tarea, actividad, artefacto, documento, agente, rol, recurso, constructor de proceso, entre otros¹.

Ahora, y sin entrar en detalles, ¿cuál es el interés y necesidad de modelar procesos? La respuesta es más bien simple si nos atenemos a esta hipótesis generalizada: comprendiendo, controlando y mejorando a los procesos, también se puede controlar y mejorar a la calidad de los productos.

Desde el punto de vista de evaluación y valoración de procesos, su visibilidad y madurez, la década del noventa ha marcado el auge en marcos conceptuales y estrategias. Sólo basta recordar algunas de ellas como CMM, [Humphrey 89, Paulk et al 93, 94], SPICE [ISO_SPICE], ISO 9001 [ISO 9001] y 9000-3 [ISO 9000-3], y Bootstrap, entre otros.

Particularmente, de interés para nuestras investigaciones es el recientemente editado estándar del proceso para evaluadores [ISO/IEC 14598-5]. Este proceso (que discutiremos en el capítulo 11) representa principalmente la vista funcional y de información de un modelo de proceso, y se apoya en el modelo definido en ISO/IEC 9126, de características de calidad de productos, y, además, es aplicable dentro de los procesos de ciclo de vida definido en ISO/IEC 12207.

¹ El lector cuenta al final con un glosario, el cual abarca una buena cantidad de conceptos de modelo de procesos y, además, conceptos de sistemas y métodos de evaluación en general y en la Web

4.2 Panorama de las principales fases, procesos y modelos de la Metodología Web-site QEM

En esta sección describiremos, para la metodología Web-site QEM, las principales fases, actividades, modelos, y algunos constructores intervinientes en el proceso de evaluación, comparación y ranquin de calidad. La figura 4.1 muestra una vista general de las fases de la metodología y de los principales pasos y constructores de proceso. Estas fases son, a saber:

- ✓ *Planificación y Programación de la Evaluación de Calidad*
- ✓ *Definición y Especificación de Requerimientos de Calidad*
- ✓ *Definición e Implementación de la Evaluación Elemental*
- ✓ *Definición e Implementación de la Evaluación Global*
- ✓ *Análisis de Resultados, Conclusión y Documentación*
- ✓ *Validación de Métricas* (no mostrada en la figura)

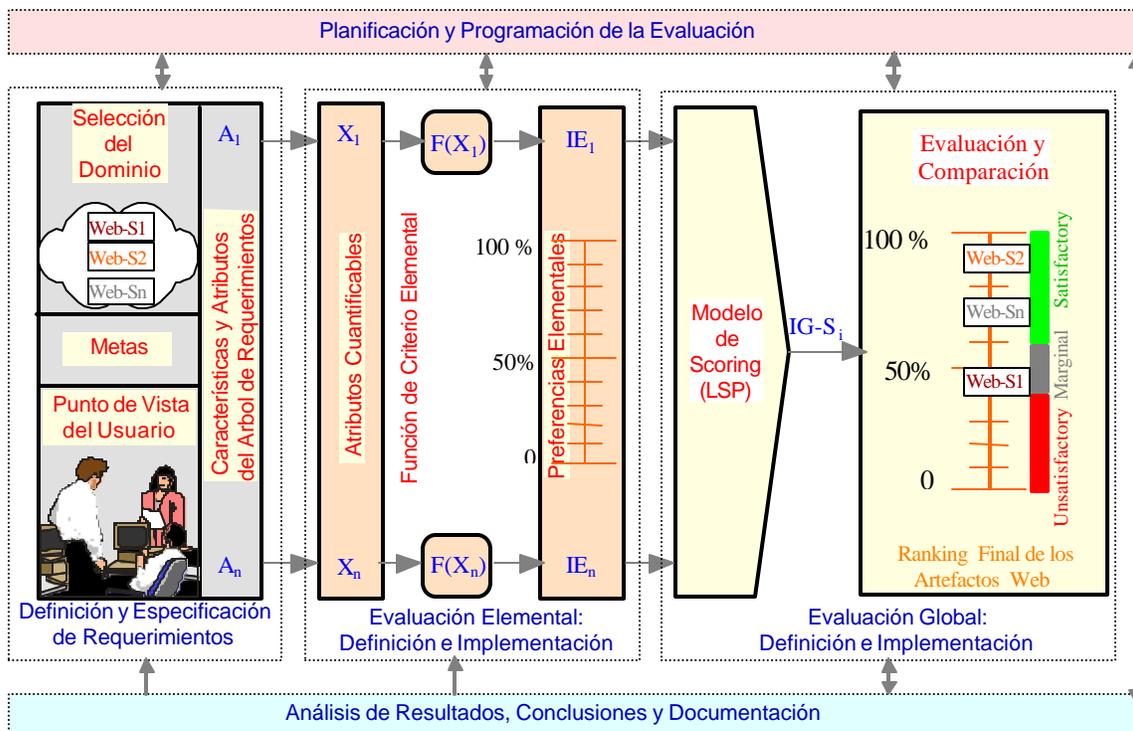


Figura 4.1 Un panorama de los principales módulos intervinientes en el proceso de evaluación y comparación usando Web-site QEM.

Con respecto a la fase de *Planificación y Programación de la Evaluación de Calidad*, la misma contiene actividades y procedimientos de soporte, con el fin de determinar objetivos estratégicos, tácticos y operativos. Esto es, permite establecer las principales estrategias y metas del proceso en un contexto organizacional; permite seleccionar un

modelo de proceso de evaluación, asignar métodos, agentes y recursos a las actividades; programar y replanificar una vez en marcha el proceso de evaluación. Si bien se tratarán algunos aspectos esta fase, no es objetivo central de este trabajo discutir las estrategias y sus actividades.

Considerando a la fase de *Definición y Especificación de Requerimientos de Calidad*, la misma trata con actividades y modelos para la elicitación, determinación, análisis y especificación de los requerimientos. A partir de un proceso de medición orientado a metas, y con el fin de evaluar, comparar, analizar, y mejorar características y atributos de artefactos Web, los requerimientos deben responder a necesidades y comportamientos de un perfil de usuario y dominio dados. El proceso de determinación de requerimientos, realizado en una mezcla de estrategias prescritivas y descriptivas, culmina con un documento que jerárquicamente especifica a todas las características y atributos cuantificables que modelan a la calidad según las necesidades del usuario. En el capítulo 5, discutiremos detalladamente a esta fase.

Con respecto a la fase de *Definición e Implementación de la Evaluación Elemental* la misma trata con actividades, modelos, técnicas y herramientas para determinar métricas y criterios de evaluación para cada atributo cuantificable. Se consideran tipos de criterios elementales, escalas, escalas de preferencia, valores críticos, y funciones para determinar la preferencia elemental, entre otros asuntos. Una vez definidos y consensuados los criterios para medir cada atributo, se debe ejecutar el proceso de recolección de datos, computar las métricas y preferencias elementales, y documentar los resultados. En el capítulo 6, discutiremos detalladamente los principales aspectos de esta fase.

Considerando a la fase *Definición e Implementación de la Evaluación Global* la misma trata con actividades, modelos, y herramientas para determinar los criterios de agregación de las preferencias de calidad elemental para producir la preferencia global, para cada sistema seleccionado. Se consideran tipos de funciones de agregación para modelar diferentes relaciones entre atributos y características, a saber: relaciones de reemplazabilidad, simultaneidad, neutralidad y diferentes niveles de polarización “y/o” (and/or). Una vez definidos y consensuados los criterios, se debe llevar a cabo el proceso de cálculo y ranquin. En el capítulo 7, discutiremos las principales tareas, modelos y procedimientos para esta fase.

Con respecto a la fase de *Análisis de Resultados, Conclusiones y Documentación* la misma trata con actividades de análisis y comparación de las preferencias de calidad elementales, parciales y globales, y, asimismo, la justificación de los resultados. Por otra parte, se utilizan herramientas y mecanismos de documentación para facilitar la interpretación de los datos y su seguimiento. En los capítulos 8 y 9, discutiremos procedimientos y herramientas para esta fase.

Por último, la *Validación* de las métricas es un proceso importante en la disciplina de evaluación de software. Podemos definirla como el proceso de asegurar que las medidas sean una caracterización numérica apropiada del atributo, mostrando que se satisfaga la condición de representación [Fenton et al 97]. Esto es, que la correspondencia entre el dominio empírico y el nuevo dominio numérico preserve a la relación funcional. Algunos aspectos de validación se irán discutiendo a lo largo de la tesis, (particularmente, le dedicaremos el capítulo 10). No obstante, algunos tópicos de esta fase, como validación de métricas predictivas, serán motivo de investigación futura.

Particularmente, en esta tesis discutiremos y ejemplificaremos por medio de dos casos de estudio [Olsina 99, Olsina et al 99a], los siguientes procesos de la metodología Web-site QEM, que son parte de algunas de las fases antes mencionadas, a saber:

- ✓ *Definiendo el Dominio y Ente para la Evaluación de la Calidad*
- ✓ *Definiendo Metas de Evaluación y Seleccionando el Perfil de Usuario*
- ✓ *Especificando Requerimientos de Calidad para artefactos Web*
- ✓ *Definiendo Criterios Elementales e Implementando Procedimientos de Medición (también llamado Determinación de la Preferencia de Calidad Elemental)*
- ✓ *Definiendo las Estructuras de Agregación e Implementando la Evaluación de Calidad Global*
- ✓ *Analizando y comparando los Resultados Parciales y Globales.*

A seguir, describimos sucintamente los procesos mencionados:

En el proceso “*Definiendo el Dominio y Ente para la Evaluación de la Calidad*”, los tomadores de decisión deben saber exactamente cuál es el dominio de la aplicación a evaluar y definir el ente a evaluar. Tres entes generales que pueden intervenir en procesos de evaluación son, a saber: procesos, productos, y recursos. Por ejemplo, para realizar un estudio con el objeto de valorar, comparar y determinar el estado del arte de la calidad sobre características específicas como usabilidad y funcionalidad, se debe seleccionar dentro de un dominio específico (como el dominio de museos, sitios académicos, comercio electrónico, etc.) un conjunto de sitios típicos (artefactos). Por otra parte, dado que las métricas y criterios que utilizaremos para ejemplificar son absolutas, se podría evaluar atributos y características de un solo artefacto.

En el paso, “*Definiendo Metas de Evaluación y Seleccionando el Perfil de Usuario*”, los evaluadores deben definir y refinar las metas y el alcance del proceso de evaluación. Ellos pueden evaluar un proyecto de desarrollo, o un sistema de información Web en la fase operativa. Pueden evaluar la calidad de un conjunto de características de parte de un sistema, o de un sistema completo, o comparar características y sus preferencias de calidad global de sistemas comparativos. Los resultados podrían ser utilizados para comprender, mejorar, controlar o predecir la calidad de artefactos Web. Por otra parte, la formulación de metas y, consecuentemente, la relativa importancia de las

características y atributos de calidad varían conforme al perfil de usuario seleccionado, entre otros factores. En términos generales, y considerando estándares como ISO [ISO/IEC 9126], nosotros consideramos tres perfiles de usuario, a un alto nivel de abstracción para dominios Web, a saber: visitantes, desarrolladores, y gerenciadore. Siguiendo un mecanismo de descomposición, en breve, podemos dividir a la categoría visitante en clases más específicas como veremos para los casos de estudio que iremos desarrollando.

En la tarea “*Especificando Requerimientos de Calidad para artefactos Web*”, los evaluadores deben acordar y especificar las características, subcaracterísticas y atributos de calidad agrupándolas en un árbol de requerimientos. Respecto de las características de calidad de más alto nivel, se sigue la misma clasificación conceptual que la prescrita en el estándar ISO [ISO/IEC 9126], y la casi idéntica clasificación dada por [IEEE Std 1061], en su Anexo informativo A, (en donde le llaman factores). Estas características de alto nivel son: **usabilidad, funcionalidad, confiabilidad, eficiencia, portabilidad, y mantenibilidad** (que serán discutidas en el próximo capítulo). Es importante destacar, sin embargo, que se puede cuantificar a la **calidad** de un producto de software, por la apropiada agregación y cuantificación de algunas o todas esas características, a partir de la medición directa e indirecta de atributos y la posterior agregación. De modo que, a partir de esas características, se derivan subcaracterísticas, y, a partir de éstas, siguiendo un proceso de descomposición recursivo, se pueden especificar atributos². Finalmente, considerando el dominio, las metas y el perfil de usuario (sus necesidades y comportamiento), se debe especificar el árbol de requerimientos de calidad. Técnicas de elicitación de requerimientos como cuestionarios y encuentros entre los participantes deben ser planificadas.

En el proceso “*Definiendo Criterios Elementales e Implementando Procedimientos de Medición*”, los evaluadores deben definir una base de criterios para la evaluación elemental; realizar el proceso de medición, y puntaje elemental. Un criterio de evaluación elemental declara y especifica cómo medir atributos cuantificables. El resultado final es una preferencia o ranquin elemental, el cual puede ser interpretado como el grado o porcentaje del requerimiento elemental satisfecho. Para cada variable medida X_i , $i = 1, \dots, n$ se define una función que representa al criterio elemental. Esta función es una correspondencia (mapeo) de los valores computados a partir del dominio empírico en el nuevo dominio numérico, y la denominamos preferencia de calidad elemental (IE_i), como se puede apreciar en la fig. 4.1 (en la sección 6.4.3 se muestran ejemplos específicos). Podemos asumir, como indicamos previamente, a IE_i como el porcentaje de requerimiento satisfecho para A_i . En este sentido, $IE_i = 0\%$ denota una situación totalmente insatisfactoria, mientras que $IE_i = 100\%$ representa una situación totalmente satisfactoria. Así, el puntaje elemental cae en uno de los tres niveles de

² Para cada atributo cuantificable A_i , podemos asociar una variable X_i , que puede tomar un valor real a partir de una medición. El valor final para el atributo es el computado a partir de la función elemental de preferencia

aceptabilidad, esto es, *insatisfactorio* (de 0 a 40%), *marginal* (desde 40 a 60%), y *satisfactorio* (desde 60 a 100%). (No obstante, el análisis de los puntajes cobra más importancia, cuando se está al final del proceso de evaluación).

En el proceso “*Definiendo las Estructuras de Agregación e Implementando la Evaluación de Calidad Global*”, los evaluadores deben establecer estructuras de agregación de preferencias elementales para producir la preferencia de calidad global. Posteriormente, se debe implementar el proceso de evaluación de modo de obtener un indicador de calidad global para cada sistema evaluado. Siguiendo la idea representada en la figura 4.1, para n atributos, la correspondiente función produce n preferencias elementales. Por lo tanto, aplicando un mecanismo de agregación paso a paso, las preferencias elementales se pueden agrupar convenientemente para producir al final el resultado global. La preferencia de calidad global representa el grado de satisfacción de todos los requerimientos explícitos e implícitos. En los casos de estudio de museos y sitios académicos usamos un modelo de puntaje lógico para estructurar y computar la calidad global de cada sitio Web. Específicamente, usamos el modelo Logic Scoring of Preference (LSP) fundado en la lógica de preferencia continua [Dujmovic et al 82; Dujmovic 96]. Sin embargo, otros modelos como los meramente lineales y aditivos [Gilb 76] podrían ser usados en este proceso cuando el sistema a evaluar no es complejo (discutidos en 3.1.2.1).

En el paso “*Analizando y comparando los Resultados Parciales y Globales*”, los evaluadores analizan, evalúan y comparan resultados parciales y globales considerando las metas y la vista de usuario establecidas. Este proceso culmina con la documentación de las conclusiones y recomendaciones. El proceso de evaluación, por medio de la metodología Web-site QEM, produce información elemental, parcial y global. La misma puede ser fácilmente analizada por medio de un modelo de seguimiento o trazabilidad, esto es, seguimiento hacia atrás (backward) y hacia delante (forward), y eficientemente empleada en actividades de toma de decisión.

4.2.1 Ejemplos a Utilizar

Para contextualizar esta discusión, a lo largo de la tesis emplearemos ejemplos y describiremos experiencias de algunos de los proyectos de evaluación realizados. Particularmente, haremos uso de una técnica de investigación con el fin de corroborar o refutar a la hipótesis nula. En nuestro caso se parte de la hipótesis que declara: “*que en sitios Web típicos, de un dominio determinado -como sitios académicos, museos, comercio electrónico, etc.-, la calidad de los artefactos (o sitios) satisfacen en general los requerimientos de calidad en consideración de un perfil de usuario. Particularmente, que cada sitio Web satisface al menos el punto crítico de aceptabilidad del 60% de la preferencia global, conforme a los requerimientos de calidad acordados*”.

Específicamente, en el párrafo anterior nos estábamos refiriendo a la técnica de investigación denominada *caso de estudio*. Los autores en [Fenton et al 97, Kitchenham et al 96], entre otros, indican que hay al menos tres técnicas o tipos de investigación que se basan en principios científicos y prácticas ingenieriles antes que en la intuición o sabiduría popular, a saber:

- ✓ *Experimento* (esto es, “research in the small”, con un nivel de control, rigor, y replicación alto)
- ✓ *Caso de Estudio* (esto es, “research in the typical” con un nivel de control menor)
- ✓ *Survey* –o Encuesta- (esto es, “research in the large”, en donde no es fácil manipular y tener control de las variables; además, la muestra es más bien grande y el estudio es generalmente retrospectivo)

En la sección 4.2.1.2 introducimos y realizamos algunas consideraciones sobre el caso de estudio para el dominio de museos, y, en la sección 4.2.1.3 ponemos en contexto el caso de estudio para el dominio de sitios académicos en la Web. Estos estudios nos permitieron utilizar la metodología propuesta y validar empíricamente sus principales procesos, métodos, modelos y herramientas.

Por otra parte, hemos realizado un survey sobre 29 atributos directamente mensurables derivados de algunas de las características como *usabilidad, funcionalidad y eficiencia*. La muestra fue seleccionada con criterios específicos (como tamaño del sitio) y la misma consistió de 24 museos, tomada de uno de los índices más relevantes en la Web, separada por continentes y países. El tamaño de la muestra se determinó sobre base no probabilística [Kendall et al 98]. El estudio se describe en el Apéndice A.3 de un modo informativo, dado que no es central para el desarrollo de esta tesis. (Solamente se utilizaron unas pocas actividades y procedimientos de Web-site QEM; sin embargo, es interesante incorporarlo como anexo dado que el estudio arrojó interesantes conclusiones y tendencias).

Finalmente, vale comentar que hemos realizado un proyecto de sitio Web bajo contrato con un ente público (el Ministerio de la Producción de la Provincia de La Pampa), cubriendo las fases de exploración y desarrollo. En este proyecto especificamos los requerimientos de calidad en etapas tempranas y evaluamos la calidad para el perfil de visitante general -la principal audiencia-, al momento de su distribución. Para el proyecto empleamos la metodología Web-site QEM como estrategia de soporte al proceso de evaluación.

4.2.1.1 Poniendo el foco en la Fase Operativa de un Sistema de Información centrado en la Web. Como se enuncia en el glosario, una fase es una agrupación de procesos fuertemente relacionados o cohesivos realizados en cierto orden. Las distintas fases pueden exhibir comportamientos diferentes. Por otra parte, se ha observado como común denominador en todo modelo de proceso de producto de software, las siguientes

tres fases: de exploración, de desarrollo y operativa (o mantenimiento) [Pressman 93].

La primera fase, denominada *de exploración*, se elicitan conceptos y requerimientos iniciales, si fuera necesario se realiza un estudio de factibilidad, y se planifica preliminarmente. La segunda fase, la *de desarrollo*, es la fase esencial de re-planificación, análisis y especificación de requerimientos detallados (funcionales y no funcionales [Davis 93]), coordinación y control, análisis del dominio, diseño, construcción, verificación y validación, integración, aseguramiento de la calidad, entre otros procesos (como cognitivos, estéticos, de autoría, etc. para proyectos de Hipermedia [Olsina 96 , 97c, 98b]). La tercera fase, denominada *fase operativa* o de vida útil de los artefactos, consiste esencialmente en tareas de documentación, configuración de cambios, mantenimiento y evolución de los mismos. En cuanto a la tarea de mantenimiento al menos podemos considerar tres tipos de mantenimiento: correctivo, adaptativo y perfectivo. Además, como indican los autores en [Powell et al 98], en los sitios Web operativos se debe introducir la idea de mantenimiento de contenido, debido a que los artefactos Web son tanto orientados a la documentación como a la aplicación de software tradicional.

Los casos de estudio fueron realizados para sitios Web en la fase operativa, y en el período de recolección de datos, en el proceso de medición, no se observaron cambios significativos, como se comentará seguidamente.

4.2.1.2 Caso de Estudio de Museos en la Web. Para analizar, comparar, y arribar a conclusiones respecto al estado del arte de la calidad sobre características específicas, (como usabilidad, funcionalidad, confiabilidad y eficiencia), se seleccionó el dominio de museos en la Web. Los mismos eran típicos y estaban publicados en Internet con al menos dos años de anterioridad con respecto a la fecha de la evaluación.

La figura 4.2 muestra las pantallas bajadas de la página principal de cada museo (en 1998). Se puede apreciar, en la fila superior a la izquierda, la “home page” del Museo del Prado (Madrid-España [Mu_Prado 98]); y a la derecha se muestra la página principal del Museo Nacional de Galería del Arte (Whashington-US [Mu_Gal_Art 98]). En la fila inferior, de izquierda a derecha, se muestran la página del Museo de Arte Metropolitano (Nueva York-US [Mu_Met 98]) y la del Museo Louvre (París-Francia [Mu_Louvre 98]). Como el lector puede observar, los sitios son bien conocidos internacionalmente, sin duda como museos en su “versión física”, tanto por su magnitud como por su relevancia. No obstante, como indicamos anteriormente, también son conocidos en su “versión virtual”.

Las imágenes de la figura fueron bajadas en el período de recolección de datos que fue efectuada desde el 15 de Septiembre, al 15 de Octubre de 1998. Es importante tener en mente la siguiente consideración: los sitios Web son artefactos que dinámicamente pueden evolucionar y siempre se accede a la última versión en línea.



Figura 4.2 En la fila superior, de izquierda a derecha, se muestra las páginas principales del Museo del Prado y del Museo Nacional de Galería del Arte. En la fila inferior, de izquierda a derecha, se muestran las del Museo de Arte Metropolitano y del Museo Louvre. Estas figuras fueron bajadas en el período de recolección de datos (desde el 15 de Septiembre, al 15 de Octubre de 1998)

Al momento de la evaluación no percibimos cambios que pudieran haber afectado al proceso de evaluación. No obstante, luego de la recolección de datos, a fines de Noviembre, el museo Louvre cambió significativamente su look & feel, y, consecuentemente, el resultado final de la evaluación de los atributos que conciernen a la interface y a características de los objetos de control navegacional, habrían cambiado con esta nueva versión. La figura 4.3 muestra, de un modo anecdótico, la nueva página principal del museo Louvre (bajada el 5-04-99). Principalmente se hace uso de marcos (frames) para separar el agrupamiento de controles del marco de presentación de contenidos.

Además, el museo Louvre ha introducido recientemente (noviembre de 1998) tours virtuales (usando tecnología VR™). Si bien se consideró dentro de los requerimientos el atributo de visita o tour guiado, al momento de la evaluación no estaba presente en ninguno de los sitios evaluados, ya que era una tecnología emergente. (Sin embargo, prueba del creciente interés por incorporar tours VR en museos, se manifiesta en la

tendencia observada en el survey –cuyos datos se recolectaron en el período noviembre-diciembre de 1998-, y que algunos de los resultados se presentan en el Apéndice A.3).

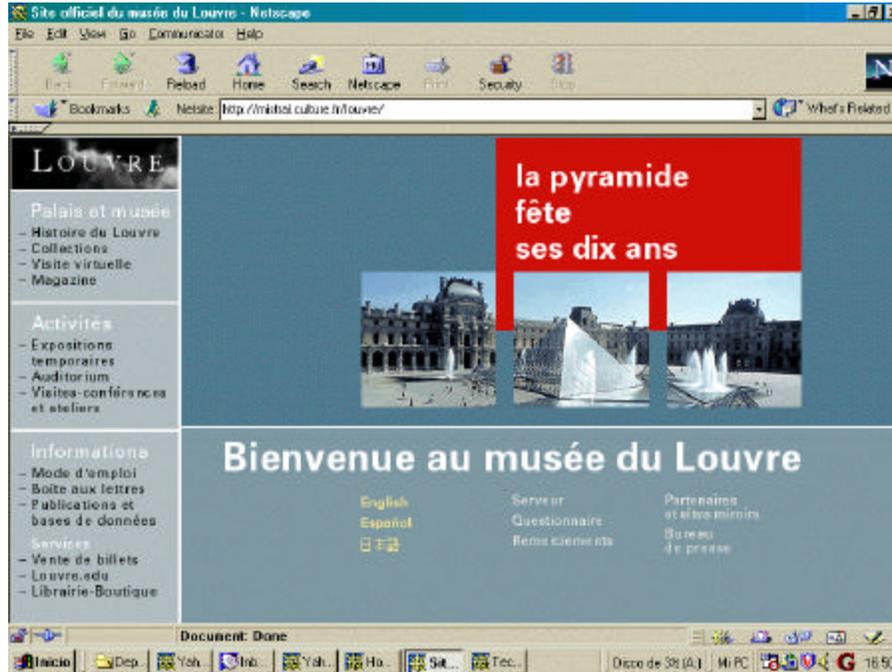


Figura 4.3 Se muestra la página principal del Museo Louvre luego del período de recolección de datos para el caso de estudio de museos, la cual cambió su apariencia (bajada el 05-04-1999)

Finalmente, el museo Metropolitano incorporó funcionalidad de búsqueda global no disponible al momento de la recolección de datos. Afortunadamente, muchos documentos de los sitios fueron bajados y almacenados.

4.2.1.3 Caso de Estudio de Sitios Web Académicos. Al diseñar el caso de estudio de sitios académicos, seleccionamos seis sitios Web, en la fase operativa, con al menos tres años de haber estado publicados en la red (esto se determinó por medio de una herramienta denominada SiteSweeper™).

La figura 4.4 muestra las pantallas bajadas de la página principal de cada sitio académico. Se puede apreciar en la fila superior las páginas de la Universidad de Stanford (US - [Un_Estados_Un 99]), y de la Universidad Tecnológica de Sidney (UTS - Australia [Un_Australia 99]). En la parte media, se muestran las pantallas de la Universidad Politécnica de Cataluña (UPC – España - [Un_España 99]), y la Universidad de Quebec en Monreal (UQAM – Canadá - [Un_Canadá 99]). Finalmente, en la fila inferior se muestran las páginas de la Universidad de Chile (UChile - [Un_Chile 99]), y la Universidad Nacional de Singapur (NUS - [Un_Singapur 99]). Estas imágenes fueron bajadas en el período de recolección de datos (desde el 22 de Enero, al 22 de Febrero de 1999).



Figura 4.4 En la fila superior, de izquierda a derecha, se muestran las páginas principales de la Universidad de Stanford y la Universidad Tecnológica de Sidney. En la parte media, de izquierda a derecha, se muestran las de Universidad Politécnica de Cataluña, y Universidad de Quebec en Montreal. Finalmente, en la fila inferior se muestran las páginas de Universidad de Chile, y Universidad Nacional de Singapur. Estas figuras fueron bajadas en el período de recolección de datos (efectuado entre el 22 de Enero y el 22 de Febrero de 1999)

Al igual que en el caso anterior, la meta principal de este caso de estudio fue comprender y comparar el nivel de cumplimiento de características y atributos de calidad a partir de un conjunto acordado de requerimientos, en consideración de

estudiantes actuales y futuros (de grado y posgrado). Particularmente, evaluamos el nivel de cumplimiento de la hipótesis según la expresamos al inicio de la sección 4.2.1.

Por otra parte, debemos hacer alguna consideración con respecto al tipo de recolección de datos. Esta actividad se puede realizar manualmente, semiautomáticamente, o de un modo totalmente automática. La mayoría de los valores para la función elemental fueron recolectados manual y observacionalmente, dado que no habría otro modo de hacerlo efectivamente. Sin embargo, la recolección de datos automática es en muchos casos el único mecanismo para obtener resultados de un modo confiable y efectivo. Este fue el caso para medir atributos como *Enlaces Rotos*, *Páginas de Acceso Rápido*, entre otros, como será ejemplificado y discutido en las secciones 6.4.3 y 6.6.

Finalmente, esta evaluación fue focalizada esencialmente en el sitio de la Universidad como un todo antes que en cada unidad académica como facultad, laboratorio o instituto.

Capítulo 5

Fase de Definición y Especificación de los Requerimientos de Calidad.

En este capítulo, consideraremos aspectos de la fase de *Definición y Especificación de Requerimientos de Calidad*. Esta fase trata con actividades y procedimientos para la elicitación, modelado y especificación de los requerimientos de calidad. A partir de un proceso de evaluación realizado en una mezcla de estrategias prescriptivas y descriptivas (el enfoque de modelo mixto de calidad), y con el fin de analizar, comparar, comprender y potencialmente mejorar características y atributos de artefactos Web, los requerimientos deben responder a necesidades y deseos de un perfil (o perfiles) de usuario y dominio establecidos.

El proceso de determinación de requerimientos culmina, por una parte, con un documento que jerárquicamente lista a todas las características y atributos cuantificables (y que modelan a la calidad según las necesidades del usuario). De manera que a partir de esas características, se derivan subcaracterísticas, y, a partir de éstas, siguiendo un proceso de descomposición recursivo, se especifican atributos. Por otra parte, para cada atributo cuantificable A_i podemos asociar una variable X_i , que puede tomar valores reales a partir de un proceso de medición. El valor final computado para el atributo se corresponderá a una preferencia elemental (definida y ejecutada en la siguiente fase – ver figura 4.1). Por lo tanto, los requerimientos de calidad quedarán completos, luego de acordar un conjunto de valores y rangos para cada atributo. (Esta actividad puede realizarse en un estilo incremental e iterativo con las actividades de la siguiente fase, de determinación de los criterios de preferencia elementales)

5.1 Definición del Dominio y Ente de Evaluación

Podemos definir al dominio de la aplicación, desde el punto de vista de la evaluación, como a un sistema real o abstracto del universo que existe independientemente del sistema de evaluación. Consiste de un conjunto de entes a los que se le atribuyen propiedades (*atributos, características*), manifiestan un comportamiento y se relacionan.

Por una parte, tenemos a los entes y relaciones propios del dominio de la aplicación; por ejemplo, para un sistema académico podemos considerar entes o clases como *Curso, Departamento, Estudiante, Carrera*, entre otras, y relaciones entre ellas, como *asiste, pertenece*, etc. Esto se puede observar en la figura 5.1 (en donde se usa la notación UML [UML 97] ligeramente adaptada para expresar atributos multi-tipados [Rossi 96]). Además, a los entes y relaciones se le atribuyen propiedades. Por ejemplo, la clase

Equipamiento tiene atributos como *modelo*, *descripción*, etc.; y la relación *enseña* tiene atributos como *primer semestre* o *segundo semestre*.

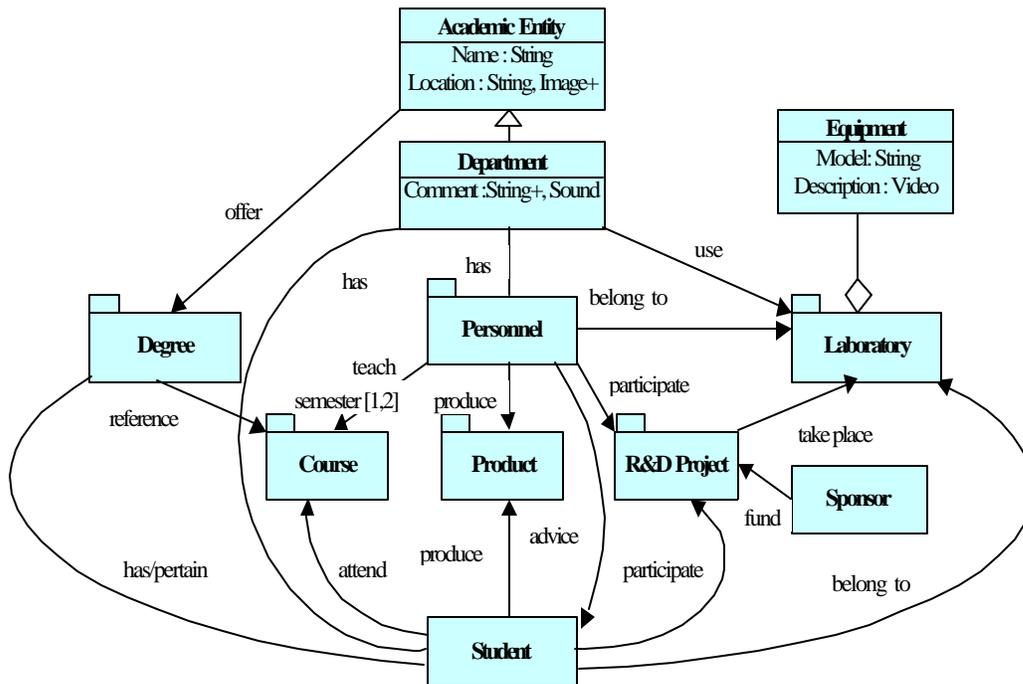


Figura 5.1 Diagrama reducido de clases y relaciones para un Sistema de Información Académico [Olsina 98c].

Este modelo conceptual es importante dado que representa y comprende al dominio de la aplicación de un sistema de información, y, por lo tanto, definirlo y representarlo, es una actividad necesaria en el proceso de desarrollo. Sin embargo, desde el punto de vista de la evaluación, es también necesario definir al ente de la evaluación.

De manera que, un proceso de evaluación debe identificar a los entes, a las características, y a los atributos a medir, para posteriormente analizar, comprender y emplear los resultados apropiadamente. Desde el punto de vista de la evaluación, en Ingeniería de Software hay tres entes o clases principales de interés [Fenton 91, Goldberg et al 95, Olsina 98d], a saber:

- ✓ Proceso: es el ente compuesto posiblemente de otros subprocesos y actividades, usado para producir artefactos;
- ✓ Artefacto: es el ente temporario o persistente que representa al producto de realizar un proceso,
- ✓ Recurso: es un ente requerido como entrada por un proceso para producir alguna salida especificada (recursos de un proyecto son: humanos, monetarios, materiales, tecnológicos, temporales)

La tabla C.1 (del apéndice C), muestra a un alto nivel de abstracción, a estos entes y a

algunas de las características y atributos de interés, personalizado para el campo de Hipermedia.

Por lo tanto, es importante diferenciar, comprender (y relacionar, cuando fuera necesario) a los entes del dominio de la aplicación, y a los entes del proceso de evaluación. A modo de ejemplo, la meta de la evaluación para el dominio de un Sistema de Información Académico centrado en la Web, podría consistir en evaluar individualmente la calidad del ente artefacto Web y compararlo con otros artefactos representativos, a partir de un conjunto de características y atributos acordados. Por lo tanto, tenemos a dos subprocesos de importancia en nuestro ejemplo: la definición y comprensión del dominio de evaluación, y la definición (y selección) de los productos a evaluar. En el primer caso es necesario que los evaluadores sean expertos (o ganen conocimiento) en el dominio de la aplicación; de este modo, comprenderán y establecerán criterios con mayor madurez respecto a las características y atributos del artefacto a evaluar. Por otra parte, para realizar un caso de estudio con el objetivo de comprender el estado del arte de la calidad de artefactos Web sobre características específicas como por ejemplo, usabilidad, funcionalidad y performance, se debe seleccionar dentro del dominio, un conjunto de sistemas típicos.

Sin embargo, estos subprocesos deben ser vistos en conjunción con las actividades de definición de las metas de evaluación y, la de selección del perfil (o punto de vista) del usuario a evaluar.

5.2 Definición de las Metas de Evaluación

Como citamos en la sección 3.1: “*Los proyectos sin metas claras no alcanzarán sus metas claramente*” [Gilb 88]. Asimismo, podemos recordar el viejo pero siempre vigente principio que dice: si no sabemos en dónde estamos parados, un mapa no ayuda para alcanzar el destino (la meta). Por lo tanto, en esta actividad los participantes del proyecto de evaluación deben definir y especificar claramente las metas y el alcance del proceso.

Los objetivos y metas del proceso de evaluación pueden ser muy variados y específicos del tipo de proyecto (la criticidad del mismo, o de alguno de sus componentes), y de la madurez de la organización. Los participantes pueden evaluar a un ente de un proyecto de desarrollo, o a un proyecto en la fase operativa. Pueden evaluar a la calidad de un par de características de un ente; por ejemplo, la confiabilidad de un artefacto o sistema en función de la cantidad y frecuencia de defectos; pueden evaluar a la calidad global de un sistema completo, en donde intervienen todas las características esenciales; o pueden comparar a la calidad global de varios sistemas a la vez. Los resultados podrían ser utilizados para comprender, controlar, mejorar o predecir a la calidad del ente (particularmente en nuestro caso, artefactos Web).

Sea por ejemplo, el caso de un artefacto en la fase operativa, el que puede permanecer sin mayores cambios en tanto sus preferencia global se percibe como satisfactoria. Pero una vez que es percibida como marginal o insatisfactoria, una acción correctiva se debe planificar. Tal acción generalmente asume la transición desde una situación actual (marginal o insatisfactoria) a una situación de mejora (meta a alcanzar).

5.2.1 Enfoques Orientados a Metas.

En este punto, es importante resaltar que ningún evaluador o desarrollador tiene tiempo de medir y analizar cualquier ente, cualquier característica, o cualquier atributo, sino que debe concentrarse en aquellos que provean un eficiente análisis o comprensión requerida, en el contexto de una organización. Por lo tanto, la pregunta es, ¿a partir de qué enfoque o estrategia seleccionamos las características observables (por ejemplo, navegabilidad, performance, usabilidad, etc.)?, ¿Cómo seleccionamos las métricas para esas características y atributos y en qué contexto las analizamos e interpretamos?

De acuerdo a investigaciones realizadas [Basili et al. 84, 94], las métricas para que sean efectivas deben estar focalizadas en metas específicas, aplicadas a todo o parte del ciclo de vida de los entes e interpretadas en función de la comprensión del contexto organizacional. Esto implica que las mediciones se deben definir bajo una estrategia de arriba hacia abajo, o también conocida como “top-down”. De estos autores surge el modelo Goal-Question-Metrics (GQM).

Para no salirnos demasiado del foco del desarrollo de este capítulo, el lector interesado puede ir al Apéndice C, en donde en la sección C.2, resaltamos aspectos básicos del modelo GQM; y, en la sección C.3 ponemos en contexto al enfoque por medio de un ejemplo utilizado en la medición de la navegabilidad de un proyecto de autoría, realizado en el año 1995 [Olsina et al 95]. Además, en la sección C.3 realizamos algunos comentarios para este enfoque, el cual es útil para considerar métricas en función de metas y objetivos. Algunas potencialidades y limitaciones son resaltadas. Particularmente nos interesa transcribir las siguientes dificultades en la utilización del modelo:

- ✓ El enfoque GQM no trata aspectos de criterios elementales para medir un atributo, ni cuestiones de selección, estructuración, y agregación de atributos y características en el proceso de evaluación.
- ✓ Además, no trata cuestiones inherentes a sistemas de complejidad media o alta en el proceso de valoración y comparación, como cantidad de atributos; relaciones entre atributos y características como simultaneidad, reemplazabilidad, y neutralidad, entre otros asuntos.

En nuestro caso, si bien no empleamos sistemáticamente el enfoque GQM para los

ejemplos introducidos en la sección 4.2.1, podemos decir que la meta consistió en "*Comprender a la calidad global de un sitio Web desde el punto de vista del visitante general*". En secciones posteriores, discutiremos cuál es la utilidad de este enfoque, en el contexto de otros modelos de descomposición de características, atributos y métricas.

5.3 Selección del Perfil de Usuario

Por otra parte, la formulación de metas y, consecuentemente, la relativa importancia de las características y atributos de calidad, varían conforme al perfil (o perfiles) de usuario seleccionado.

5.3.1 Perfiles de Usuario prescritos por el Estándar ISO/IEC 9126.

En términos generales, y teniendo como marco conceptual al estándar ISO [ISO/IEC 9126], nosotros consideramos tres perfiles de usuario, a un alto nivel de abstracción para dominios Web, a saber: visitantes, desarrolladores, y gerenciadore. Siguiendo un mecanismo de categorización y descomposición podemos dividir, por ejemplo a la categoría visitante en clases más específicas como veremos para los estudio que iremos desarrollando.

Particularmente, el estándar citado afirma que la relativa importancia de las características de calidad (como usabilidad, funcionalidad, confiabilidad, eficiencia, portabilidad, y mantenibilidad) varían dependiendo del punto de vista considerado y de la criticidad de los componentes del artefacto a evaluar. Por ejemplo, la visión del usuario (que nosotros denominamos punto de vista o perfil del visitante, para el dominio Web), concierne al interés de los mismos en usar el artefacto de software, como así también su performance, su eficiencia, su facilidad de uso, entre otros aspectos. Los usuarios no están interesados en características internas o de desarrollo de los artefactos (sin embargo, atributos internos contribuyen a la calidad de uso).

En cambio, la visión de calidad del desarrollador debe considerar no sólo los requerimientos del producto para la visión del usuario sino también la calidad para los productos intermedios resultantes de las actividades de la fase de desarrollo. Para poder evaluar a la calidad de los productos intermedios de las diferentes actividades de la fase de desarrollo, los evaluadores deberán considerar las métricas apropiadas para la misma característica de acuerdo al proceso. Por otra parte, los desarrolladores están preocupados en características de calidad del producto como mantenibilidad y portabilidad.

Finalmente, la visión de calidad del gerenciador es una visión integradora, que necesitará incorporar requerimientos de negocio a las características individuales. Por ejemplo, un gerente puede estar interesado en el óptimo equilibrio entre la mejora de

un producto dentro de costos y tiempos establecidos.

5.3.2 Perfiles de Usuario considerando las Metas y el Dominio de Aplicación en la Web.

En la sección 4.2.1.2 introducimos y realizamos algunas consideraciones sobre el caso de estudio para el dominio de museos en la Web. El perfil de usuario seleccionado para este estudio fue el del visitante, y, en particular, el del visitante general (esto es, casuales o intencionales). Como indicamos en la sección anterior, siguiendo un mecanismo de categorización y descomposición podemos dividir, por ejemplo, a la categoría visitante en clases más específicas, a saber: visitantes generales y visitantes expertos. A su vez, según estudios realizados sobre el dominio de museos [Furano et al 97, Garzotto et al 97], los visitantes generales, pueden ser descompuestos en dos audiencias más específicas: visitantes casuales y visitantes intencionales.

Según Garzotto et al., un *visitante casual* se define como a la audiencia que ingresa al sitio por azar, y permanece probablemente un corto lapso de tiempo, en donde la aplicación está pensada para ser usada probablemente una sola vez. Esta audiencia puede diferir en los niveles de experticia en el uso de tecnologías Web. Por otra parte, un *visitante intencional* se define como a la audiencia que tiene al menos algún conocimiento o manifiesta algún interés en el dominio de museos, y desea probablemente informarse o aprender más acerca del mismo. Su permanencia en el sitio es generalmente mayor que la de una audiencia casual. Por último, un *visitante experto* se define como a la audiencia que es especialista en el dominio de museos, como investigadores de la historia del arte, entre otros. Su permanencia en el sitio es generalmente mayor que la de una audiencia casual ([Garzotto et al 97] pp. 42). Asimismo, según Furano et al. un visitante experto es aquél que se encuentra interesado generalmente en todo el contenido del museo (en la Web) y en todos los itinerarios provistos tanto en la página principal como en las restantes ([Furano et al 97] pp. 122).

De manera que para la definición de los requerimientos de calidad para el dominio de museos en la Web se consideró el punto de vista del *visitante general* con el fin de evaluar y comparar la calidad global de los cuatro sitios típicos, en la fase operativa (a saber: Museo del Prado, Museo Nacional de Galería del Arte, Museo de Arte Metropolitano y Museo del Louvre –ver figura 4.2).

Finalmente, en la sección 4.2.1.3 pusimos en contexto el caso de estudio para el dominio de sitios académicos en la Web. El perfil de usuario seleccionado para este estudio fue el del visitante, en consideración de estudiantes actuales y futuros (de grado y posgrado). En cuanto al punto de vista del visitante, tres audiencias han sido identificadas [Lowe et al 98, Schwabe et al 95a], a saber: estudiantes actuales y futuros; personal académico, esencialmente profesores e investigadores; y patrocinadores de la investigación. Debe quedar claro que las necesidades de información para el perfil del

estudiante, interesado primeramente en carreras, cursos, servicios e infraestructura académica en general, es generalmente distinta de los requerimientos de un gerente que desea financiar un proyecto de investigación, y el cual estará primeramente interesado como visitante en el perfil del proyecto, sus antecedentes y producción científica realizada; el nivel de sofisticación de equipamiento de los laboratorios e institutos de investigación, y los recursos humanos calificados con que cuenta. Esta división orientada a las diferentes audiencias, está claramente establecida en la estructura del sitio Web de la Universidad Tecnológica de Sidney [Un_Australia 99].

5.4 Representación de las Características y Atributos de Calidad: Modelos de Calidad

En este paso los evaluadores y demás participantes del proceso deben acordar y especificar las características y atributos de calidad y agruparlas en un árbol de requerimientos, conforme a las metas, el dominio, y el perfil de usuario seleccionados.

La idea bastante generalizada de la calidad como el producto que está libre de defectos, es una concepción muy reducida para nuestros objetivos de evaluación¹ y tampoco coincide con la idea global de calidad que se forma intuitivamente un usuario con el perfil seleccionado (ni con la concepción integral de calidad propuesta por los modelos de calidad que seguidamente discutiremos).

Definimos a la calidad de un ente como al conjunto de las características de dicho ente (artefacto, proceso, o recurso) que le confiere la capacidad de satisfacer requerimientos o necesidades explícitas e implícitas, en consideración de un perfil de usuario. En el mismo sentido, podemos alternativamente describir a la calidad como a una característica de muy alto nivel, que se la puede medir mediante métricas directas e indirectas y procesos de agregación, de utilidad en el proceso de evaluación y comparación de ciertas propiedades y relaciones atribuidas a un ente.

Para nuestro proceso de evaluación, como ya lo hemos mencionado, el ente de interés es el artefacto Web. Además, es importante destacar que la calidad de un artefacto Web se la puede cuantificar mediante la apropiada descomposición de algunas o todas las características de un conjunto prescrito, esto es, a partir de la medición directa de un conjunto de atributos (derivados a partir de esas características) y por medio de un proceso de agregación y cómputo de preferencias elementales, para producir

¹ En este sentido Fenton et al [Fenton et al 97], realizan una interesante observación con respecto a la idea de calidad relacionada con pocos defectos. Comentan que bajo ciertas circunstancias las medidas basadas en defectos pueden ser útiles, pero que no siempre se puede asumir que sean indicadores exactos de calidad (como percibidos por los usuarios). La limitación subyace en que los defectos descubiertos durante la actividades de testeo, inspección o compilación de programas, pueden llegar a no producir fallas en la fase de operación. Así, un alto nivel de defectos no siempre es un indicador de baja calidad, y un bajo nivel de defectos no siempre podrá significar un alto nivel de calidad.

preferencias de calidad parciales y global. (Tal es, de un modo general, el procedimiento propuesto e implementado en el presente trabajo).

Finalmente, nos interesa recalcar que dado que la calidad es una composición de muchas características y sub-características el concepto de calidad es, consecuentemente, representado en modelos que muestran dichos componentes y sus relaciones primarias.

En la sección siguiente describimos algunos modelos de calidad de software (como los de McCall y Boehm) que han ganado su aceptación en la comunidad de Ingeniería de Software. Posteriormente, en la sección 5.4.2, describimos y comentamos las características de calidad prescritas en el estándar ISO [ISO/IEC 9126], y la casi idéntica clasificación dada por [IEEE Std 1061], en su anexo informativo A. En la sección 5.4.2.1 describiremos los puntos salientes del nuevo modelo de calidad de la ISO [ISO/IEC 9126-1] el cual aún no ha sido publicado oficialmente. En la sección 5.4.3 presentaremos nuestro enfoque de modelo de calidad; y en la sección 5.4.4 vemos un mecanismo de selección de importancia relativa de características. A seguir, en 5.5 presentamos conceptos para un árbol de requerimientos general e independiente de dominios de artefactos Web específicos. Finalmente, en la sección 5.6 ejemplificamos la representación del árbol para el caso de estudio de sitios académicos.

5.4.1 Los Primeros Modelos de Calidad de Producto

Aunque a fines de la década del 80 e inicios de la del 90 se ha puesto mucho énfasis en los conceptos de calidad de producto y satisfacción del usuario desde diversos enfoques, y particularmente a la valoración y certificación de la calidad de procesos con los bien conocidos marcos conceptuales de evaluación como CMM y SPICE (entre otros), sin embargo, también es conocido que los modelos de calidad ya eran reconocidos en la comunidad científica a fines de la década del 70 como los descriptos por McCall [McCall et al 77] y Boehm [Boehm et al 78].

Estos modelos describen a la calidad del producto usando un enfoque de descomposición top-down. En la figura 5.2 el lector puede apreciar el modelo de calidad de McCall; y en la figura 5.3 se puede apreciar la visión del modelo de calidad del producto según Boehm. El modelo de McCall fue originariamente desarrollado para la Fuerza Aérea de US y se promocionó su uso para evaluar la calidad del software dentro del DoD. No obstante, otros modelos se han usado dentro del mismo departamento.

En estos modelos, los evaluadores se concentran en los atributos de calidad claves para el producto software, en consideración de un punto de vista de usuario. Esos atributos claves son llamados en el modelo de McCall, *factores de calidad* (ver columna *Factor* en la fig. 5.2).

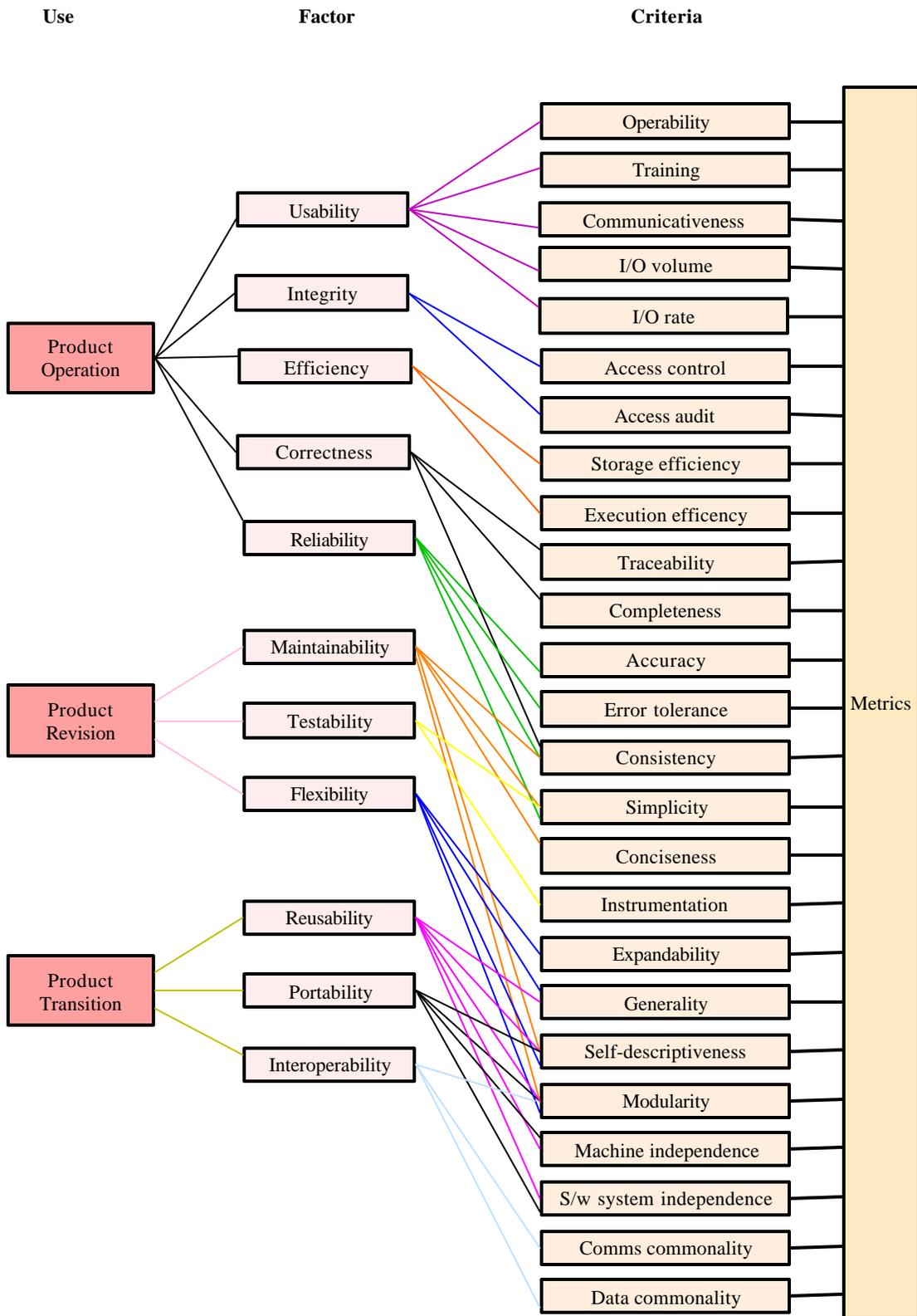


Figura 5.2 Modelo de Calidad de Software de McCall [McCall et al 77]

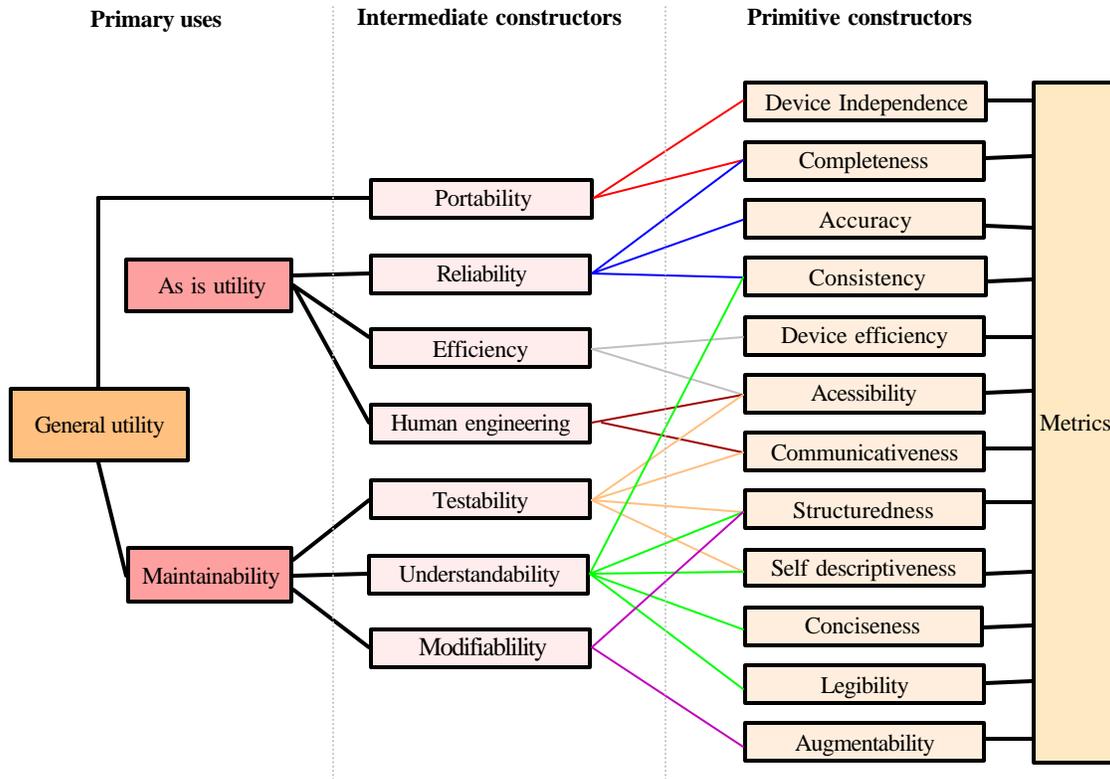


Figura 5.3 Modelo de Calidad de Software de Boehm [Boehm et al 78]

Por ejemplo, para personalizar el modelo de calidad con el objetivo de valorar un producto en operación (ver columna *Use*), los evaluadores seleccionan los factores “*Efficiency*” y “*Reliability*”. Pero estos factores son todavía de muy alto nivel para poder ser medidos directamente; por lo tanto, se descomponen en atributos de más bajo nivel llamados *criterios de calidad*. Siguiendo con el ejemplo, para el factor “*Efficiency*” se cuenta con el criterio “*Storage Efficiency*” y el criterio “*Execution Efficiency*”.

Algunas veces, se requiere un nivel de descomposición mayor, en el cual se asocia el criterio de calidad con un conjunto de métricas directamente mensurables.

5.4.1.1 Otros Modelos o Marcos Conceptuales de Descomposición. En este punto es importante resaltar otros marcos conceptuales para representar factores, atributos y métricas como así también el modelo subyacente de descomposición. La figura 5.4 muestra el marco conceptual conforme al estándar IEEE 1061 (en [IEEE Std 1061] pp. 5) para la descomposición de métricas de calidad de software.

El diseño es flexible dado que permite agregar, modificar y borrar factores, subfactores y métricas. Cada nivel puede ser expandido a varios subniveles. En el primer nivel de la

jerarquía, a nivel de factor, comienza con el establecimiento de los requerimientos de calidad asignando varias características (factores²) de calidad. Todos los requerimientos se deben acordar con los integrantes del proyecto, y se deben establecer las definiciones. Asociado a cada factor hay una métrica o valor establecido, que servirá posteriormente para saber si el factor ha sido alcanzado.

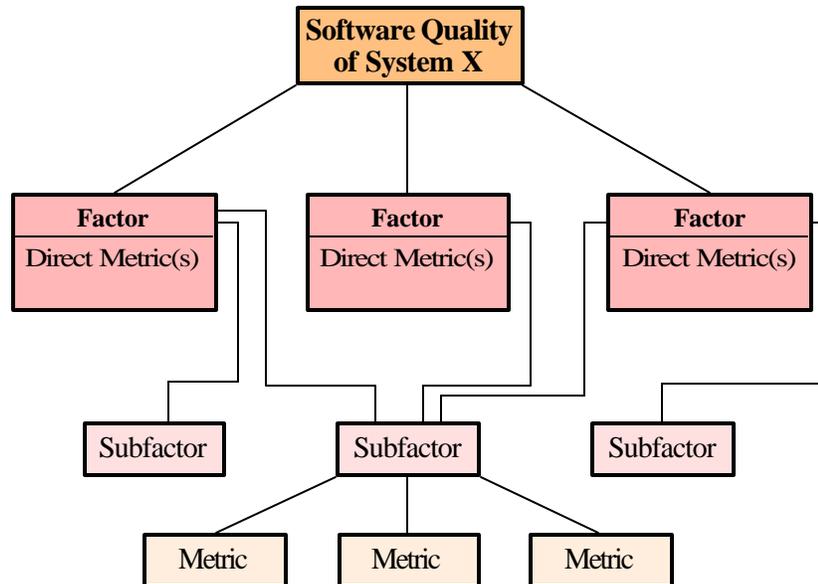


Figura 5.4 Marco conceptual para la descomposición de métricas de calidad de software conforme al estándar IEEE 1061

En el segundo nivel de la jerarquía, están los subfactores que son atributos concretos del software con mayor significado que los factores para el personal técnico como analistas, desarrolladores, personal de mantenimiento, entre otros. La descomposición de factores en subfactores, facilita objetivos de comunicación entre el gerente y el personal técnico, en consideración de los objetivos de calidad.

Finalmente, en el tercer nivel, los subfactores se descomponen en métricas usadas para medir productos o procesos durante el ciclo de vida de desarrollo. Los valores de los factores (y sus métricas directas), no están disponibles o son costosas de recolectar en etapas tempranas de la fase de desarrollo. Por lo tanto, las métricas al tercer nivel que sean validadas con las métricas directas, serán usadas para estimar el valor de los factores en etapas tempranas. Este marco conceptual facilita, de arriba hacia abajo (– según la cita, en IEEE 1061 pp. 6):

- ✓ El establecimiento de requerimientos de calidad en términos de factores, por parte de los gerentes, en etapas tempranas del ciclo de vida

² Un factor de calidad se define (según este estándar) como un atributo orientado a la gerencia, y que contribuye a la calidad de software, a diferencia de un subfactor que está orientado más bien al personal técnico.

- ✓ Comunicación de los factores acordados, en términos de subfactores, al personal técnico
- ✓ Identificación de métricas que estén relacionadas a factores y subfactores.

Y de abajo hacia arriba, posibilita al personal técnico y gerencial obtener retroalimentación:

- ✓ Al evaluar los procesos y productos de software, a nivel de métricas elementales
- ✓ Al analizar los valores de las métricas obtenidos para estimar y valorar los factores de calidad

A pesar de los claros beneficios indicados que tiene el utilizar un marco conceptual como el descrito, al momento de su implementación, nos encontraremos con semejantes dificultades a las informadas en el modelo GQM (comentadas en 5.2.1). No obstante, dicho estándar en su anexo informativo, prescribe un conjunto de factores, subfactores y sus relaciones aunque no muestra relaciones a niveles inferiores.

5.4.2 Modelos de Calidad de Producto Prescritos en los Estándares ISO 9126 e IEEE 1061

Las ventajas de contar con un modelo de calidad universal serían varias, pero esencialmente nos permitiría evaluar y comparar productos, potencialmente, sobre la misma base.

En 1992 fue publicado un estándar internacional para la evaluación de la calidad de producto de software, llamado *“Information technology – Software product evaluation – Quality characteristics and guidelines for their use”*; o también conocido como ISO 9126 (o ISO/IEC 9126). El mismo tiene sus raíces, principalmente, en el modelo de calidad de producto propuesto por McCall, mostrado en la sección anterior.

Este estándar describe a la calidad del software, con mínimo solapamiento, a partir de seis características generales. *“Esas características proveen una línea base para ulteriores refinamientos y descripciones de la calidad del software”*. Las características fueron mencionadas previamente, y ahora son definidas transcribiéndolas de su fuente original:

- ✓ **funcionalidad**, y que textualmente la define: *“A set of attributes that bear on the existence of a set of functions and their specified properties. The functions are those that satisfy stated or implied set of users”*
- ✓ **confiabilidad**, y que textualmente la define: *“A set of attributes that bear on the capability of software to maintain its level of performance under stated conditions for a stated period of time”*

- ✓ **usabilidad**, y que textualmente la define: “*A set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users.*”
- ✓ **eficiencia**, y que textualmente la define: “*A set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions*”
- ✓ **mantenibilidad**, y que textualmente la define: “*A set of attributes that bear on the effort needed to make specified modifications*”
- ✓ **portabilidad**, y que textualmente la define: “*A set of attributes that bear on the ability of software to be transferred from one environment to another*”

Como idea central podemos extraer que esta versión del estándar ISO de características de calidad de producto software, provee una base conceptual muy importante al prescribir dichas características a un alto nivel de abstracción, pero si bien sugiere subcaracterísticas de un modo informativo (en su anexo A), “... *no provee [prescribe] subcaracterísticas y métricas ni métodos para la medición del puntaje y evaluación*” ([ISO/IEC 9126], pp 1).

En la tabla 5.1, el lector cuenta con una descripción sintética que hemos realizado de las características y subcaracterísticas de calidad del estándar ISO, por medio de la formulación de una pregunta central.

Del mismo modo, en la tabla 5.2, el lector cuenta con una descripción sintética de los factores y subfactores conforme al anexo A (informativo) del estándar IEEE 1061. Claramente el estándar IEEE declara que no prescribe características (factores), ni subcaracterísticas, ni métricas.

Sin embargo, es importante comprobar que en lo que respecta a la definición de características en ambos trabajos, coinciden significativamente en su contenido, manifestándose algunas diferencias en las subcaracterísticas informadas (o subfactores), para alguna de las características. Si bien no profundizaremos en las diferencias encontradas³, en la nota al pie se encuentran algunos tipos y ejemplos.

³ Las diferencias entre ambos estándares son de distinto tipo, entre las que podemos enumerar: **a)** de términos: a.1) tiene el mismo nombre, pero su significado es ligeramente distinto –p.ej., el término atributo-; a.2) tiene distinto nombre, pero su significado es idéntico –p.ej., los términos factor y característica-; **b)** de alcance semántico –p.ej., en ambos anexos informativos el concepto de usabilidad, se expande en subcaracterísticas con algunas diferencias o agregados (ver las tablas 5.1 y 5.2 y/o las fuentes originales); otro ejemplo es el concepto de portabilidad, en donde las diferencias son aún más marcadas: en el anexo de ISO, se informan las siguientes subcaracterísticas *Adaptabilidad, Instalabilidad, Conformidad, Reemplazabilidad*; en tanto que en el del IEEE, se especifican los siguientes subfactores *Independencia de Hardware, Independencia de Software, Instalabilidad, Reusabilidad*

Tabla 5.1 Características y subcaracterísticas de calidad con una descripción sintética conforme al estándar ISO/IEC 9126

Característica	Pregunta Central	Subcaracterística	Pregunta Central
Funcionalidad	Las funciones y propiedades satisfacen las necesidades explícitas e implícitas; esto es, el qué ... ?	Adecuación	Tiene el conjunto de funciones apropiadas para las tareas especificadas?
		Exactitud	Hace lo que fue acordado en forma esperada y correcta?
		Interoperabilidad	Interactúa con otros sistemas especificados?
		Conformidad	Está de acuerdo con las leyes o normas y estándares, u otras prescripciones?
		Seguridad de Acceso	Previene accesos no autorizados a los datos y programas?
Confiabilidad	Puede mantener el nivel de rendimiento, bajo ciertas condiciones y por cierto tiempo?	Nivel de Madurez	Con qué frecuencia presenta fallas por defectos o errores?
		Tolerancia a fallas	Si suceden fallas, como se comporta en cuanto a la performance especificada?
		Recuperabilidad	Es capaz de recuperar datos en caso de fallas?
Usabilidad	El software, es fácil de usar y de aprender?	Comprensibilidad	Es fácil de entender y reconocer la estructura y la lógica y su aplicabilidad?
		Facilidad de Aprender	Es fácil de aprender a usar?
		Operabilidad	Es fácil de operar y controlar?
Eficiencia	Es rápido y minimalista en cuanto a uso de recursos, bajo ciertas condiciones?	Comportamiento con respecto al Tiempo	Cuál es el tiempo de respuesta y performance en la ejecución de la función?
		Comportamiento con respecto a Recursos	Cuántos recursos usa y durante cuánto tiempo?
Mantenibilidad	Es fácil de modificar y testear ?	Analisisabilidad	Es fácil diagnosticar una falla o identificar partes a modificar?
		Modificabilidad	Es fácil de modificar y adaptar?
		Estabilidad	Hay riesgos o efectos inesperados cuando se realizan cambios?
		Testeabilidad	Son fáciles de validar las modificaciones?
Portabilidad	Es fácil de transferir de un ambiente a otro?	Adaptabilidad	Es fácil de adaptar a otros entornos con lo provisto?
		Instalabilidad	Es fácil de instalar en el ambiente especificado?
		Conformidad	Adhiere a los estándares y convenciones de portabilidad?
		Reemplazabilidad	Es fácil de usarlo en lugar de otro software para ese ambiente?

Tabla 5.2 Factores y subfactores con una descripción sintética conforme al anexo A (informativo) del estándar IEEE 1061

Característica	Pregunta Central	Subcaracterística	Pregunta Central
Eficiencia	Es rápido y minimalista en cuanto a uso de recursos, bajo ciertas condiciones ?	Economía con respecto al Tiempo	Realizar funciones especificadas bajo ciertas condiciones y en el marco de tiempo apropiado?
		Economía con respecto a Recursos	Realiza funciones especificadas bajo ciertas condiciones usando cantidad apropiada de recursos?
Funcionalidad	Las funciones y propiedades satisfacen todas las necesidades explícitas e implícitas de los usuarios?	Compleitud	Posee las necesarias y suficientes funciones para satisfacer a los requerimientos del usuario?
		Correctitud	Especifica todas las funciones?
		Seguridad	Detecta y previene uso ilegal, destrucción de recursos, pérdida o filtración de información?
		Compatibilidad	El nuevo software se puede instalar sin cambiar el ambiente y las condiciones?
		Interoperabilidad	Se conecta y opera fácilmente con otros sistemas?
Mantenibilidad	Es fácil de modificar y testear ?	Correctibilidad	Es fácil corregir errores y tratar con las demandas de usuarios?
		Expandibilidad	Es fácil de mejorar y modificar la eficiencia de las funciones?
		Testeabilidad	El software, es fácil de testear?
Portabilidad	Es fácil de transferir de un ambiente a otro?	Independencia de Hardware	El software, depende de entornos de hardware específicos?
		Independencia de Software	El software, depende de entornos de software específicos?
		Instalabilidad	Es fácil de ajustar el software para ese nuevo entorno?
		Reusabilidad	El software es fácil de reusar en otras aplicaciones?
Confiabilidad	Puede mantener el nivel de rendimiento, bajo ciertas condiciones y por cierto tiempo?	No-deficiencia	No contiene errores?
		Tolerancia a Errores	Si suceden fallas, como se comporta en cuanto a la performance especificada? Posee funciones de recuperación?
		Disponibilidad	Se mantiene operable en presencia de fallas del sistema?
Usabilidad	El software, es fácil de usar y de aprender?	Comprensibilidad	Es fácil de comprender?
		Facilidad de Aprender	Se minimiza el esfuerzo para comprender el software?
		Operabilidad	La operación está conforme al objetivo, contexto, y factores ergonómicos como color, forma?
		Nivel de Comunicación	Se diseña el software conforme a las características psicológicas de los usuarios?

Finalmente, el estándar internacional ISO está destinado para ser aplicado en la definición de requerimientos de calidad y evaluación de productos de software en general. Puede ser usado por las partes intervinientes en la adquisición, desarrollo, uso, soporte, mantenimiento y auditoría de software. Actualmente, en el seno de la organización ISO, está en borrador el estudio de un nuevo marco conceptual de modelo de calidad. Este será el tema de la siguiente subsección.

5.4.2.1 El Modelo de Calidad Actualizado en el Estándar ISO 9126-1⁴. La revisión del estándar 9126 antes descrito, comenzó a mediados de 1994, y aún hoy, es un trabajo en progreso. Retiene las mismas seis características de alto nivel que la versión previa, aunque ahora *prescribe* semejantes subcaracterísticas a un segundo nivel, preservando los mismos nombres y conceptos (detallados sucintamente en la tabla 5.1) con algunos agregados. Particularmente, cada característica cuenta con la subcaracterística *Conformidad* (o *Compliance*) y, para el caso de *Usabilidad*, se le ha agregado la subcaracterística *Grado de Atracción* (o *Attractiveness*). De manera que el modelo de calidad y el mecanismo de descomposición de características y subcaracterísticas es semejante a la versión de 1991.

En cuanto a la definición de las características es más abarcativa, en tanto que la parte de la frase de inicio, que decía: “*A set of attributes that bear on...*” (ver sección 5.4.2) ha sido modificada por “*The capability of the software product to...*” permitiendo una interpretación en términos de calidad interna y externa del producto a ser medido. Pero además, la revisión del estándar introduce nuevos conceptos. Por ejemplo, las subcaracterísticas pueden ser medidas por métricas internas o externas, definiendo los conceptos de calidad interna, externa y calidad en uso. Una métrica interna de un atributo (interno) es un valor numérico que siempre involucra al ente en sí, ya sea obtenido por una métrica directa o indirecta; en cambio, una métrica externa es el valor resultante de aplicar una métrica indirecta y siempre involucra al ente y su comportamiento con el entorno. Por otra parte, las métricas de calidad de uso miden el grado con que un producto cumplimenta las necesidades de usuarios específicos para lograr las metas con efectividad, productividad, seguridad y satisfacción.

Los documentos de trabajo o borradores 9126-2 y 9126-3 proveen un conjunto sugerido de métricas externas e internas respectivamente, a ser usado con el modelo de calidad descrito en el 9126-1. Además, en el draft 9126-4 se están definiendo las métricas correspondientes al concepto de calidad de uso.

Un marco conceptual del modelo de calidad considerando diferentes visiones de la misma y sus relaciones, es como se aprecia en la figura 5.5 (extractado del documento

⁴ El nuevo estándar ISO/IEC 9126, consta de cuatro documentos [ISO/IEC 9126-1, 9126-2, 9126-3, 9126-4], los cuales al momento de esta escritura, están en modo borrador (draft).

ISO/IEC 9126-1).

Como se observa a la izquierda de la figura, la evaluación del producto de software es uno de los procesos en el ciclo de vida. La calidad del software puede ser evaluada midiendo atributos internos (típicamente, medidas estáticas de productos intermedios), o puede ser evaluada midiendo atributos externos (típicamente, medidas del comportamiento del código cuando se está ejecutando). Además, el objetivo de un producto es que tenga el efecto requerido en un contexto de uso particular.

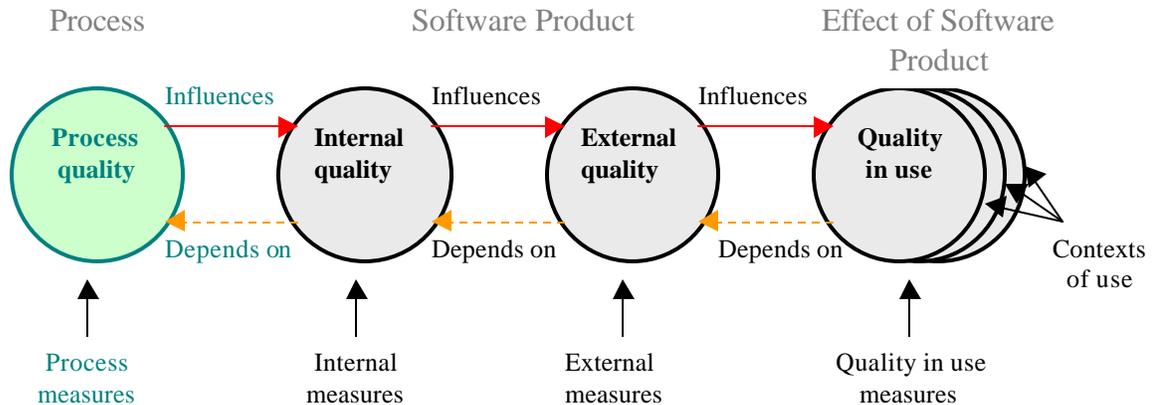


Figura 5.5 Marco Conceptual para el Modelo de calidad especificado en ISO/IEC 9126-1

Conforme al marco conceptual de la figura, la calidad de un proceso (es decir, como aquellos descritos en un modelo de proceso como en [Boehm 88, ISO/IEC 12207], entre otros) contribuye a mejorar la calidad del producto, y, a su vez, la calidad del producto contribuye a mejorar la calidad en uso. Por lo tanto, evaluando y mejorando los procesos es un medio para mejorar la calidad de los productos, y valorando y mejorando la calidad del producto es un medio para mejorar la calidad en uso. Similarmente, evaluando la calidad en uso se puede proveer retroalimentación para mejorar al producto y, evaluando al producto se puede proveer retroalimentación para mejorar al proceso.

De este modo, se requiere de procesos de evaluación de calidad apropiados para que soporten la tarea de medición durante las fases de desarrollo y mantenimiento. Por otra parte, atributos internos del producto apropiados son un prerequisite para alcanzar el comportamiento externo requerido, y de la misma manera, un comportamiento externo apropiado es un prerequisite para alcanzar la calidad de uso.

Finalmente, la calidad de un producto de software debe ser evaluada usando un modelo de calidad definido. El modelo de calidad se debe utilizar cuando se establecen las metas de evaluación para un producto o productos intermedios. Tanto la calidad de

software como la calidad en uso se deben descomponer jerárquicamente en un modelo conformado en base a características y subcaracterísticas.

5.4.3 Seleccionando el Enfoque de Modelo de Calidad

Además, es importante comentar las observaciones realizadas por Fenton y Pfleeger [Fenton et al 97], en lo que respecta a los dos posibles enfoques a adoptar al momento de modelar la calidad de software, a saber:

- ✓ *El enfoque de modelo fijo de calidad*: en la que se asume que todas las características (o factores) y demás componentes necesarios para monitorear un proyecto, son un subconjunto de aquellos publicados en los modelos bien conocidos (McCall, Boehm, ISO 9126, IEEE 1061)
- ✓ *El enfoque “Defina su propio modelo de calidad”*: en la que se acepta la concepción general que la calidad está compuesta de muchas características y atributos, pero en donde no se adopta un modelo de calidad establecido. Por el contrario, el modelo se define por consenso con todas las partes intervinientes, para un proyecto y producto dado. En conjunto, se define el mecanismo de descomposición (posiblemente construido en base a un modelo existente), y se acuerdan las características, atributos (y los criterios de medición) y las relaciones entre atributos, subcaracterísticas y métricas (esta estrategia pragmática ha sido encabezada por Tom Gilb [Gilb 88] y Kitchenham & Walker [Kitchenham et al 89]). Por otra parte, es oportuno considerar al modelo GQM como una herramienta de utilidad para este enfoque.

Por último, dado el estado del arte en modelos de calidad de software, en donde todavía no se ha obtenido un amplio consenso a nivel de subcaracterísticas más allá del segundo nivel y sus relaciones para el producto software, y, dadas algunas de las características distintivas del producto Web respecto de las aplicaciones tradicionales (que discutimos en el capítulo 2), es que consideramos conveniente introducir la noción de un enfoque de modelo mixto.

La noción de un *enfoque de modelo de calidad mixto* es a la vez, pragmático y flexible. Se parte de un modelo fijo (en nuestro caso, a partir del modelo de calidad ISO 9126) en la que se asume que todas las características necesarias para monitorear un proyecto de evaluación de calidad de producto, son un subconjunto de las seis características publicadas en el estándar (ver las características y subcaracterísticas comentadas en la tabla 5.1). Por otra parte, a nivel de subcaracterísticas (segundo nivel) se definen por consenso entre evaluadores y demás partes intervinientes (posiblemente construido en consideración de subcaracterísticas prescriptas en el estándar para cada característica). En conjunto, se define el mecanismo de descomposición restante, y se acuerdan las

subcaracterísticas de niveles inferiores, los atributos, (los criterios de medición) y las relaciones entre atributos, subcaracterísticas y características.

5.4.4 Seleccionando Características conforme al Perfil de Usuario

En este paso, los evaluadores deben seleccionar el subconjunto de características de primer nivel, conforme a las metas y el perfil de usuario seleccionado.

Los perfiles de usuario a un alto nivel de abstracción fueron introducidos en la sección 5.3. Estos perfiles son, a saber: visitantes, desarrolladores, y gerenciadore. Siguiendo un mecanismo de especialización consideramos a la categoría visitante compuesta de dos clases: visitantes expertos y generales.

Para el proceso de evaluación utilizado en los casos de estudio, la meta general consistió en "*Comprender a la calidad global de un sitio Web desde el punto de vista del visitante general*". Para este perfil de usuario⁵ y para esta meta, de las seis características prescriptas en el estándar ISO solamente cuatro de ellas resultaron de relevancia para el proceso de evaluación. Las características de significativa importancia fueron *usabilidad, funcionalidad, confiabilidad y eficiencia*. Para tal fin, se realizó un estudio de campo, con tres grupos de estudiantes con conocimiento de los conceptos y uso de sitios Web para determinar la relativa importancia de las características para el perfil de usuario citado y para sitios académicos. Los estudiantes eran del cuarto año de la carrera Analista Programador en Computación, de la Fac. de Ingeniería de la UNLPam, luego de haber cursado la materia Laboratorio. Primeramente, se les informó de los conceptos de cada característica y de los conceptos de las potenciales subcaracterísticas intervinientes. Por medio de una encuesta cerrada se determinó la importancia promedio que le otorgaron a cada una de las seis características a partir de una escala de 1 a 10. El valor se determinó mediante el cociente:

Peso relativo_Caract_n = Prom(Caract_n) / Suma de todos los Promedios.

A su vez, la suma de todos los pesos relativos de todas las características quedó normalizada a uno. Este estudio corroboró la hipótesis de la mínima importancia de las características *Mantenibilidad* y *Portabilidad* para el perfil dado. El mayor peso fue dado para las características *Usabilidad* y *Funcionalidad* y en menor grado para *Eficiencia* y *Confiabilidad*, con pesos semejantes (de este estudio y del consenso de los evaluadores surgieron los pesos que participaron en el cómputo de la preferencia global, como se puede ver en la figura 7.8). Encuestas similares se realizaron para determinar la relativa importancia de subcaracterísticas y atributos (aunque no fue realizada para todos los elementos).

⁵ Particularmente, el perfil de usuario seleccionado para el estudio de sitios académicos fue el del visitante general, en consideración de estudiantes actuales y futuros, de grado y posgrado.

5.5 Hacia un Arbol de Requerimientos de Calidad General para dominios Web

Un aspecto importante, que si bien hemos avanzado en parte aún es motivo de futura investigación, reside en determinar un árbol estándar que pueda ser reusado y personalizado para distintos subdominios de aplicaciones, dentro del dominio de la Web. Contar con un árbol de requerimientos general pero a su vez fácilmente adaptable a diversos dominios, sería de gran beneficio para tareas de aseguramiento y control de calidad posibilitando potencialmente la automatización de las mismas [Olsina et al 99d].

En la siguiente tabla podemos observar las semejanzas entre los modelos de calidad establecidos en ambos estándares, aunque con ligeras diferencias a nivel de subcaracterísticas, como apuntábamos en la sección 5.4.2.

En la versión ISO (de 1991) está ausente la idea de nivel de comunicación, refiriéndose al grado con que el artefacto software es diseñado de acuerdo con características psicológicas de los usuarios, incluyendo aspectos ergonómicos como color, sonido, estilo, entre otros. Usabilidad es una característica esencial en las aplicaciones Web, debido al carácter interactivo, al carácter de orientación a documentación y publicación, y, en definitiva de orientación al usuario que posee [Nielsen 93, Nielsen_Alert, Rosenfeld et al 98]. En la versión reciente de ISO, se ha introducido la subcaracterística *Grado de Atracción* la cual expresa el grado en que el usuario le agrada el producto software, por ejemplo, en operación, es decir “*the capability of the software product to be attractive to the user*” [ISO/IEC 9126-1].

Tabla 5.3 Descomposición de subcaracterísticas para Usabilidad, según se observa para los tres modelos.

ISO 9126	IEEE 1061	Modelo de Calidad para la Web
Título: Usabilidad	Título: Usabilidad	Título: Usabilidad
Tipo: Característica	Tipo: Característica	Tipo: Característica
Sub-característica/s:	Sub-característica/s:	Sub-característica/s:
Comprensibilidad	Comprensibilidad	Comprensibilidad (Global del Sitio)
Facilidad de Aprender	Facilidad de Aprender	Mecanismo de Ayuda y Retroalimentación
Operabilidad	Operabilidad	Aspectos de Interfaces
Grado de Atracción ⁶	Nivel de Comunicación	Aspectos Estéticos y de Estilo

Nuestra categorización de subcaracterísticas para *Usabilidad*, se basó en parte en las guías establecidas, pero además, debido a aspectos distintivos del artefacto Web y debido, por otra parte, a que los estudios estuvieron centrados en dominios específicos antes que generales, se tuvo en cuenta necesidades y comportamientos para el perfil de usuario seleccionado, según el enfoque de modelado de calidad mixto antes descripto.

⁶ Introducido en la versión draft del estándar ISO/IEC 9126-1

Sin embargo, el lector puede apreciar algunas semejanzas, por ejemplo, con respecto a la subcaracterística *Comprensibilidad* que pasaremos a comentar.

Cuando un visitante entra por primera vez a la página principal de un sitio Web, puede tener en mente encontrar una pieza de información, o no. Hay dos funciones que permiten moverse potencialmente por un sitio: búsqueda y exploración. Para facilitar un modelo mental de usuario efectivo con respecto a todo el sitio Web, tanto de su estructura como de su contenido, es decir, para obtener una comprensión global del sitio, hay subcaracterísticas y atributos como el *Esquema de la Organización Global*, *Visitas Guiadas*, *Calidad del Sistema de Etiquetado*, entre otros, que contribuyen a la *Comprensibilidad* del mismo. Por ejemplo, la subcaracterística *Esquema de la Organización Global* puede contener atributos como *Tabla de Contenidos*⁷, *Indices* (de distinto tipo, como cronológico, alfabético, temático, geográfico, mixtos) y *Mapa del Sitio*, que favorecen, por una parte, obtener una visión rápida de la estructura y, por otra parte, la navegación y exploración directa a páginas con contenido de información. Además, contar con una visita guiada, puede brindar rápidamente a la audiencia un panorama del contenido, sin tener que navegar por los distintos subsitios o contextos de navegación.

En el párrafo anterior hemos introducido la idea de distintos niveles de descomposición (más allá del segundo), y hemos citado algunos atributos. El concepto de atributo en nuestro trabajo es el de una característica elemental de un ente (que no soporta más descomposición, esto es, la hoja del árbol), y al que se lo puede medir mediante la aplicación de una métrica directa (o en algunos casos, indirectas, como discutiremos en el capítulo 10).

Finalmente, muchos de los atributos y subcaracterísticas citados y que veremos en el ejemplo siguiente, se pueden considerar suficientemente generales y válidos para diversos dominios de aplicaciones Web. No obstante, se requiere profundizar esta línea de investigación.

5.6 Características, Subcaracterísticas y Atributos en la Web: un Ejemplo de Arbol de Requerimientos de Calidad.

En esta sección, mostraremos el árbol de requerimientos de calidad, correspondiente al dominio Web de una unidad académica como lo es una universidad. Específicamente, presentaremos los requerimientos de calidad utilizados para especificar el caso de estudio de sitios académicos [Olsina 99 a, c].

⁷ En los capítulos sucesivos –no en éste- iremos definiendo a los atributos. Además, presentaremos una plantilla o marco conceptual para especificar criterios de evaluación, escalas, etc.

3. Confiabilidad

3.1 No Deficiencia

- 3.1.1 Errores de Enlaces
 - 3.1.1.1 *Enlaces Rotos*
 - 3.1.1.2 *Enlaces Inválidos*
 - 3.1.1.3 *Enlaces no Implementados*
- 3.1.2 Errores o Deficiencias Varias
 - 3.1.2.1 *Deficiencias o cualidades ausentes debido a diferentes navegadores (browsers)*
 - 3.1.2.2 *Deficiencias o resultados inesperados independientes de browsers (p.ej. errores de búsqueda imprevistos, deficiencias con marcos (frames), etc.)*
 - 3.1.2.3 *Nodos Destinos (inesperadamente en Construcción)*
 - 3.1.2.4 *Nodos Web Muertos (sin enlaces de retorno)*

4. Eficiencia

4.1 Performancia

- 4.1.1 *Páginas de Acceso Rápido*

4.2 Accesibilidad

- 4.2.1 Accesibilidad de Información
 - 4.2.1.1 *Soporte a Versión sólo Texto*
 - 4.2.1.2 Legibilidad al desactivar la Propiedad Imagen del Browser
 - 4.2.1.2.1 *Imagen con Título*
 - 4.2.1.2.2 *Legibilidad Global*
- 4.2.2 Accesibilidad de Ventanas
 - 4.2.2.1 *Número de Vistas considerando Marcos (frames)*
 - 4.2.2.2 *Versión sin Marcos*

Figura 5.6 Arbol de requerimientos de calidad para el caso de estudio de sitios académicos, especificando dos de la cuatros características de alto nivel, a saber: *Confiabilidad* y *Eficiencia*.

En la figura 5.6 se consideran subcaracterísticas y atributos derivados de las características de más alto nivel denominadas *Confiabilidad* y *Eficiencia*; asimismo, y conforme a los estándares y enfoques previamente discutidos, en la figura 5.7 se detallan subcaracterísticas y atributos derivados de las características *Usabilidad* y *Funcionalidad*. Alrededor de ciento veinte características y atributos fueron modelados. Estas características prescriptas a un alto nivel de abstracción, les da a los evaluadores un marco conceptual para especificar requerimientos de calidad proveyendo una base firme para posteriores refinamientos. Aplicando el mecanismo de descomposición descripto, las características se pueden descomponer en múltiples niveles de subcaracterísticas, hasta arribar a las hojas del árbol, es decir, a un conjunto de atributos cuantificables. De este modo, la característica *Funcionalidad* se descompone en tres sub-características, dos bastante independientes del dominio como son *Aspectos de Búsqueda* y *Recuperación*, *Aspectos de Navegación* y *Exploración* y una tercera dependiente, denominada *Aspectos del Dominio orientados al Estudiante*.

De la misma manera, un mecanismo de descomposición semejante se aplicó para *Confiabilidad* y *Eficiencia*. Por ejemplo, a la característica *Eficiencia* se la dividió en los subfactores *Performancia* y *Accesibilidad*.

Concentrándonos en *Aspectos del Dominio orientados al Estudiante* se observan dos subcaracterísticas principales *Relevancia de Contenido* y *Servicios On-line*. Como el lector puede apreciar, en este estudio se evaluaron aspectos que van desde información de unidades académicas; información de inscripción, de carreras y cursos; información de infraestructura académica, hasta servicios en línea como información de aprobación de cursos y listados; servicios de páginas Web, FTP y grupo de noticias

1. Usabilidad

1.1 Comprensibilidad Global del Sitio

- 1.1.1 Esquema de Organización Global
 - 1.1.1.1 *Mapa del Sitio*
 - 1.1.1.2 *Tabla de Contenidos*
 - 1.1.1.3 *Indice Alfabético*
- 1.1.2 *Calidad en el Sistema de Etiquetado*
- 1.1.3 *Visita Guiada Orientada al Estudiante*
- 1.1.4 *Mapa de Imagen (Campus/Edificio)*

1.2 Mecanismos de Ayuda y Retroalimentación en línea

- 1.2.1 Calidad de la Ayuda
 - 1.2.1.1 *Ayuda Explicatoria Orientada al Estudiante*
 - 1.2.1.2 *Ayuda de la Búsqueda*
- 1.2.2 Indicador de Última Actualización
 - 1.2.2.1 *Global (de todo el sitio Web)*
 - 1.2.2.2 *Restringido (por sub-sitio o página)*
- 1.2.3 Directorio de Direcciones
 - 1.2.3.1 *Directorio E-mail*
 - 1.2.3.2 *Directorio TE-Fax*
 - 1.2.3.3 *Directorio Correo Postal*
- 1.2.4 *Facilidad FAQ*
- 1.2.5 Retroalimentación
 - 1.2.5.1 *Cuestionario*
 - 1.2.5.2 *Libro de Invitados*
 - 1.2.5.3 *Comentarios/Sugerencias*

1.3 Aspectos de Interfaces y Estéticos

- 1.3.1 *Cohesividad al Agrupar los Objetos de Control Principales*
- 1.3.2 Permanencia y Estabilidad en la Presentación de los Controles Principales
 - 1.3.2.1 *Permanencia de Controles Directos*
 - 1.3.2.2 *Permanencia de Controles Indirectos*
 - 1.3.2.3 *Estabilidad*
- 1.3.3 Aspectos de Estilo
 - 1.3.3.1 *Uniformidad en el Color de Enlaces*
 - 1.3.3.2 *Uniformidad en el Estilo Global*
 - 1.3.3.3 *Guía de Estilo Global*
- 1.3.4 *Preferencia Estética*

1.4 Misceláneas

- 1.4.1 *Soprote a Lenguaje Extranjero*
- 1.4.2 *Atributo "Qué es lo Nuevo"*
- 1.4.3 *Indicador de Resolución de Pantalla*

2. Funcionalidad

2.1 Aspectos de Búsqueda y Recuperación

- 2.1.1 Mecanismo de Búsqueda en el Sitio Web
 - 2.1.1.1 Búsqueda Restringida
 - 2.1.1.1.1 *de Personas*
 - 2.1.1.1.2 *de Cursos*
 - 2.1.1.1.3 *de Unidades Académicas*
 - 2.1.1.2 *Búsqueda Global*
- 2.1.2 Mecanismos de Recuperación
 - 2.1.2.1 *Nivel de Personalización*
 - 2.1.2.2 *Nivel de Retroalimentación en la Recuperación*

2.2 Aspectos de Navegación y Exploración

2.2.1 Navegabilidad

- 2.2.1.1 Orientación
 - 2.2.1.1.1 *Indicador del Camino*
 - 2.2.1.1.2 *Etiqueta de la Posición Actual*
- 2.2.1.2 *Promedio de Enlaces por Página*

2.2.2 Objetos de Control Navegacional

- 2.2.2.1 Permanencia y Estabilidad en la Presentación de los Controles Contextuales (Subsitio)
 - 2.2.2.1.1 *Permanencia de los Controles Contextuales*
 - 2.2.2.1.2 *Estabilidad*
- 2.2.2.2 Nivel de Desplazamiento
 - 2.2.2.2.1 *Desplazamiento Vertical*
 - 2.2.2.2.2 *Desplazamiento Horizontal*

2.2.3 Predicción Navegacional

- 2.2.3.1 *Enlace con Título (enlace con texto explicatorio)*
- 2.2.3.2 *Calidad de la Frase del Enlace*

2.3 Aspectos del Dominio orientados al Estudiante

2.3.1 Relevancia de Contenido

- 2.3.1.1 Información de Unidad Académica
 - 2.3.1.1.1 *Indice de las Unidades*
 - 2.3.1.1.2 *Sub-sitios de las Unidades*
- 2.3.1.2 Información de Inscripción
 - 2.3.1.2.1 *Información de los Requerimientos de Ingreso/Admisión*
 - 2.3.1.2.2 *Formulario para Rellenar/Bajar*
- 2.3.1.3 Información de Carreras
 - 2.3.1.3.1 *Indice de Carreras*
 - 2.3.1.3.2 *Descripción de Carrera*
 - 2.3.1.3.3 *Plan de Carrera/Oferita de Cursos*
 - 2.3.1.3.4 Descripción de Cursos
 - 2.3.1.3.4.1 *Comentarios*
 - 2.3.1.3.4.2 *Programa Sinté./Anal.*
 - 2.3.1.3.4.3 *Programación Cursos*

2.3.1.4 Información de Servicios al Estudiante

- 2.3.1.4.1 *Indice de Servicios*
- 2.3.1.4.2 *Información de Salud*
- 2.3.1.4.3 *Información de Becas*
- 2.3.1.4.4 *Información de Residencias*
- 2.3.1.4.5 *Información Cultural/Deport.*

2.3.1.5 Información de Infraestructura Académica

- 2.3.1.5.1 *Información de Bibliotecas*
- 2.3.1.5.2 *Información de Laboratorios*
- 2.3.1.5.3 *Información Resultados I+D*

2.3.2 Servicios On-line

- 2.3.2.1 *Información Aranceles, Aprobación de Cursos.*
- 2.3.2.2 *Servicio de Páginas Web*
- 2.3.2.3 *Servicio FTP*
- 2.3.2.4 *Servicio de Grupo de Noticias*

Figura 5.7 Arbol de Requerimientos de Calidad especificando dos de las cuatro características de alto nivel, a saber: Usabilidad y Funcionalidad

Por último, dado un proceso de evaluación y en consideración de un dominio Web, podríamos ver fácilmente que no necesariamente todos los atributos deben existir simultáneamente; podría ser necesario un *Mapa del Sitio*, o una *Tabla de Contenidos*, o un *Índice*. Es más, distintos tipos de índice podrían ser reemplazables dependiendo de los requerimientos específicos. Los índices orientados a la temática o alfabéticos podrían ser mejores en ciertas circunstancias que un índice cronológico; además, podría especificarse más de uno tipo de índice. (Web-site QEM permite modelar relaciones de simultaneidad y de reemplazabilidad tomando en cuenta los pesos relativos y niveles de polarización and/or, como se verá en el capítulo 7).

Del mismo modo que la relación de reemplazabilidad arriba mencionada, se podría modelar a la relación de simultaneidad; por ejemplo, en la subcaracterística *Aspectos de Búsqueda y Recuperación* (ver en la fig. 5.7, la característica *Funcionalidad*). Para un visitante dado, se podría requerir dos tipos de búsqueda, a saber: *Búsqueda Restringida* (o “scoped search” [Nielsen_Alert]) y *Búsqueda Global*; es decir, podría ser necesario una búsqueda personalizada para investigar por palabras claves aspectos de una colección de un museo, o cursos de una carrera, en tanto que también podría ser necesaria una búsqueda global para encontrar documentos o registros que pudieran estar en cualquier subsitio de todo el sitio Web.

5.7 Plantilla de Referencia de Variables y Parámetros

El proceso de determinación de requerimientos culmina, por una parte, con un documento que jerárquicamente lista a todas las características y atributos cuantificables y que modelan a la calidad según las necesidades específicas del perfil de usuario (figuras 5.6 y 5.7). Por otra parte, los requerimientos de calidad quedarán completos, luego de acordar un conjunto de valores, variables, parámetros y criterios para cada atributo cuantificable A_i (esta actividad puede realizarse en un estilo incremental e iterativo con las actividades de la siguiente fase, de determinación de los criterios elementales).

En este punto de la discusión, es oportuno destacar que la disponibilidad de parámetros, y valores de referencia o línea base estándar para métricas de Hipermedia y en la Web son, en la mayoría de los casos, inexistentes. Esta investigación debe ayudar a establecer valores actuales de características y atributos, que sirvan de referencia para posteriores estudios en el área.

Por una parte, luego de computar la función de preferencia elemental que modela el requerimiento del atributo A_i (a partir de la métrica $m: A_i \rightarrow X_i$), el valor de preferencia caerá en uno de los tres niveles de aceptabilidad o barras de calidad, esto es, *insatisfactorio* (de 0 a 40%), *marginal* (desde 40 a 60%), y *satisfactorio* (desde 60 a 100%). Este es el esquema de categorías de aceptabilidad que adoptamos por consenso

y observando otros modelos como ISO e IEEE.

La figura 5.8 muestra los diferentes rangos de satisfacción de requerimientos de usuario según ISO 9126. Con respecto al estándar IEEE 1061, sólo se indica que los valores de las métricas son usados para clasificar al software en las categorías de “aceptable”, “marginal” e “inaceptable”.

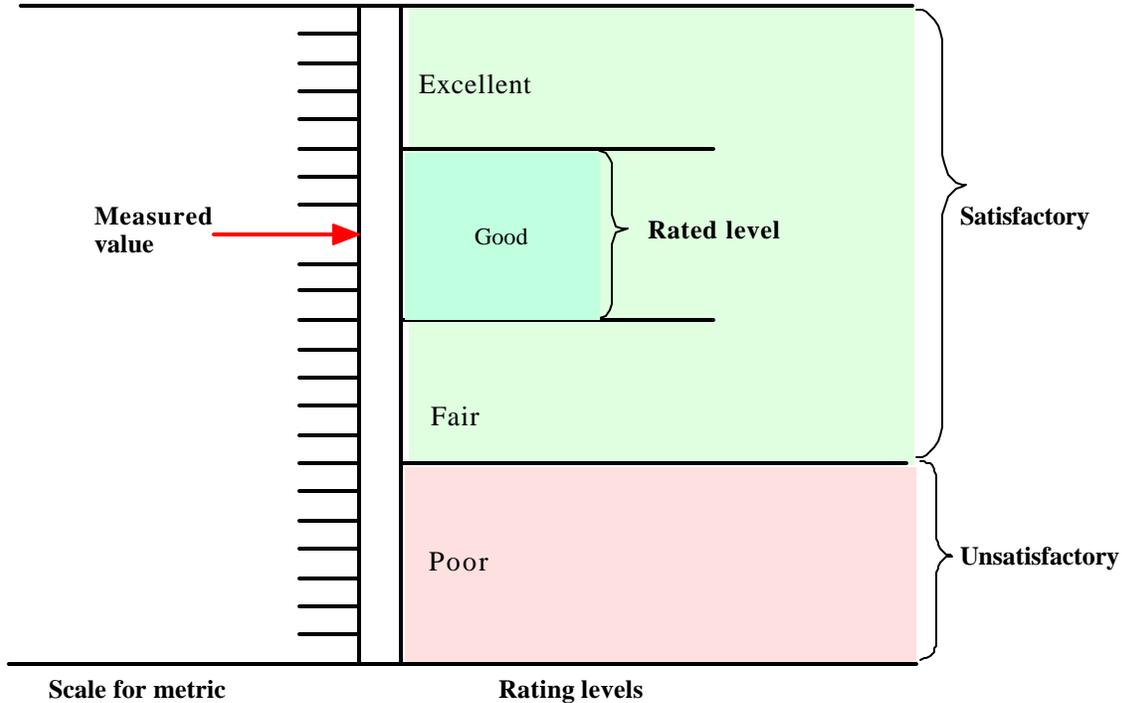


Figura 5.8 Valor medido y niveles de aceptabilidad, según ISO 9126.

Por otra parte, el valor de la preferencia elemental se obtiene a partir de una función que responde a un criterio. Existen diversos tipos de criterios elementales como veremos en la sección 6.3. Lo importante es que el criterio debe satisfacer la necesidad del requerimiento de evaluación en consideración de factores de mayor o menor criticidad, riesgo, complejidad del atributo, costos, entre otros [ISO/IEC 14598-5]. Además, para ayudar en la determinación del criterio para el requerimiento A_i y su mayor o menor precisión, es necesario (o sería deseable) conocer, unidades de medición, valores corrientes, valor planificado, el mejor y el peor valor esperado. Podemos utilizar para tal fin una *Plantilla de Referencia de Variables y Parámetros* (semejante a las Plantillas de Calidad –o “Quality Templates”– propuesta por [Gilb 88]) como una herramienta de utilidad para especificar información deseada de cada atributo de calidad interviniente.

Por ejemplo, para el atributo 4.1.1 (*Páginas de Acceso Rápido*) hemos especificado la siguiente Plantilla de Referencia de Calidad (ver tabla 5.4) en consideración del caso de estudio de sitios académicos (el esquema propuesto está pensado para ser reusable). Asimismo, en la tabla 5.5 especificamos la plantilla correspondiente al atributo *Soporte a Lenguaje Extranjero*, codificado como 1.4.1 en el árbol de requerimientos para el caso

de estudio de museos en la Web (ver figura 9.1).

Tabla 5.4 Ejemplo de Plantilla de Referencia de Variables y Parámetros para el atributo **Páginas de Acceso Rápido**

Aspecto	Descripción	Valores, Parámetros, Ejemplos
Código y Nombre del Atributo	El código debe ser único, en correspondencia con el árbol de requerimientos citado	4.1.1 (figura 5.6) Páginas de Acceso Rápido
Definición	Mide el tamaño de todas las páginas (estáticas) del sitio Web considerando todos sus componentes gráficos, tabulares y textuales. El tamaño de cada página se especifica como una función del tiempo de espera y de la velocidad constante establecida. Se especifican categorías de tiempo de espera.	Ejemplo: se especifica como tamaño umbral aceptable de tamaño de cada página, el de 35,2 Kb. Una página de este tamaño requiere 20 segundos para ser bajada a una taza de 14.400 bps
Fecha de Medición	Se especifica la fecha (o rango entre dos fechas), en la que se realizará la recolección de datos (si corresponde) y se computará la variable de calidad para los parámetros planificados.	F=[dd-mm-aaaa] o desde F ₁ a F ₂
Parámetros (y Variables) Disponibles	Velocidad	[bps] 14400; 19600; 28800
	Rango de tiempo aceptable para que un usuario no se impacienta [IEEE WPG]	[segundos] 0 < t ≤ 20
	Rango de tamaños máximos aceptables de cada página para un tiempo aceptable (extraído el 27-Ene-1999 de [IEEE WPG])	[Kbytes] 30 < s < 45
Parámetros Planificados	Velocidad	[bps] 14400
	Tiempo aceptable Tiempo poco aceptable Tiempo rara vez aceptable	[segundos] 0 < t ₁ ≤ 20 20 < t ₂ ≤ 40 t ₃ > 40
	Tamaño máximo de cada página para un tiempo aceptable	[Kbytes] 35,2
Mayor grado de Satisfacción	El mayor valor de X, dentro de los números reales, obtenido de computar la función elemental y que se traducirá en una preferencia de calidad del 100%	Si X = 100 → IE ₁ = 100%
Función elemental	$X = ((X_1 - 0.4 X_2 - 0.8 X_3) / (X_1 + X_2 + X_3)) * 100$	X ₁ representa el número de páginas dentro de un tiempo que va desde 0 < t ₁ ≤ 20 seg. ; y X ₂ representa el número de páginas dentro de un tiempo que va desde 20 < t ₂ ≤ 40, y X ₃ representa el número de páginas dentro de un tiempo de espera donde: t ₃ > 40 seg.
Referencias	Distintas fuentes de referencia	Ver la Plantilla del atributo del mismo código en sección 6.4.3 para el caso de estudio de sitios académicos

Comentarios Rangos de Aceptabilidad de Preferencia de Calidad	Satisfactorio	[%] 60 < IE <= 100
	Insatisfactorio	[%] 0 < IE <= 40
	Marginal	[%] 40 < IE <= 60

Tabla 5.5 Ejemplo de Plantilla de Referencia de Variables y Parámetros para el atributo *Soporte a Lenguaje Extranjero*

Aspecto	Descripción	Valores, Parámetros, Ejemplos
<i>Código y Nombre del Atributo</i>	El código debe ser único, en correspondencia con el árbol de requerimientos citado	1.4.1 (figura 9.1) Soporte a Lenguaje Extranjero
<i>Definición</i>	Este atributo modela el número de lenguajes naturales extranjeros soportados por sitios de museos típicos, y, además, el nivel de soporte para cada lenguaje. No se computa el lenguaje nativo del sitio web, como lenguaje extranjero.	Ejemplos: el sitio del museo Louvre (al Oct-1998) posee soporte parcial a tres lenguajes extranjeros (japonés, español e inglés). Por otra parte, el sitio del museo del Prado posee soporte total del lenguaje inglés
<i>Fecha de Medición</i>	Se especifica la fecha (o rango entre dos fechas), en la que se realizará la recolección de datos (si corresponde) y se computará la variable de calidad para los parámetros planificados.	F=[dd-mm-aaaa] o desde F ₁ a F ₂
<i>Parámetros (y Variables) Disponibles</i>	Cantidad de Lenguajes Extranjeros, para el nivel de soporte S _i Nivel de soporte, a saber: Total (en todo el sitio); Parcial (algunos subsitios del sitio); Mínimo (algunas páginas o documentos).	N _i S _i i=(1 .. n) para n=3 S ₁ =0.2 -> soporte mínimo; S ₂ =1 -> soporte medio; S ₃ =2 -> soporte total.
<i>Valores Planificados</i>	Se indican algunos casos modelados, no de modo extensivo.	Caso 1) N ₁ = 0; N ₂ = 0; N ₃ = 2 y S ₃ = 2 Caso 2) N ₁ = 0; N ₂ = 1; N ₃ = 1 y S ₂ = 1; S ₃ = 2 Caso 3) N ₁ = 0; N ₂ = 2; N ₃ = 0 y S ₂ = 1 Caso 4) N ₁ = 5; N ₂ = 2; N ₃ = 0 y S ₁ = .2; S ₂ = 1
<i>Mayor grado de Satisfacción</i>	El mayor valor de X, dentro de los números reales, obtenido de computar la función elemental y que se traducirá en una preferencia de calidad del 100%	Si X >= 100 → IE _i = 100%
<i>Función elemental</i>	$X = 30 * \hat{a}_i S_i * N_i$	Si X resulta mayor al valor 100, entonces se computa a X = 100
<i>Referencias</i>	Distintas fuentes de referencia	Ver la Plantilla del atributo del mismo código para el documento con el caso de estudio de museos en la Web (9.2.2)

5.8 Consideraciones Finales

Para muchos casos prácticos, y a partir de requerimientos y usuarios específicos, se puede necesitar medir solamente un par de subcaracterísticas de calidad, utilizando para tal fin, por ejemplo, un enfoque GQM, y/o modelos y procedimientos que son de público conocimiento. Sin embargo, para evaluar, comprender, comparar y establecer un ranquin de artefactos Web, con cualidades de sistemas complejos, afirmamos que es importante una visión integral de la calidad como la que aquí se discute.

De esta manera, para elicitar y especificar el árbol de requerimientos de características, subcaracterísticas y atributos de calidad, para determinar los criterios de medición elementales e implementarlos, para realizar la agregación apropiada para producir un indicador final y global, se requiere una metodología integral y robusta como la propuesta (ver sección 4.1.1). Web-site QEM, es esencialmente una metodología integral, flexible y robusta, y cubre la mayor parte de las actividades del proceso de evaluación y comparación de artefactos Web (como seguiremos viendo en los próximos capítulos).

Con todo, la meta de este proceso de evaluación consiste en comprender a la calidad global de sitios Web típicos, es decir, evaluarlos individualmente y compararlos. El concepto de comparación de sistemas (por ej. artefactos Web) está estrechamente relacionado al concepto de evaluación. Dado que la preferencia de calidad global es un cuantificador del grado de cumplimiento de un conjunto de características y atributos de un artefacto con respecto a los requerimientos de calidad establecidos, comparar y establecer un ranquin será establecer un valor respecto de los indicadores absolutos y globales obtenidos para cada sistema. Es importante notar en este punto (aunque aún no se han discutido con detenimiento los criterios elementales y globales de evaluación y los procedimientos de medición y cómputo), que la comparación absoluta de sistemas, permite simultáneamente responder al menos a estas preguntas:

- ✓ ¿Cuál es el indicador global de calidad de *este* sistema, por ejemplo, para *éste* sitio Web?
- ✓ ¿Cómo se compara globalmente *este* sistema evaluado, con el *otro* sistema (u otros) respecto de los mismos requerimientos?
- ✓ ¿Qué características están mejor diseñadas en éste respecto de aquel otro?
- ✓ ¿En dónde residen las fortalezas y las debilidades en los sistemas comparados?

Por lo tanto, la base de la comparación (y justificación) es el indicador absoluto del sistema (y sus indicadores absolutos parciales y elementales).

La comparación absoluta se diferencia de la comparación relativa en que la meta de ésta última consiste solamente en determinar los indicadores relativos de los sistemas comparados sin evaluar la calidad (o la característica que corresponda) de cada sistema

individual e independiente. Así, una comparación relativa entre sistemas produce un orden o puntaje relativo, que no puede ser interpretado individualmente dado que no representan el grado de cumplimiento de los requerimientos absolutos.

El árbol de requerimientos especificado en las figuras 5.6 y 5.7 y la plantilla como la mostrada en la tabla 5.4 servirá como base para establecer criterios absolutos de evaluación, que discutiremos seguidamente.

Capítulo 6

Fase de Definición e Implementación de la Evaluación Elemental

Con respecto a la fase de *Definición e Implementación de la Evaluación Elemental* la misma trata con actividades, modelos, técnicas, heurísticas y herramientas para determinar criterios de evaluación para cada atributo cuantificable y realizar el proceso de medición. Particularmente, nos interesa la calidad de artefactos Web como característica de estudio. No obstante, otros indicadores como costo, o relación costo-calidad pueden ser de interés para la investigación en evaluación de sistemas. Por ejemplo, en la evaluación de la compra de una computadora, intuitivamente una persona observa el valor del costo final que suman todos los componentes requeridos, y, además, la calidad de los mismos (como atributos de calidad del monitor, tipo del microprocesador –características como velocidad, procesamiento multimedia, etc.-, tipo y capacidad del disco rígido, entre otros.)

En este capítulo, se consideran diferentes tipos de criterios de calidad elemental, escalas (y representación gráfica como escala de preferencia), valores y rangos críticos, y funciones para determinar la preferencia elemental, entre otros asuntos. Una vez definidos y consensuados los criterios para medir cada atributo, se debe ejecutar el proceso de medición, es decir, la recolección de datos, el cómputo de las variables y las preferencias elementales, y la documentación de los resultados.

Este capítulo está destinado a esos tópicos y a la ejemplificación de los mismos.

6.1 Criterio de Evaluación Elemental para Atributos

A partir del árbol de requerimientos (como el de la figura 5.7), para cada atributo cuantificable A_i (u hoja del árbol) debemos asociar y determinar una variable X_i , que tomará un valor real a partir de un proceso de medición (la métrica debe ser válida, como veremos en el capítulo 10). Además, para cada variable X_i computada, por medio de un criterio elemental, producirá una preferencia elemental IE_i . Este resultado final, elemental, se puede interpretar como el grado o porcentaje del requerimiento del usuario satisfecho para el atributo A_i .

Un criterio de evaluación elemental ayuda a comprender y especificar cómo medir atributos cuantificables. Nosotros estamos interesados en la determinación de la preferencia de calidad elemental para cada atributo, de manera que por medio de un proceso de agregación podamos obtener un valor numérico global para el producto a evaluar (y que denominaremos la preferencia de calidad global del producto –esto será

discutido en el capítulo 7)

6.1.1 Criterio de Preferencia de Calidad Elemental

Para cada variable de calidad medida X_i , $i = 1, \dots, n$ se define una función que representa al criterio elemental.

Por definición un criterio elemental es una correspondencia del valor de la variable de calidad X_i en el valor de la preferencia (o indicador) elemental de calidad IE_i . En términos generales, el valor medido de la variable es un número real:

$$X_i \in R_i \subset R$$

El valor de la preferencia de calidad elemental es también un número real pero perteneciente al intervalo unitario I , de manera que:

$$IE_i \in I, \quad i = 1, \dots, n, \quad I = [0, 1]$$

En una interpretación rigurosa la preferencia elemental de calidad es el grado de verdad en la declaración que afirma “el valor de la variable de calidad X_i satisface completamente el requerimiento de calidad del i -ésimo criterio elemental”. Consecuentemente, la preferencia de calidad elemental representa el grado de satisfacción de un requerimiento o necesidad de usuario. Con frecuencia, en vez de usar el intervalo unitario es útil emplear la escala porcentual de $[0, 100\%]$. En este sentido se interpreta a la preferencia como el porcentaje del requerimiento satisfecho.

Desde un punto de vista analítico, el criterio elemental se define como la función:

$$F_i : R_i \rightarrow I \text{ en donde } IE_i = F_i(X_i), \quad X_{i \min} \leq X_i \leq X_{i \max}$$

Sea el tiempo t_i que representa el tiempo de respuesta promedio (por ejemplo, el tiempo de respuesta promedio necesitado por el sistema para realizar un tipo de consulta a una base de datos). Entonces, el criterio elemental para dicho atributo puede ser definido como:

$$\begin{aligned} IE_i &= 1 \text{ (o } 100\%) && \text{si } t_i \leq t_{i \min} \\ IE_i &= 0 \text{ (o } 0\%) && \text{si } t_i \geq t_{i \max} \\ IE_i &= (t_{i \max} - t_i) / (t_{i \max} - t_{i \min}) && \text{si } t_{i \min} \leq t_i \leq t_{i \max} \end{aligned} \quad (1)$$

Según lo expresado, el criterio declara que el evaluador está completamente satisfecho ($IE_i = 100\%$) si el tiempo de respuesta medido para el atributo A_i , es menor o igual al menor (mejor) tiempo $t_{i \min}$. Una situación totalmente insatisfactoria se da ($IE_i = 0\%$), si el tiempo promedio de respuesta de la consulta es más grande o igual al mayor (peor)

tiempo $t_{i \max}$. Si el tiempo cae en un valor intermedio entre $t_{i \min}$ y $t_{i \max}$ el requerimiento es parcialmente satisfecho. Para determinar el valor correspondiente al grado o porcentaje de satisfacción para este último caso, se puede emplear interpolación lineal [Dujmovic 96].

La figura 6.1 muestra una representación gráfica en los ejes cartesianos para el criterio elemental anterior, y la tabla 6.1 representa la plantilla de referencia de variables y parámetros útiles para determinar la variable de calidad en el proceso de medición. La figura 6.2 muestra un criterio elemental para el atributo *Capacidad de Almacenamiento Secundario* (disco rígido) considerando, por ejemplo, un componente de los requerimientos en el proceso de evaluar la compra de una PC.

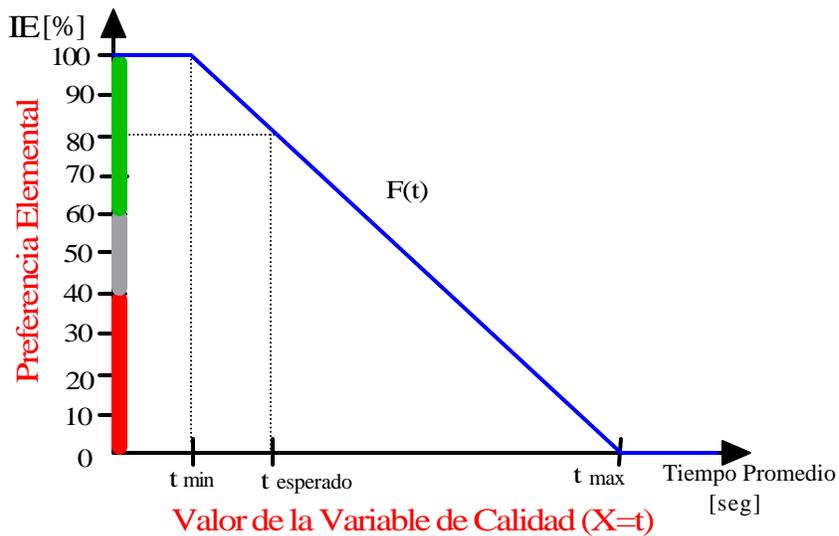


Figura 6.1 Criterio Elemental para el atributo Tiempo de Respuesta Promedio para un tipo de Consulta a BD

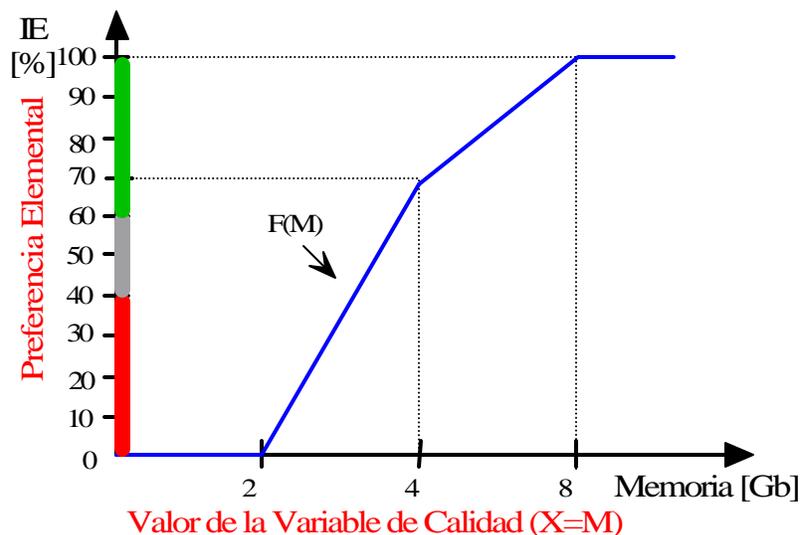


Figura 6.2 Criterio Elemental para el atributo Capacidad de Almacenamiento Secundario (p. ej. para disco rígido)

Tabla 6.1 Plantilla de Referencia de Variables y Parámetros para el atributo Tiempo de Respuesta Promedio (p. ej. en un tipo de Consulta a una BD)¹.

Aspecto	Descripción	Valores, Parámetros, Ejemplos
<i>Código y Nombre del Atributo</i>	El código debe ser único, en correspondencia con el árbol de requerimientos citado	X.X.X Tiempo de Respuesta Promedio
<i>Definición</i>	Este atributo modela el tiempo de respuesta promedio para un tipo de consulta a una base de datos. Particularmente, la consulta es para N registros de una tabla.	(Ver Referencias para tipos de consultas)
<i>Fecha de Medición</i>	Se especifica la fecha, en la que se realizará la recolección de datos (si corresponde) y se computará la variable de calidad para los parámetros planificados.	F=[dd-mm-aaaa] o desde F ₁ a F ₂
<i>Parámetros Disponibles</i>	El estado actual para el tipo (T ₁) de consulta a BD	[segundos] 0,9
<i>Valores Planificados</i>	Peor Caso Mejor Caso Esperado	[segundos] 3 [segundos] 0,5 [segundos] 0,8
<i>Mayor grado de Satisfacción y otros valores de relevancia</i>	El mayor valor de X obtenido de computar la función elemental y que se traducirá en una preferencia de calidad del 100%	Si X ≤ 0,5 → IE _i = 100% Si X = 0,8 → IE _i = 80% Si X ≥ 3 → IE _i = 0%
<i>Función elemental</i>	X = Promedio (t₁ ...t_n)	Realizar veinticinco consultas para el tipo (T ₁) de consulta computando el promedio.
<i>Referencias</i>	Distintas fuentes de referencia	Ver documento X.x en donde se especifica tipos diferentes de consultas a bases de datos.

La figura 6.3 muestra un panorama del proceso de determinación de la preferencia de calidad elemental antes discutido. La entrada es el valor de la variable de calidad X_i, que conforme al criterio elemental correspondiente, permite obtener el valor del indicador o preferencia de calidad elemental IE_i.

Para determinar el valor de X_i se debe utilizar la función o métrica² adecuada conforme al criterio seleccionado. La elección del criterio de evaluación elemental nos permitirá computar valores de X, con mayor o menor precisión y objetividad (como veremos en la sección 6.3).

¹ Nota: Este atributo no fue empleado en los casos de estudio y es meramente ilustrativo (por lo tanto algunos valores son indicativos)

² Remitimos al lector a las secciones 10.1 y 10.2, en donde se discuten propiedades y un modelo conceptual para métricas.

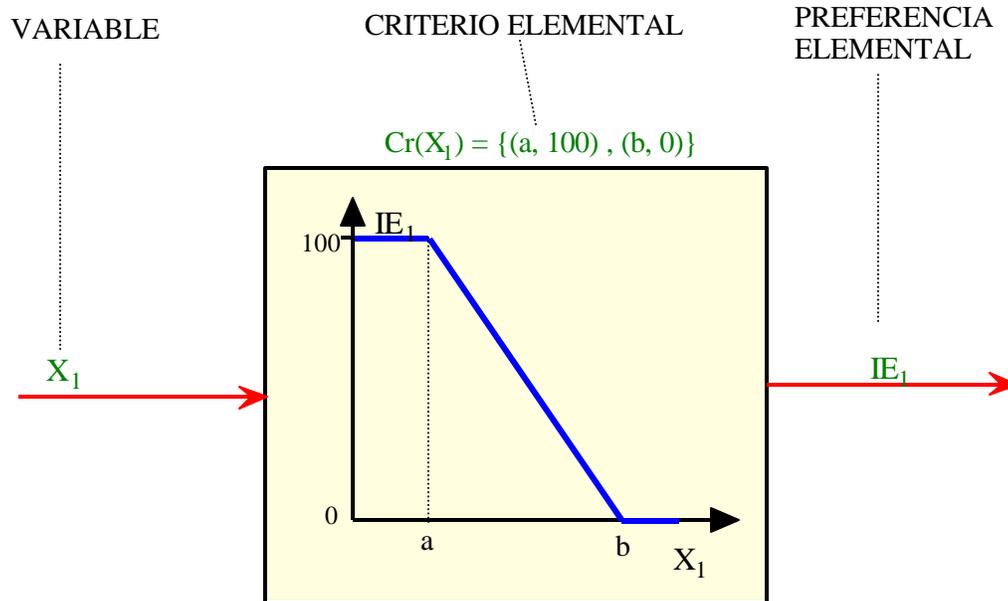


Figura 6.3 Panorama del proceso de determinación de la preferencia de calidad elemental.

6.2 Representación Notacional de los Criterios

Se pueden identificar al menos cuatro tipos diferentes de notaciones para representar a los criterios elementales [Dujmovic et al 82, Gilb 76], a saber:

- ✓ Notación gráfica (de líneas, barras, etc)
- ✓ Notación en escala de preferencia
- ✓ Notación de los puntos de coordenadas relevantes
- ✓ Notación analítica

Las figuras 6.1 y 6.2, representan una notación gráfica (de líneas) de los atributos respectivos. En el caso en que el criterio no sea continuo sino discreto podemos utilizar un gráfico de barras para mostrar el conjunto de valores y sus preferencias respectivas. Sin embargo, una notación reducida es la escala de preferencia. En ésta se muestran los puntos de mayor interés y su preferencia correspondiente. En la figura 6.4 se aprecia la escala de preferencia para el atributo *Capacidad de Almacenamiento Secundario*.

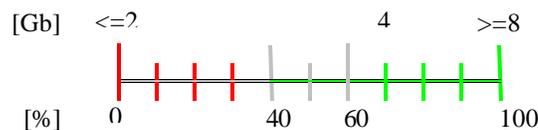


Figura 6.4 Escala de Preferencia para el atributo Capacidad de Almacenamiento Secundario

En muchos casos, es suficiente una representación notacional de las coordenadas de los puntos más relevantes. Por ejemplo, debido a la interpolación lineal entre dos puntos, una representación de las coordenadas de los puntos relevantes para el criterio de la figura 6.1 es la siguiente:

$$\text{CrE}(t_i) = \{ (t_{i \min}, 100), (t_{i \max}, 0) \} \text{ o, si tenemos en cuenta los valores de la tabla 6.1}$$

$$\text{CrE}(t_i) = \{ (0,5, 100), (0,8, 80), (3, 0) \}$$

Una interpretación en lenguaje natural de dichos valores y sus preferencia asociadas es: “un tiempo promedio de respuesta menor o igual a 0,5 segundos satisface totalmente el requerimiento; 0,8 seg. satisface el 80% del requerimiento; y, un tiempo mayor o igual a 3 segundos es totalmente inaceptable”

Además, se puede apreciar una notación analítica en (1), en la sección anterior.

Finalmente, no todos los tipos de notación son adecuados para todas las circunstancias. En los casos de estudio realizados, debido a la cantidad de atributos intervinientes, en la mayoría de los casos empleamos la notación de escala de preferencias para documentar a los atributos, debido a su poder visual y al espacio reducido que ocupa (en la sección 6.6 retomaremos el tema de documentación).

6.3 Tipos de Criterios de Preferencia de Calidad Elemental

Como indicábamos anteriormente, la elección del tipo de criterio de evaluación elemental resulta de importancia en consideración de los niveles de precisión, objetividad y facilidad de uso. El nivel de precisión depende del grado de criticidad de alguno o de todos los componentes del producto en un proyecto de evaluación.

Dos tipos básicos de criterios elementales son los *absolutos* y los *relativos*, y, dentro de los primeros se pueden descomponer en *criterios con variables continuas*, y *criterios con variables discretas*. La figura 6.5 muestra una jerarquía detallada de los tipos de criterios.

Un criterio de evaluación elemental absoluto es aquél que se emplea para determinar la preferencia absoluta de un atributo de un artefacto, y que no está relacionado con indicadores de otros sistemas comparativos. Un criterio absoluto se diferencia de uno relativo en que la meta de este último consiste solamente en la determinación de los indicadores relativos de los sistemas comparados sin evaluar la calidad (o la característica que corresponda) de cada sistema de un modo individual e independiente. Por lo tanto, una comparación relativa entre sistemas produce un orden o ranquin relativo, que no puede ser interpretado individualmente dado que no representan el grado de cumplimiento de los requerimientos absolutos.

En la siguientes subsecciones discutiremos los criterios elementales absolutos según la figura 6.5.

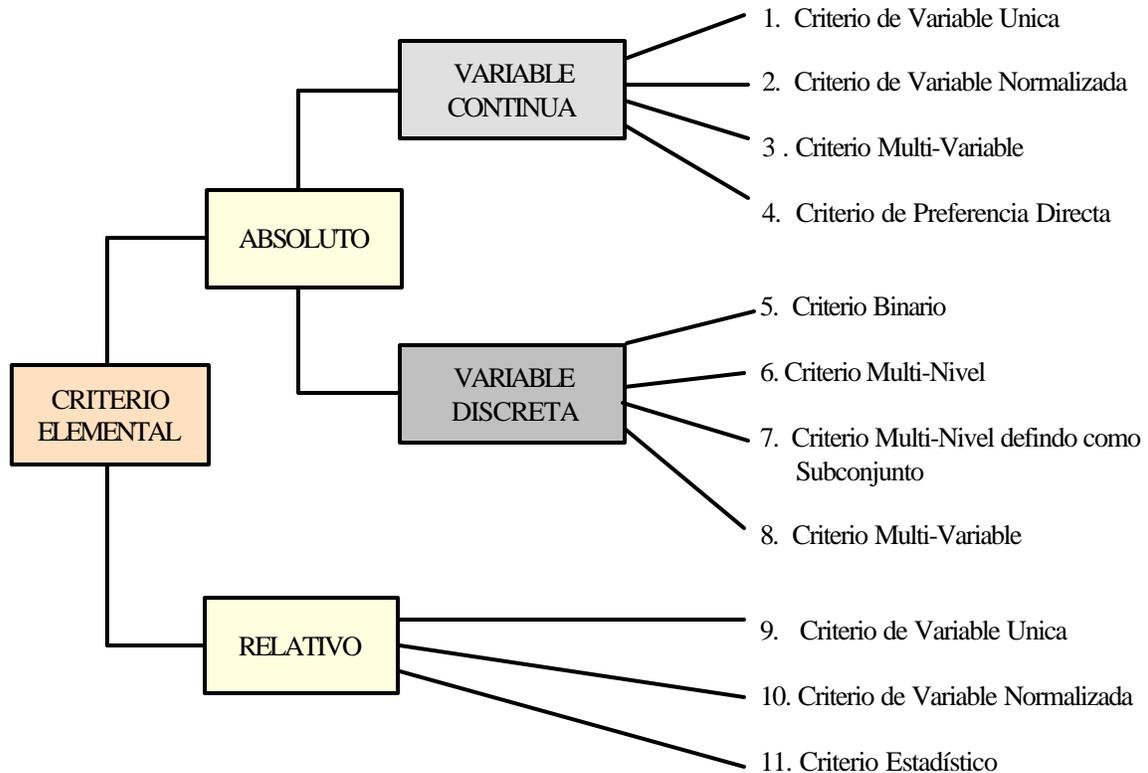


Figura 6.5 Taxonomía de tipos de criterio elementales

6.3.1 Criterios Elementales Absolutos con Variable Continua

6.3.1.1 Criterio de Variable Unica. Este es un criterio elemental común. Se asume que la variable X es única y continua, como por ejemplo, el tiempo medio entre dos fallas; el tiempo total transcurrido de un programa de prueba (benchmarking); el tiempo activo de un microprocesador durante una prueba, etc.

Con el fin de determinar el criterio elemental, el primer paso consiste en definir el rango de valores de interés para la evaluación de la variable continua. El siguiente paso, consiste en determinar las coordenadas de los puntos más relevantes y su preferencia de calidad. Si por ejemplo se requiere determinar el tiempo total transcurrido de un programa de prueba en una computadora, un criterio elemental podría ser:

$$CrE(t_i) = \{ (0,5 , 100) , (0,8 , 80) , (3 , 0) \} \text{ en unidades [seg. , \%]}$$

6.3.1.2 Criterio de Variable Normalizada. Este es un criterio elemental que se suele utilizar para evaluar la relación entre dos atributos con criterios absolutos de un mismo

sistema. Por ejemplo, si se requiere medir la utilización del microprocesador (U_m) dado un programa de prueba, este se puede medir por medio de los siguientes indicadores:

X_i = tiempo activo de un microprocesador durante una prueba

X_j = tiempo total transcurrido de un programa de prueba (benchmarking)

de donde, $U_m = X_i / X_j$

Por ejemplo, se empleó este tipo de criterio para determinar la preferencia de calidad del atributo *Imagen con Título* (cuyo código es 4.2.1.2.1, conforme a la figura 5.7). La fórmula para computar la variable es:

$$X = 100 - (AAR * 100/TAR) = 100 - X_1$$

en donde, **AAR** representa la cantidad de referencias ALT ausentes en los objetos gráficos; **TAR** la cantidad total de objetos que deben referenciar a la propiedad ALT; X_1 representa el % de ausencia de la propiedad ALT para el atributo del ente, y **X** representa el % de presencia de disponibilidad de la propiedad ALT

La siguiente figura muestra la escala de preferencia para dicho atributo, conforme al criterio antes descrito.

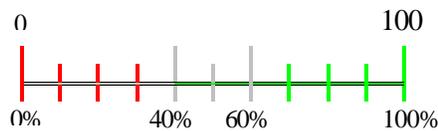


Figura 6.6 Escala de Preferencia para el atributo Imagen con Título

6.3.1.3 Criterio de Multi-variables Continuas. En este tipo de criterio, la variable X es resultante de algunas otras variables y constantes (el valor de X corresponde a una métrica indirecta).

Por ejemplo, se empleó este tipo de criterio para determinar la preferencia de calidad del atributo *Soporte a Lenguaje Extranjero* (cuyo código es 1.4.1, conforme a la figura 9.1). La fórmula para computar la variable se especificó en la tabla 5.5, y la preferencia de calidad se verá en la plantilla de la sección 6.4.

6.3.1.4 Criterio de Preferencia de Calidad Directa. Este tipo de criterio es subjetivo y basado en la experiencia y criterio de los evaluadores. Desde el punto de vista de la precisión y objetividad, es el peor criterio, debido a que se pueden introducir errores de valoración intencionales y/o involuntarios.

No obstante, dentro de los requerimientos algunos atributos sólo podrán comprobarse de un modo subjetivo, a partir del juicio de evaluadores expertos. Es decir, puede ser

difícil y costoso modelar la descomposición del “atributo” para determinar la preferencia de calidad. El criterio para la variable X se mapea en una preferencia trivial cuyas coordenadas son:

$$\text{CrE}(X_i) = \{ (0, 0), (100, 100) \}$$

Por ejemplo, para el atributo “*Uniformidad en el Estilo Global*” del sitio, puede ser conveniente medirlo mediante un criterio de preferencia directa y subjetiva. Realizarlo de otro modo, podría implicar mayores costos dada la relativa baja prioridad del atributo o la alta complejidad para descomponerlo. Sin embargo, dicho atributo es un elemento necesario en la composición para determinar la calidad global del artefacto.

De todos modos, la afirmación previa: “sólo podrán comprobarse de un modo subjetivo” debe interpretarse no de un modo categórico, sino relativo al contexto, relevancia, complejidad, costos y precisión requerida para el atributo. Generalmente, un atributo medido con un criterio de preferencia directa y subjetivo, se lo podría descomponer, en otros objetivos, más alguno subjetivo. En el caso del atributo mencionado, se podría observar si el sitio emplea fondos con imágenes y colores semejantes; si los colores del texto se mantienen, etc., y luego realizar una valoración subjetiva (y observacional) del mantenimiento del estilo en todos los subsitios del sitio.

Por último, la estrategia debe ser minimizar y tener bajo control el empleo de este criterio y, además, justificar su uso (ver sección 3.1.1).

6.3.2 Criterios Elementales Absolutos con Variable Discreta

6.3.2.1 Criterio Binario. Este criterio es el más simple de los criterios discretos y absolutos. El criterio para la variable binaria X se mapea en una preferencia elemental cuyas coordenadas son:

$$\text{CrE}(X_i) = \{ (0, 0), (1, 100) \}$$

En donde un valor de $X_i = 0$ se interpreta como la ausencia del atributo de calidad; en cambio un valor de $X_i = 1$, se interpreta como la presencia o disponibilidad del mismo. Este es el criterio utilizado para evaluar muchos de los atributos del árbol de requerimientos mostrados en las figuras 5.6 y 5.7. Por ejemplo, este fue el caso para determinar la preferencia de calidad de los atributos *Mapa del Sitio*, *Tabla de Contenidos*, e *Indice*, entre otros. La figura siguiente muestra la escala de preferencia respectiva.

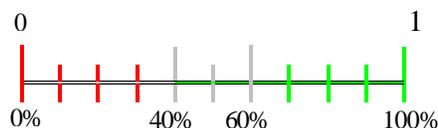


Figura 6.7 Escala de Preferencia para criterios binarios

6.3.2.2 Criterio de Multi-nivel. Este criterio es una generalización del criterio binario. La variable discreta puede tomar más de dos valores, cada uno de los cuales se corresponde a una preferencia de calidad. Por ejemplo, para el atributo *Soporte a Versión sólo Texto* (código 4.2.1.1), la variable discreta X se mapea en valores de preferencias cuyas coordenadas son :

$$\text{CrE}(X_i) = \{ (0, 0), (1, 60), (2, 100) \}$$

En donde un valor de $X_i = 0$ se interpreta como la ausencia del atributo de calidad; en cambio un valor de $X_i = 1$, se interpreta como la presencia parcial de la versión sólo texto; y, finalmente, un valor de $X_i = 2$, se interpreta como la presencia total de la versión sólo texto para todo el sitio Web.

Una codificación incremental de este tipo representa un incremento en la calidad elemental (por lo tanto, el tipo de escala es ordinal en donde se cumplen las propiedades de transitividad y completitud –sección 10.1.2).

6.3.2.3 Criterio de Multi-nivel definido como Subconjunto. Este criterio es un multi-nivel definido como un subconjunto de los números naturales (en una escala estrictamente ordinal). La variable discreta puede tomar más de dos valores, cada uno de los cuales se corresponde a una preferencia de calidad. Por ejemplo, para el atributo *Búsqueda de Personas*³ (código 2.1.1.1.1), la variable discreta X se mapea en valores de preferencias cuyas coordenadas son :

$$\text{CrE}(X_i) = \{ (0, 0), (1, 60), (2, 100) \}$$

En donde el listado de valores para X_i es como sigue:

0 = ausencia del mecanismo de búsqueda restringida;

1 = *búsqueda básica*: mecanismo de búsqueda por nombre/apellido;

2 = 1 + *búsqueda extendida o avanzada*⁴: mecanismo de búsqueda por unidad académica, y/o disciplina o materia, y/o TE, etc.

6.3.2.4 Criterio de Multi-variables discretas. Este criterio permite agrupar varias variables discretas y modelar el resultado en una única variable X. De este modo se puede reducir la cantidad de criterios elementales. Sea el conjunto de variables discretas D_1, \dots, D_h , entonces se puede definir una variable compuesta X, también discreta, como función de las anteriores, a saber:

³ En el capítulo 10 se realiza la validación de la métrica de este atributo

⁴ Para el atributo *Búsqueda Global* del sitio, se puede emplear igual criterio; no obstante, la búsqueda básica puede ser por medio de palabras claves, y la búsqueda avanzada, igual que la anterior pero incorporando operadores conjuntivos, disjuntivos, de negación, entre otras posibilidades como filtros.

$$X = F(D_1, \dots, D_n), \text{ y } X \in \{ X_1, \dots, X_n \}$$

La siguiente tabla muestra un ejemplo de correspondencia de preferencia elemental dadas tres variables discretas y binarias.

Tabla 6.2 *Ejemplo de un criterio absoluto y discreto con multi-variables binarias*

D₁	D₂	D₃	IE_i [%]
0	0	0	IE_{i1}
0	0	1	IE_{i2}
0	1	0	IE_{i3}
0	1	1	IE_{i4}
1	0	0	IE_{i5}
1	0	1	IE_{i6}
1	1	0	IE_{i7}
1	1	1	IE_{i8}

Por ejemplo, este criterio se podría emplear para modelar la preferencia de los atributos discretos y binarios *Mapa del Sitio*, *Tabla de Contenidos*, e *Indice*. Sin embargo, en los casos de estudio realizados utilizamos una función de agregación que modela la relación lógica de reemplazabilidad.

6.3.3 Otros Criterios Elementales

Finalmente, otros criterios podrían emplearse para modelar las preferencias de sistemas competitivos; en particular, los criterios relativos mostrados en la figura 6.5. Los mismos son de utilidad cuando se relacionan variables entre sistemas competitivos para determinar el indicador de calidad relativo, para cada sistema. No nos extenderemos en la explicación de estos criterios, debido principalmente a que no fueron utilizados en los casos de estudio que estamos desarrollando; sin embargo, el lector interesado puede encontrar información en la literatura citada (a saber, [Dujmovic et al 82, Fenton et al 97], entre otros).

6.4 Un Marco Descriptivo para Representar a las Características y Atributos de Calidad en consideración de sus Criterios de Preferencia de Calidad: Ejemplos.

A seguir, presentamos un marco de especificación descriptivo y jerárquico para representar a las características y atributos a partir de un árbol de requerimientos. Discutiremos un conjunto de características y atributos siguiendo una plantilla de especificación cuya estructura regularmente está compuesta por todos o algunos de estos elementos: título, código (del árbol de requerimientos), tipo de elemento (atributo, característica o subcaracterística), característica de más alto nivel, super y

subcaracterística (en caso que corresponda), definición y comentarios, tipo de criterio de preferencia elemental, plantilla de referencia a variables y parámetros, escala de preferencia, preferencias de entrada, pesos, operador aritmético/lógico, tipo de recolección de datos, herramienta utilizada, ejemplo/s de uso y valor/es computado/s [Olsina et al 99b].

En la figura 6.8 se muestran tres tipos de plantillas: para el componente característica, y para el componente subcaracterística, y para el atributo (elemento de más bajo nivel en la jerarquía).

<p>Título: <u>Código:</u></p> <p>Tipo: Característica</p> <p><u>Sub-característica/s (Código/s):</u></p> <p><u>Definición / Comentarios:</u></p> <p><u>Modelo para determinar el Cómputo Global :</u></p> <p><u>Herramienta Empleada:</u></p> <p><u>Peso:</u></p> <p><u>Operador Aritmético / Lógico:</u></p> <p><u>Ejemplo/s:</u></p> <p><u>Valor/es de Preferencia/s Computado/s:</u></p> <p>a)</p>	<p>Título: <u>Código:</u></p> <p>Tipo: Atributo</p> <p><u>Característica de más Alto Nivel (Código):</u></p> <p><u>Super-característica (Código):</u></p> <p><u>Definición / Comentarios:</u></p> <p><u>Tipo de Criterio Elemental:</u></p> <p><u>Plantilla de Referencia de Variables y Parámetros:</u></p> <p><u>Escala de Preferencia:</u></p> <p><u>Tipo de Recolección de Datos:</u></p> <p><u>Herramienta Empleada:</u></p> <p><u>Ejemplo/s:</u> <u>Peso:</u></p> <p><u>Valor/es de Preferencia/s Computado/s:</u></p> <p>b)</p>
<p>Título: <u>Código:</u></p> <p><u>Super-característica (Código):</u></p> <p><u>Sub-característica/s (Código/s):</u></p> <p><u>Definición / Comentarios:</u></p> <p><u>Modelo para determinar el Cómputo Parcial:</u></p> <p><u>Peso:</u></p> <p><u>Ejemplo/s:</u></p> <p>c)</p>	<p>Tipo: Subcaracterística</p> <p><u>Atributo/s (Código/s):</u></p> <p><u>Herramienta Empleada:</u></p> <p><u>Operador Aritmético/Lógico:</u></p> <p><u>Valor/es de Preferencia/s Computado/s:</u></p>

Figura 6.8 a) Plantilla para especificar a una característica de más alto nivel; **b)** Plantilla para especificar a un atributo; **c)** Plantilla para especificar a una subcaracterística

Para ejemplificar lo anterior, tomaremos características y atributos del árbol de requerimientos de calidad para el caso de estudio de sitios académicos.

6.4.1 Especificación de Algunas Características de Calidad

En la presente subsección, se especifican dos características de alto nivel, *Usabilidad* y *Eficiencia*, a partir del árbol de requerimientos de calidad de las figuras 5.6 y 5.7.

Título: *Usabilidad*; **Código:** 1 ; **Tipo:** *Característica*

Sub-característica/s: *Comprensibilidad Global del Sitio (1.1); Mecanismos de Ayuda y Retroalimentación en línea (1.2); Aspectos de Interface y Estéticos (1.3); Misceláneas (1.4)*

Definición / Comentarios: Es una característica de calidad de producto de alto nivel, -que se la puede medir mediante cálculo a partir de métricas directas e indirectas-, y representa la capacidad o potencialidad del producto para ser utilizado, comprendido y operado por los usuarios, además de ser atractivo. Incluye principalmente a subcaracterísticas como comprensibilidad, operabilidad, facilidad de aprendizaje y comunicatividad, entre otras subcaracterísticas como estéticas y de estilo que hacen del artefacto que sea agradable de usar.

Es importante remitirnos, como marco conceptual, a la definición dada por el estándar ISO [ISO/IEC 9126-1], que dice en pp. 6: “*The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions*”. Y la definición dada por [IEEE Std 1061], en su Anexo informativo A, en pp. 19, que dice: “*An attribute that bears on the effort needed for use (including preparation for used and evaluation of results), and on the individual assessment of such use by users*”.

Modelo para determinar el Cómputo Global: modelo LSP (en nuestro caso), o modelo meramente Aditivo;

Herramienta Empleada: Automática (WebQEM_Tool), desarrollada para computar operadores lógicos de preferencia continua y su agregación.

Peso: con un peso de 0.3 –ver el esquema de la figura 7.8

Operador Aritmético/Lógico: C --

Ejemplo/s: Se utilizó como parte constituyente del modelo de calidad en dos casos de estudio [Olsina 99, Olsina et al 99c] y un survey [Lafuente et al 99]

Valor/es de Preferencia/s Computado/s: hace referencia a una entrada (código 1 de la tabla 7.1), con el valor o valores obtenidos (en caso que el sitio a evaluar sea uno o más).

Título: *Eficiencia*; **Código:** 4 ; **Tipo:** *Característica*

Sub-característica/s: *Performance (4.1); Accesibilidad (4.2)*

Definición / Comentarios: Es una característica de calidad de producto de alto nivel, -que se la puede medir mediante cálculo a partir de métricas directas e indirectas-, y principalmente representa a la relación entre el grado de performance del artefacto y la cantidad de recursos (tiempo, espacio, etc.) usados bajo ciertas condiciones.

Es importante remitirnos, como marco conceptual, a la definición dada por la ISO [ISO/IEC 9126-1], que dice en pp. 7: “*The capability of the software product to provide appropriate performance, relative to the amount of resources used, under stated conditions*”. Y la definición dada por [IEEE Std 1061], en su Anexo informativo A, en pp. 19, que dice: “*An attributes that bears on the relationship of the level of performance to the amount of resources used under stated conditions*”.

Modelo para determinar el Cómputo Global/Parcial: modelo LSP, o modelo meramente Aditivo;

Herramienta Empleada: Automática (WebQEM_Tool), desarrollada para computar operadores lógicos de preferencia continua y su agregación.

Peso: con un peso de 0.3 –ver el esquema de la figura 7.8

Operador Aritmético/Lógico: C --

Ejemplo/s: Se utilizó como parte constituyente del modelo de calidad en dos casos de estudio [Olsina 99, Olsina et al 99c].

Valor/es de Preferencia/s Computado/s: hace referencia a una entrada (código 4 de la tabla 7.1), con el valor o valores obtenidos (en caso que el sitio a evaluar sea uno o más).

6.4.2 Especificación de Algunas Subcaracterísticas de Calidad

Se especifican dos subcaracterísticas de calidad, *Comprensibilidad Global del Sitio* y *Aspectos de Búsqueda y Recuperación*, a partir del árbol de requerimientos de calidad de la figura 5.7.

Título: *Comprensibilidad Global del Sitio*; Código: 1.1 ; Tipo: Subcaracterística

Super-Característica: Usabilidad (1)

Sub-característica/s: Esquema de Organización Global (1.1.1)

Atributo/s: Calidad del Sistema de Etiquetado (1.1.2), Visita Guiada Orientada al Estudiante (1.1.3); Mapa de Imagen del Campus (1.1.4)

Definición / Comentarios: Es una característica que representa a todas aquellas facilidades que permiten a la audiencia, tener una rápida comprensión tanto de la estructura organizativa, como del contenido del sitio Web como un todo, facilitando el rápido acceso y recorrido del mismo y sus componentes. Por tal razón, los atributos y subcaracterísticas se hallan principalmente en la página principal o en los primeros niveles del sitio.

Modelo para determinar el Cómputo Parcial: modelo LSP, o meramente Aditivo;

Herramienta Empleada: Automática (WebQEM_Tool), desarrollada para computar operadores lógicos de preferencia continua y su agregación.

Peso: con un peso de 0.35 –ver el esquema de la figura 7.7 a)

Operador Aritmético/Lógico: A

Valor/es de Preferencia/s Computado/s: ver tabla 7.1

Ejemplo/s: Se utilizó como parte constituyente del modelo de calidad en dos casos de estudio [Olsina 99, Olsina et al 99c] y un survey [Lafuente et al 99]

Título: *Aspectos de Búsqueda y Recuperación*; Código: 2.1 ; Tipo: Subcaracterística

Super-Característica: Funcionalidad (2)

Sub-característica/s: Mecanismo de Recuperación (2.1.2); Mecanismo de Búsqueda (2.1.1)

Atributo/s: No posee

Definición / Comentarios: Es una característica que modela el mecanismo que permite a la audiencia, tener un modo directo de encontrar información u ocurrencias de documentos a partir de palabras o frases claves, y operadores o filtros. Además, permite recuperarla conforme a ciertas preferencias de usuario y obtener asimismo retroalimentación de lo sucedido.

Modelo para determinar el Cómputo Parcial: modelo LSP, o meramente Aditivo;

Herramienta Empleada: Automática, desarrollada para computar operadores lógicos de preferencia continua y su agregación.

Peso: con un peso de 0.3 –ver el esquema de la figura 7.7 b)

Operador Aritmético/Lógico: C --

Ejemplo/s: idem al ejemplo anterior.

Valor/es de Preferencia/s Computado/s: ver tabla 7.1

6.4.3 Especificación de Atributos

En la presente subsección, se especifican cerca de una docena de atributos intervinientes en el modelo de calidad, a partir del árbol de requerimientos de las figuras 5.6 y 5.7. Principalmente, se describen atributos con distintos tipos de criterios de preferencia elemental o mecanismos de recolección de datos. (En el apéndice A.1.1, el lector puede analizar algunos otros atributos especificados para sitios académicos, y en la sección 9.2.2, para los sitios de museos)

Título: *Mapa del Sitio*; Código: 1.1.1.1; Tipo: *Atributo*

Característica de más Alto Nivel: *Usabilidad*

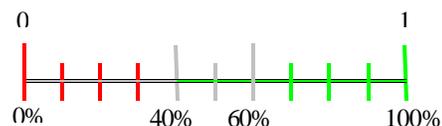
Super-característica: *Esquema de Organización Global*

Definición / Comentarios: Un mapa del sitio es una representación con componentes gráficos, que muestra la estructura o arquitectura global (a menudo jerárquica) del sitio Web como un todo.

Con frecuencia, el término se usa intercambiamente por los términos tabla de contenido o índice. Como indican los autores en [Rosenfeld et al 98], la definición anterior excluye a una tabla de contenido y a un índice por el uso de componentes gráficos para resaltar aspectos estructurales e incrementar la atracción estética. De esta manera, un mapa del sitio, presenta la información de un modo que va más allá de la representación textual, no obstante, permitiendo navegación directa a partir de sus elementos enlazados, del mismo modo que la tabla de contenidos e índices.

Tipo de Criterio Elemental⁵: es un criterio binario, discreto y absoluto: sólo se pregunta si está disponible (1) o si no está disponible (0).

Escala de Preferencia:



⁵ En algunas especificaciones de atributos de este trabajo se hace uso del campo *Plantilla de Referencia de Variables y Parámetros* (ver figura 6.8b y sección 5.7), en otras, simplemente se adiciona la información en el campo de criterio. Sin embargo, la idea sería que cada atributo tenga su plantilla asociada o enlazada, principalmente, en caso de transformar esta información en un documento hiperenlazado.

Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s: Un ejemplo, como se ilustra en la parte izquierda de la figura 6.9, es el mapa del sitio de la UPC (<http://www.upc.es/catala/index/index.htm>) produciendo una preferencia elemental de calidad del 100%.

Peso: con un peso de 0.2 –ver el esquema de la figura 7.7 a)

Valor/es Computado/s: hace referencia a una entrada (como la de la tabla 6.3), con el valor o valores obtenidos (en caso que el sistema a evaluar sea uno o más).

Título: *Mapa de Imagen del Campus*; **Código:** 1.1.4; **Tipo:** Atributo

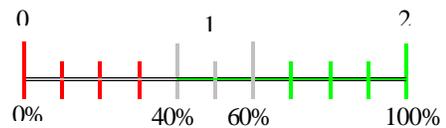
Característica de más Alto Nivel: Usabilidad

Super-característica: Comprensibilidad Global del Sitio

Definición / Comentarios: Un mapa de imagen es una imagen con áreas cliqueables las que proveen puntos de partida (o anchors) directas a contenidos (tal vez navegables), y/o funciones de ampliación (zooming); o que permiten visitas guiadas con capacidad de realidad virtual (VR tours). No obstante, por ejemplo para mapas de imagen de pisos y salas de museos o campus de universidades, se considera el caso en que se permita enlaces de navegación directos o indirectos; esto es, el mapa de imagen tiene referencias que pueden ser navegadas indirectamente, por medio de hipertexto u otros objetos gráficos cliqueables.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde si se evalúa la disponibilidad del mapa de imagen del campus y/o edificio, entonces: 0 = no disponible; 1 = disponible, permitiendo sólo referencias no enlazadas y/o zooming; 2 = disponible, permitiendo navegación enlazada directa o indirectamente y/o zooming, y/o visita virtual.

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

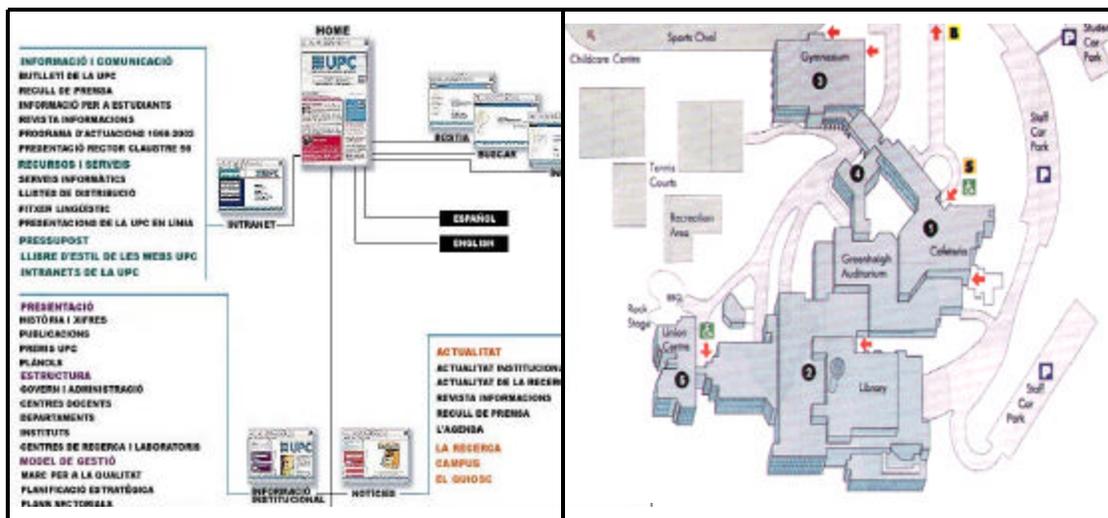


Figura 6.9 Vista parcial del mapa del sitio de la Universidad Politécnica de Cataluña, en la parte izquierda de la figura, y a la derecha, una sección del mapa de imagen de la Universidad Tecnológica de Sidney.

Ejemplo/s:

1) Un ejemplo es el mapa de imagen del campus de la UTS como se ilustra parcialmente en el costado derecho de la figura 6.9. La preferencia elemental de calidad computada es 100%. (Ver el URL: <http://www.uts.edu.au/div/publications/maps/kg.html>).

2) otro ejemplo es el mapa del campus de NUS (<http://www.nus.edu.sg/vr/>) que permite visitar y navegar diferentes edificios y lugares por medio de una visita virtual. La preferencia elemental computada es 100%.

Título: *Soporte a Lenguaje Extranjero*; **Código:** 1.4.1; **Tipo:** Atributo

Característica de más Alto Nivel: Usabilidad

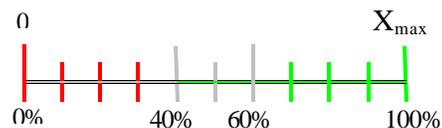
Super-característica: Misceláneas

Definición / Comentarios: Este atributo modela el número de lenguajes extranjeros soportados por un sitio (sitios de dominios de aplicación como académicos, museos, comercio electrónico, etc.). Además, especifica el nivel de soporte para cada lenguaje, a saber: total (todas las páginas del sitio), parcial (algunos subsitios del sitio), o mínimo (algunas páginas o documentos de algunos subsitios). No se computa obviamente el lenguaje nativo del sitio Web, como lenguaje extranjero.

Tipo de Criterio Elemental: es un criterio multi-variable, continuo y absoluto;

Plantilla de Referencia de Variables y Parámetros: Ver Tabla 5.5

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s:

1) Un ejemplo es el soporte a lenguajes extranjeros del sitio de la UPC (editado en catalán, como lenguaje nativo), el que tiene soporte parcial en español e inglés, produciendo una preferencia elemental del 90%.

2) En el caso de estudio de museos, el Louvre (París) obtuvo la mayor preferencia elemental, debido a que poseía soporte parcial a tres lenguajes extranjeros (español, japonés e inglés) [Olsina 99], en tanto que el museo del Prado obtuvo una marca menor, debido a que soportaba totalmente sólo a inglés. Finalmente, Galería del Arte soportaba mínimamente 4 lenguajes, en tanto que el museo Metropolitano (Nueva York) no tenía soporte a idiomas extranjeros.

Título: *Búsqueda de Personas*; **Código:** 2.1.1.1.1; **Tipo:** Atributo

Característica de más Alto Nivel: Funcionalidad

Super-característica: Búsqueda Restringida (scoped search)

Definición / Comentarios: Algunas veces, tiene sentido brindar al usuario con la facilidad de búsqueda restringida a un subsitio o parte de un sitio, debido a que el mismo es altamente cohesivo o distintivo del resto de la información del sitio Web global. Por ejemplo, para un visitante del museo en la Web puede con frecuencia necesitar ambos tipos de búsquedas: global y restringida. Es decir, podría necesitar

buscar información de colecciones por autor y escuela (scoped search), en tanto que una búsqueda global dentro del sitio también podría ser necesaria con el fin de investigar el espacio de información global.

Jakob Nielsen, en su columna editada electrónicamente (denominada Alertbox, en el artículo *Search and You May Find* <http://www.useit.com/alertbox/9707b.html>), dice: “Search is one of the most important user interface elements in any large website. As a rule of thumb, sites with more than about 200 pages should offer search. Guidelines for search include:

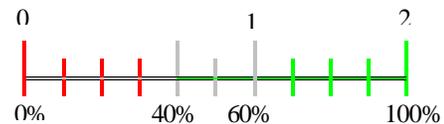
- ✓ a search button on every page
- ✓ global search (searching all of the site) is better than scoped search
- ✓ boolean queries should be relegated to a secondary "advanced search" page

Our usability studies show that more than half of all users are search-dominant, about a fifth of the users are link-dominant, and the rest exhibit mixed behavior. The search-dominant users will usually go straight for the search button when they enter a website: they are not interested in looking around the site; they are task-focused and want to find specific information as fast as possible. In contrast, the link-dominant users prefer to follow the links around a site: even when they want to find specific information, they will initially try to get to it by following promising links from the home page. Only when they get hopelessly lost will link-dominant users admit defeat and use a search command. Mixed-behavior users switch between search and link-following, depending on what seems most promising to them at any given time but do not have an inherent preference.

Despite the primacy of search, web design still needs to be grounded in a strong sense of structure and navigation support: all pages must make it clear where they fit in the larger scheme of the site. First, there is obviously a need to support those users who don't like search or who belong to the mixed-behavior group. Second, users who get to a page through search still need structure to understand the nature of the page relative to the rest of the site. They also need navigation to move around the site in the neighborhood of the page they found by searching: it is rare that a single page holds all the answers or even that the search found the most relevant page, so users need to see related pages “.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto, definido como subconjunto. Podemos decir que: 0 = no disponible mecanismos de búsqueda restringida; 1 = búsqueda básica: mecanismo de búsqueda por nombre/apellido; 2 = 1 + búsqueda extendida o avanzada: mecanismo de búsqueda (con filtros) por unidad académica, y/o disciplina o materia, y/o TE, etc.

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s:

1) Un ejemplo digno a ser destacado es la búsqueda de personas del sitio de la universidad de Stanford (<http://sin.stanford.edu:2000/frame?person>) como se ilustra en la figura 6.10. La preferencia elemental computada es 100%.

2) Otros ejemplos son las Universidades de Chile (<http://www.sisib.uchile.cl/docentes/>) y la UQAM (<http://www.repertoire.uqam.ca/>) las que obtuvieron una preferencia elemental de calidad de 60 y 100%, respectivamente.

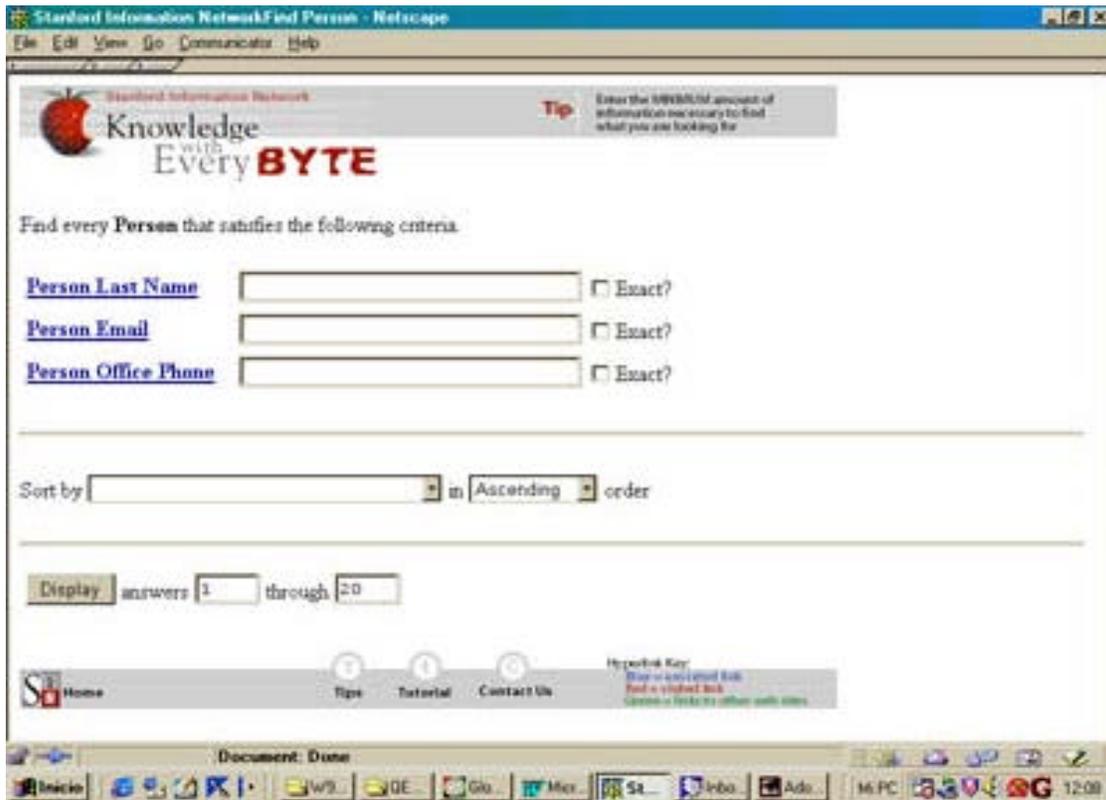


Figura 6.10 Pantalla bajada del sitio de la Universidad de Stanford en donde se puede apreciar las facilidades de Búsqueda de Personas (people scoped search), y de personalización de la recuperación de ocurrencias.

Título: *Enlaces Rotos*; **Código:** 3.1.1; **Tipo:** Atributo

Característica de más Alto Nivel: *Confiabilidad*

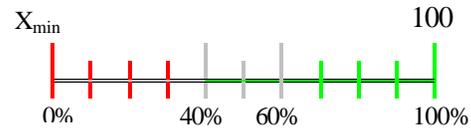
Super-característica: *Errores de Enlaces*

Definición / Comentarios: Este atributo representa básicamente a los enlaces encontrados que conducen a nodos destino ausentes (también llamados enlaces ausentes o pendientes)

Jakob Nielsen, en su columna editada electrónicamente (denominada Alertbox: <http://www.useit.com/alertbox/980614.html>), dice: “6% of the links on the Web are broken according to a recent survey by Terry Sullivan's All Things Web. Even worse, linkrot in May 1998 was double that found by a similar survey in August 1997. Linkrot definitely reduces the usability of the Web, being cited as one of the biggest problems in using the Web by 60% of the users in the October 1997 Gvu survey. This percentage was up from "only" 50% in the April 1997 survey. Users get irritated when they attempt to go somewhere, only to get their reward snatched away at the last moment by a 404 or other incomprehensible error message”.

Tipo de Criterio Elemental: es un criterio de variable normalizada, continuo y absoluto; en donde si BL=Número de enlaces rotos encontrados. TL = Número total de enlaces del sitio. La fórmula para computar la variable es: $X = 100 - (BL * 100/TL) * 10$; donde, si $X < 0$ entonces $X = 0$.

Escala de Preferencia:



Tipo de Recolección de Datos: Automatizado

Ejemplo/s: Por ejemplo, el sitio de la Universidad Nacional de Singapur produce una preferencia del 68.06 %. El valor numérico real, computado a partir de la anterior fórmula, es como sigue: $100 - ((970 * 100) / 30883) * 10 = 68.06\%$

Título: Páginas de Acceso Rápido; Código: 4.1.1; Tipo: Atributo

Característica de más Alto Nivel: Eficiencia

Super-característica: Performancia

Definición / Comentarios: Para este atributo, se mide el tamaño de todas las páginas (estáticas) del sitio Web considerando todos sus componentes gráficos, tabulares y textuales. El tamaño de cada página se especifica como una función del tiempo de espera y de la velocidad mínima establecida para una línea de comunicación dada.

Se especifica un tamaño umbral aceptable, para el tamaño total de cada página, por ejemplo, el de 35,2 Kb. Una página de este tamaño requiere 20 segundos para ser bajada a una taza de 14,400 bps. Ese es el tiempo aceptable que un usuario debe esperar, sin que se ponga impaciente. Estudios indican que:

“Users tend to become annoyed when a page takes longer than 20 seconds to load. This means it is best to limit the total of the file sizes associated with a page, including graphics, to a maximum of 30 – 45 kilobytes to assure reasonable performance for most users.”, extraído el 27-Ene-1999 [IEEE WPG], sección *PERFORMANCE-Designing for performance-file sizes* http://www.ieee.org/web/developers/style/perfrm_1.htm)

Un estudio previo, indicaba:

“In mid-1997, a study found that the mean size of Web pages was 44 kilobytes. This is more than five times too big for optimal response time for ISDN users, so even when more people get mid-band connections, the Web will be much too slow. Also note that 44KB is 30% larger than even the most generous size limit for modem users”. [Nielsen_Alert] -Size Limits for Web Pages- <http://www.useit.com/alertbox/sizelimits.html>

Tipo de Criterio Elemental: es un criterio multi-variable, continuo y absoluto.

Plantilla de Referencia de Variables y Parámetros: Ver Tabla 5.4

Escala de Preferencia:

Tipo de Recolección de Datos: Automatizado

Ejemplo/s:

1) Considerando el sitio de UTS, la herramienta utilizada para recolectar y documentar esta medición elemental (SiteSweeper), informó: “*You specified a total download size limit of 35.2K bytes per page. A page this size requires about 20 seconds to download at 14.4 Kbps. Of the 18.872 pages on your site, 2.210 pages (12 %) have a total download size that exceeds this threshold*”. Teniendo en cuenta la anterior fórmula y los datos obtenidos de la herramienta (ver figura 6.11), el siguiente cómputo: $(16662 - 0.4 * 1850 - 0.8 * 440) / 18872$, produjo una preferencia elemental de 82 %.

2) Un ejemplo de performance extraordinario, es el del sitio Web de Stanford que arrojó una preferencia elemental de 100% (ninguna página superó el umbral de 35,2 Kb)

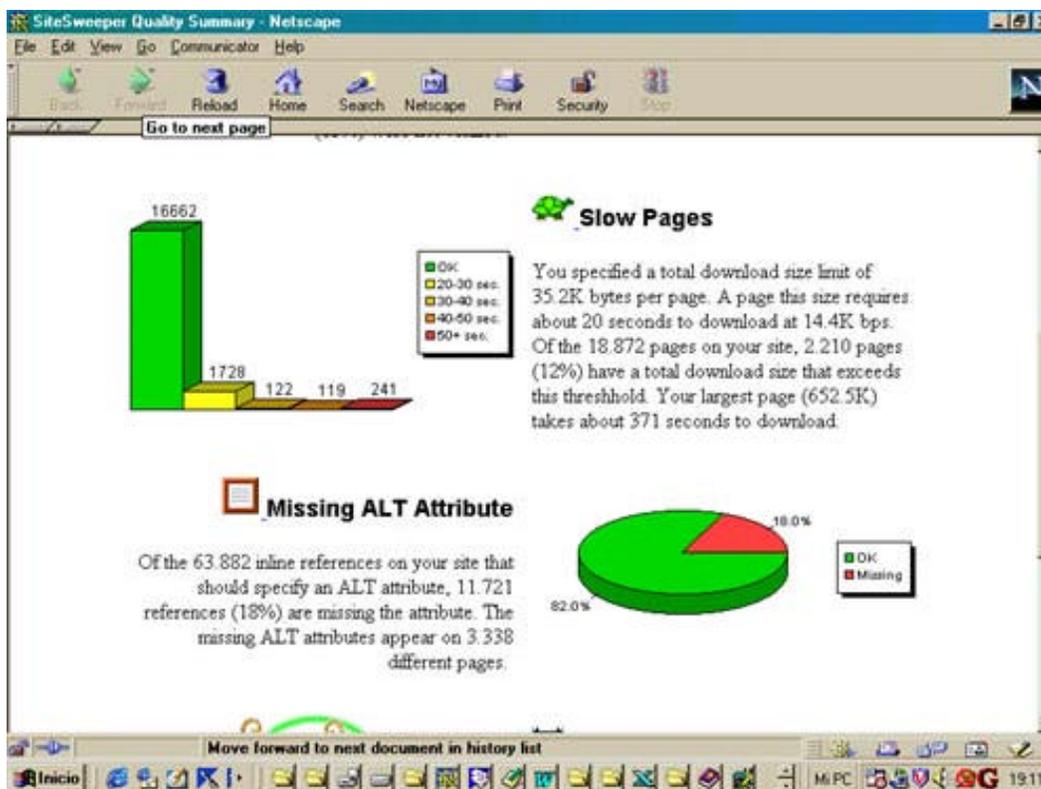
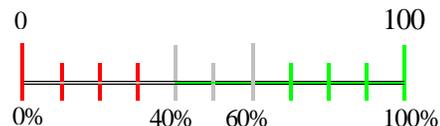


Figura 6.11 Pantalla tomada del informe “*Quality Page*” (usando la versión SiteSweeper 2.0), que muestra las diferentes categorías de tiempo de acceso a páginas, y los porcentajes de presencia del parámetro ALT, en los documentos fuentes HTML, del sitio de la UTS.

Título: *Imagen con Título*; **Código:** 4.2.1.2.1; **Tipo:** *Atributo*

Característica de más Alto Nivel: *Eficiencia*

Super-característica: *Legibilidad al desactivar la Propiedad Imagen del Browser*

Definición / Comentarios: Se debe proveer texto alternativo para cada imagen u objeto gráfico dado que comunica información visual. Este atributo mide el porcentaje de

presencia de la etiqueta <ALT> de manera de incluir texto alternativo a la imagen, cuando se desactiva la propiedad de *ver imágenes* del navegador. La presencia de este atributo favorece la legibilidad de la página. No obstante, esta presencia no garantiza la calidad del texto alternativo. El texto pudo haber sido generado automáticamente al editar la página con herramientas como FrontPage, etc.

Como comentario, es importante considerar las guías provistas por la organización W3C, en el documento en línea denominado “WAI Accessibility Guidelines” [W3C 99]. En lo que respecta a este atributo, dice: “A.1 Provide alternative text for all images, applets, and image maps”. Entre otras cosas dice: “Text is considered accessible to almost all users since it may be handled by screen readers, non-visual browsers, braille readers, etc. It is good practice, as you design a document containing non-textual information (images, graphics, applets, sounds, etc.) to think about supplementing that information with textual equivalents wherever possible”.

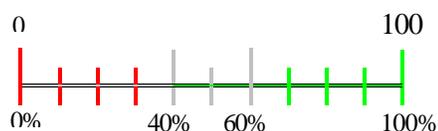
Tipo de Criterio Elemental: es un criterio de variable normalizada, continuo y absoluto; en donde si AAR= Referencia ALT ausente. TAR=Número total de objetos que deben referenciar a la propiedad ALT. La fórmula para computar la variable es: $X = 100 - (AAR * 100/TAR)$

Escala de Preferencia:

Tipo de Recolección de Datos: Automatizado

Ejemplo/s:

1) Un ejemplo se muestra en la figura 6.11 para el sitio de UTS. La herramienta reportó: “Of the 63.882 inline references on your site that should specify an ALT attribute, 11.721 references (18%) are missing the attribute. The missing ALT attributes appear on 3.338 different pages”. La preferencia computada, a partir de la fórmula anterior es del 81.65%



Título: **Legibilidad Global**; Código: 4.2.1.2.2; Tipo: **Atributo**

Característica de más Alto Nivel: **Eficiencia**

Super-característica: **Legibilidad al desactivar la Propiedad Imagen del Browser**

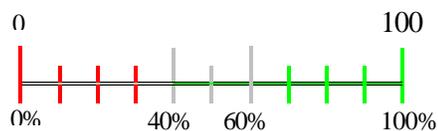
Definición / Comentarios: Este atributo representa la preferencia de calidad en consideración del nivel de legibilidad global del sitio, cuando se desactiva la propiedad de *ver imágenes* del navegador.

Como fue comentado en el atributo *Imagen con Título*, la presencia de objetos con texto no garantiza la calidad del texto alternativo.

Tipo de Criterio Elemental: es un criterio de preferencia directa

Escala de Preferencia:

Tipo de Recolección de Datos: Manual, Observacional



Ejemplo/s: Un ejemplo, es el sitio de la UPC que alcanzó una preferencia global directa mayor respecto de los restantes sitios y que es de 80%.

6.5 Implementación de las Mediciones Elementales.

Una vez que los criterios de calidad para los atributos han sido definidos y acordados, es preciso recolectar los datos y computar las variables y preferencias elementales. Un aspecto importante de este proceso es la calidad de los datos recolectados. Para ello es necesario tener en cuenta, en principio, las siguientes propiedades (ver [Fenton et al 97], capítulo 5), entre otras:

- ✓ *La medida debe ser correcta*: los datos deben ser recolectados conforme al criterio establecido y a las reglas específicas para el atributo en cuestión. Por ejemplo, para el atributo *Enlaces Rotos*, si se establece que se considere sólo aquellos nodos destinos que están ausentes, no debiera computarse aquellos que no están disponibles por una inaccesibilidad temporaria del servidor donde reside la página destino del enlace, u otra causa (como acceso no permitido por seguridad).
- ✓ *La medida debe ser replicable y consistente*: la medida es repetible e insensible a pequeños cambios en el entorno, herramientas, y observadores (ver además [ISO/IEC 14598-5], sección 5.3).
- ✓ *La medida debe estar asociada a un período de tiempo*: la medida puede variar, por ejemplo, dado el carácter evolutivo de un artefacto (como la Web); por lo tanto se debe controlar y especificar (estampar) el tiempo de medición.
- ✓ *La medida debe ser precisa*: se deben establecer escalas, tipos de escala, rangos y niveles de precisión conforme al criterio establecido.

Por otra parte, debemos considerar el tipo de recolección de datos. Esta actividad se puede realizar de un modo manual (observacional), semiautomático, o automático.

6.5.1 Herramientas de Soporte a la Automatización en el Proceso de Recolección de Datos.

Muchos de los valores para las variables de los atributos especificados, fueron obtenidos manual u observacionalmente, dado que no había hasta ese momento otro modo de hacerlo efectivamente. Sin embargo, la recolección de datos automática es en muchos casos el único mecanismo para obtener resultados de un modo confiable y efectivo. Este fue el caso para medir atributos como *Enlaces Rotos*, *Páginas de Acceso Rápido*, entre otros, como se mostró en la sección 6.4.3 (y en el Apéndice A).

De tres herramientas analizadas dentro de algunas de las existentes a ese momento (para W98 y WNT), SiteSweeper 2.0 cubrió nuestras expectativas. Este tipo de herramienta no sólo recolecta datos en línea sino que también genera informes textuales, gráficos e

hipertextuales, que facilitan el ulterior análisis y documentación. Por ejemplo, la pantalla mostrada en la figura 6.11 es una de las páginas del informe generado para “Quality Page” en donde se aprecian las diferentes categorías de tiempo de acceso a páginas conforme a parámetros previamente establecidos. Además, se muestran los porcentajes de presencia del parámetro ALT, en los documentos fuentes HTML. Del mismo modo, recolecta datos para cantidad de enlaces, enlaces rotos, enlaces promedio por página, entre otros. La siguiente figura, muestra una pantalla en donde informa sobre enlaces rotos, para el caso del sitio académico de la Universidad de Stanford. (Además, claramente se puede observar que ninguna página superó el umbral de 35,2 K de tamaño).

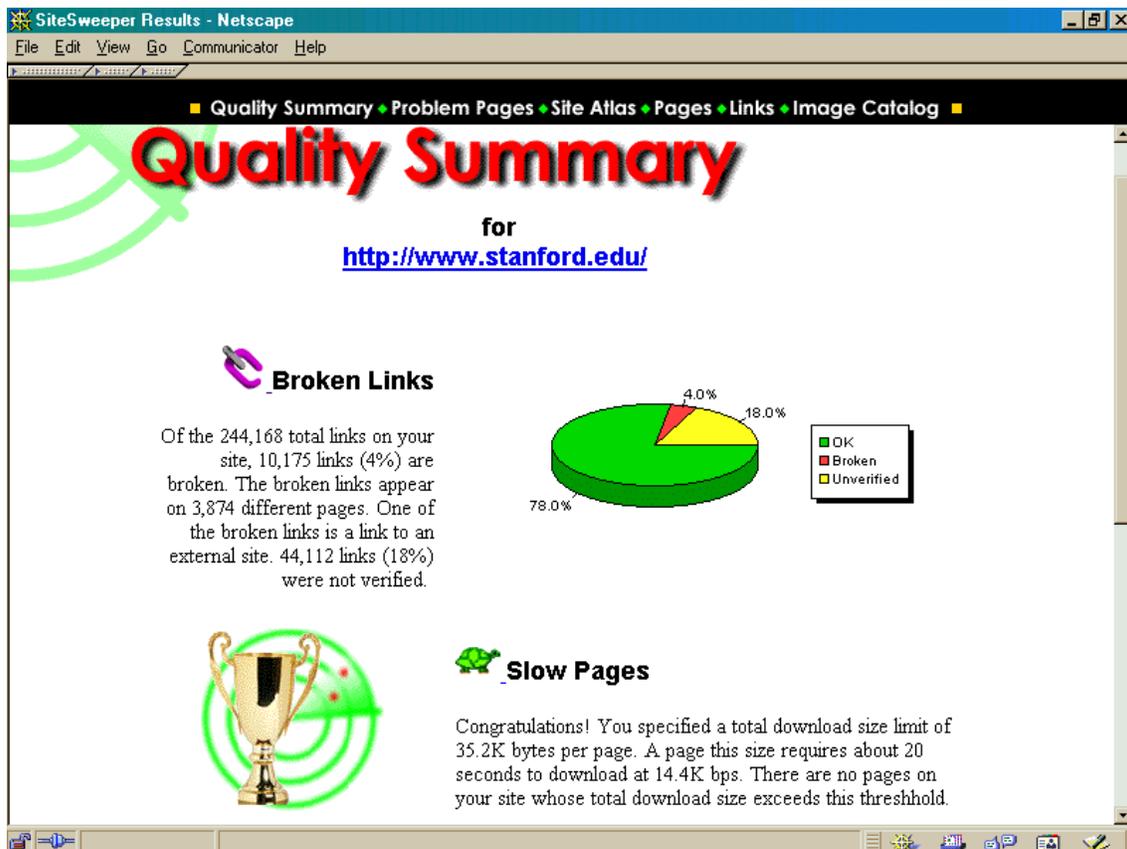


Figura 6.12 Pantalla tomada del informe “Quality Link” (usando la versión SiteSweeper™ 2.0) del sitio de la Universidad de Stanford.

La herramienta informa que de 244.168 enlaces del sitio 3.874 están ausentes. La porción de los no verificados consiste de aquellos enlaces a los que no se tiene acceso (acceso o permiso denegado, etc.), pero no son en principio nodos ausentes.

Desde inicios de 1999, estamos desarrollando una herramienta que permita automatizar muchas de las métricas antes mencionadas (y otras no soportadas por la herramienta SiteSweeper, como análisis de atributos de frames, mantenimiento de color de enlaces, etc.) y que permita además generar informes. Los estrategias de programación utilizadas

para analizar códigos HTML e integrar facilidades se basan en los lenguajes de programación WebL™ (<http://www.research.digital.com/SRC/WebL/index.html>), diseñado específicamente para manipular objetos y atributos en el formato HTML, y el ambiente Java (JDK 2.0™). La herramienta de automatización de métricas se denomina Web-siteMA, y se integrará a un ambiente que comentaremos en el próximo capítulo.

6.6 Generación de Informes.

Como productos a partir del proceso de evaluación elemental debemos considerar a la documentación generada, útil para las actividades de análisis, recomendaciones, y toma de decisión.

Tabla 6.3 Resultados parciales de las preferencias de calidad elemental, para los seis sitios de universidades

	UPC España	Info	Uchile Chile	Info	UTS Australia	Info	NUS Singapur	Info	Stanford USA	Info	UQAM Canada	Info
Usabilidad												
1.1.1.1	100	1	0	0	0	0	0	0	0	0	0	0
1.1.1.2	100	1	0	0	100	1	100	1	100	1	0	0
1.1.1.3	0	0	0	0	100	1	0	0	100	1	0	0
1.1.2	90		90		90		80		90		80	
1.1.3	0	0	0	0	100	1	0	0	100	1	0	0
1.1.4	100	1	100	1	100	1	100	1	50	0.5	100	1
1.2.5.1	0	0	0	0	100	1	100	1	0	0	0	0
1.2.5.2	0	0	0	0	0	0	0	0	0	0	0	0
1.2.5.3	100	1	100	1	0	0	100	1	0	0	100	1
1.4.1	90	90	0	0	0	0	0	0	0	0	90	90
Funcionalidad												
2.1.1.1.1	60	1	100	2	60	1	100	2	100	2	100	2
2.1.1.1.2	0	0	0	0	100	2	0	0	100	2	0	0
2.1.1.1.3	0	0	0	0	0	0	0	0	100	2	100	2
2.1.1.2	60	1	60	1	60	1	0	0	100	2	100	2
2.1.2.1	50	0,5	50	0,5	0	0	0	0	100	1	50	0,5
2.1.2.2	50	0,5	0	0	0	0	0	0	100	1	100	1
2.2.1.1.1	0	0	0	0	0	0	0	0	70	0,7	0	0
2.2.1.1.2	70	1	70	1	70	1	70	1	70	1	100	2
Confiabilidad												
3.1.1.1	0	-29	75.02	75.02	74.1	74.1	68.06	68.06	58.32	58.32	0	-10
3.1.2.2	100	0	100	0	100	0	100	0	100	0	100	0
Eficiencia												
4.1.1	75.3		50.46		82		51.46		100		83.44	
4.2.1.1	0	0	0	0	0	0	0	0	0	0	0	0
4.2.1.2.1	34,38		45,36		81,65		36,22		47,29		53,15	
4.2.1.2.2	80		50		50		60		50		50	

Básicamente, la documentación puede estar compuesta por un documento hipertextual, cuyos nodos son las plantillas de representación de características, subcaracterísticas y atributos (las especificadas en el documento de requerimientos de calidad); la plantilla de referencia de variables y parámetros de cada atributo; y, una vez implementada la

evaluación elemental, la tabla de valores computados para la preferencia de calidad elemental, en caso en que intervenga más de un sistema a evaluar (sino el valor puede quedar registrado en el campo correspondiente de la plantilla). Cada atributo medido tendrá su entrada en una fila de la tabla (que fue implementada en la planilla de cálculo Excel).

La anterior tabla (6.3) muestra algunos de los resultados obtenidos para las preferencias de calidad elemental, para cada uno de los seis sitios de universidades que integraron el caso de estudio (en el Apéndice A, en la sección A.1.2, se encuentran registrados todos los valores).

6.7 Análisis de las Preferencias Elementales.

En la tabla 6.3, registramos de un modo parcial los resultados para las preferencias elementales obtenidas de los atributos que son parte de alguna de las características como *Usabilidad*, *Funcionalidad*, *Confiabilidad* y *Eficiencia*. A pesar que sólo son resultados elementales donde aún no se ha aplicado ningún mecanismo de agregación para computar valores parciales y globales, no obstante, se pueden realizar interesantes observaciones en consideración de los indicadores elementales.

Por ejemplo, se puede observar que dos de los seis sitios investigados, no tiene resuelto un *Esquema de Organización Global*. (Esta subcaracterística -1.1.1- favorece la rápida comprensión tanto de la estructura global como del contenido del sitio, principalmente cuando los visitantes ingresan al nodo raíz de la estructura del sitio Web, -generalmente con topología jerárquica). Es decir, no tienen disponibles atributos como *Mapa del Sitio* (ver 1.1.1.1 en la tabla 6.3) , o *Tabla de Contenidos* (1.1.1.2), o *Indice Global* (1.1.1.3) .

Por otro lado, en consideración de aquellos factores que favorecen a la *Comprensibilidad Global del Sitio* (ver la plantilla correspondiente, en la sección 6.4.2) sólo las Universidades de Stanford y la Tecnológica de Sidney disponen del atributo *Visita Guiada Orientada al Estudiante* (1.1.3). En ambos casos la preferencia de calidad es excelente (logrando un puntaje de 100%), pero la visita implementada en UTS es sencillamente sobresaliente; no sólo dispone de un tour orientado al estudiante sino que además contiene una visita personalizada y personificada para cada unidad académica. (El visitante puede accederla en la tabla de contenidos del sitio, en la etiqueta “*For Students*”, y en el enlace “*Virtual Open One day*”).

Además, todas las universidades disponen del atributo *Mapa de Imagen del Campus* (1.1.4); solamente el mapa de imagen del campus de Stanford no es tan directo y fácil para accederlo (se sale además de contexto) y, por otra parte no está bien estructurado (obteniendo el 50% de la preferencia). Es importante recordar que una marca de preferencia de calidad de 50 % cae en la región de aceptabilidad denominada *marginal*

(en la barra de calidad gris, cuyos valores están en el rango $40 < IE \leq 60$). Por lo tanto, una recomendación es que se debiera planificar una mejora para ese atributo.

Una de las subcaracterísticas deseables a tener en cuenta en el diseño de sitios Web es la de la retroalimentación por parte de los usuarios; mecanismo que podría permitir modificar el sitio a los diseñadores, según las sugerencias y preferencias de los visitantes. Los atributos que modelan este concepto son los *Comentarios y Sugerencias* (1.2.5.3), el *Libro de Visitas* (1.2.5.2), y los *Cuestionarios* (o surveys, codificado como 1.2.5.1). El lector puede apreciar que todos los sitios disponen de al menos uno de estos atributos, excepto el del sitio de la Universidad de Stanford.

Finalmente, otro aspecto de importancia para el dominio de museos y sitios académicos (principalmente para los casos típicos que estamos evaluando), es el soporte a multi-lenguajes. El atributo *Soporte a Lenguajes Extranjeros* (1.4.1), modela este concepto en donde sólo dos sitios académicos cumplen satisfactoriamente con el mismo.

Con respecto a la característica de alto nivel denominada *Funcionalidad*, hay dos funciones principales para moverse dentro de un sitio Web con el fin de encontrar información, a saber: búsqueda y exploración (o browsing). A su vez, en consideración de los mecanismos de búsqueda, identificamos dos tipos: restringida y global. La subcaracterística *Búsqueda Restringida* (2.1.1.1) está compuesta por tres atributos según se aprecia en la figura 5.7. Por ejemplo, nosotros encontramos que todos los sitios tienen, con mayor o menor soporte, el atributo *Búsqueda de Personas* (2.1.1.1.1); sin embargo, no todos los sitios disponen del atributo deseable de *Búsqueda de Cursos* (2.1.1.1.2).

Del mismo modo, el lector puede apreciar en la tabla 6.3, los resultados de los indicadores de calidad elementales para *Enlaces Rotos* (3.1.1.1), correspondiente a la característica de alto nivel denominada *Confiabilidad*; o los atributos *Páginas de Acceso Rápido* (4.1.1), e *Imagen con Título* (4.2.1.2.1), correspondientes a la característica de alto nivel denominada *Eficiencia*.

Capítulo 7

Fase de Definición e Implementación de la Evaluación Global.

7.1 Fase de Evaluación Global

En la fase de *Definición e Implementación de la Evaluación Global* se trata con actividades, modelos, procedimientos y herramientas para determinar los criterios de agregación de las preferencias de calidad elemental (obtenidas en la fase anterior, a partir del árbol de requerimientos), para producir la preferencia global para cada sistema de información interviniente.

Se consideran 17 tipos de funciones lógicas de agregación (que representan diferentes niveles de polarización “y/o”), para modelar diferentes relaciones entre atributos y características, como por ejemplo, relaciones de reemplazabilidad, simultaneidad, neutralidad, simétricas y asimétricas. Al final del proceso se obtiene un valor numérico real (entre 0 y 100), y se establece un ranquin entre los sistemas evaluados. A este valor lo denominamos el indicador de calidad global IG_i para el i -ésimo sistema evaluado.

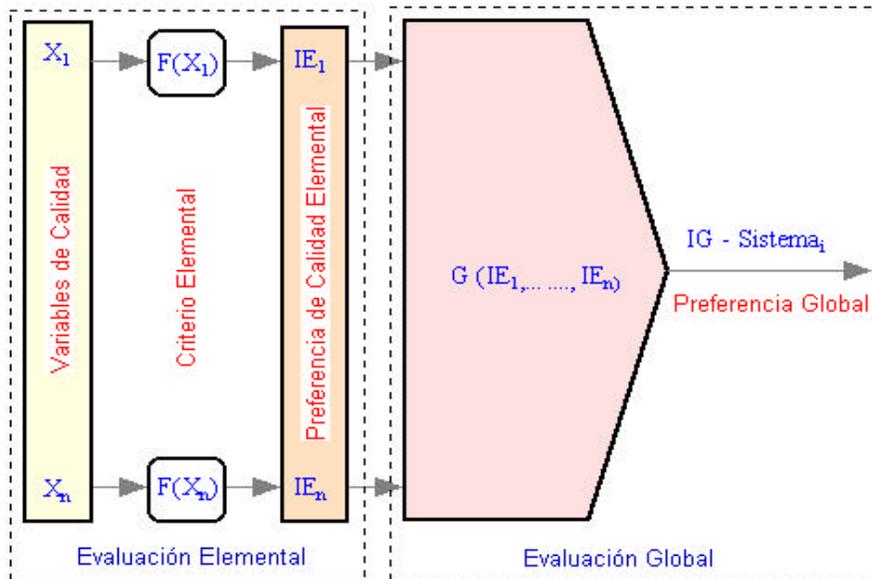


Figura 7.1 Esquema que representa la obtención de la Calidad Global para cada sistema seleccionado a partir de los Indicadores Elementales.

Observando el esquema de la figura 7.1 (ver también fig. 4.1), para n variables asociadas a n atributos directa o indirectamente cuantificables, la correspondiente función elemental produce n indicadores o preferencias elementales. Seguidamente,

aplicando un mecanismo de agregación (o composición) paso a paso, las preferencias elementales se pueden agrupar convenientemente para producir finalmente el indicador o preferencia global. La preferencia de calidad global representa el grado de satisfacción de todos los requerimientos de calidad explícitos e implícitos.

Con todo, para poner en contexto esta discusión, sería oportuno remitirnos a la presentación del problema (sección 3.1), en donde formulamos un conjunto de preguntas y respuestas. Una de las preguntas era la siguiente:

¿Las decisiones y procedimientos fundamentales de la evaluación, comparación y ranquin, están centradas en la intuición de los evaluadores, o están cimentadas en modelos de estructuración de características y atributos, y, además, en modelos y métodos cuantitativos de agregación y cálculo de puntaje?

Luego decíamos:

- ✓ *La evaluación está basada en métricas y modelos de calidad y las decisiones y procedimientos están centrados en modelos y métodos cuantitativos.*
- ✓ *El artefacto consiste, desde el punto de vista de la evaluación, de un buen número de subsistemas o componentes, los que a su vez estos se descomponen en varios elementos. Además, existen relaciones complejas entre elementos y subsistemas.*

Y surgió esta otra cuestión:

- ✓ *Para evaluar la calidad global de un artefacto, para sistemas complejos como la Web, ¿es conveniente basarse en un modelo de agregación de atributos y de puntaje meramente lineal y aditivo, o se requiere algo más robusto y flexible?*

Posteriormente, en la sección 3.1.2.1 presentamos dos modelos de agregación y puntaje. Específicamente, discutimos algunas fortalezas y debilidades de los modelos de puntaje aditivos y lineales con respecto a los modelos de decisión multiatributos no-lineales (remitimos al lector a dicha sección).

En los casos de estudio de museos y sitios académicos usamos un modelo de puntaje lógico para estructurar y computar al indicador de calidad global de cada sitio Web. El mismo responde al enfoque de modelo de decisión multiatributos no-lineales denominado modelo de Agregación Lógica de Preferencias (que es una traducción conveniente de sus siglas en inglés, LSP, o Logic Scoring of Preference) [Dujmovic et al 82; Dujmovic 96].

El presente capítulo está destinado a discutir estos temas y a ejemplificarlos por medio del caso de estudio que estamos desarrollando.

7.2 El Modelo de Agregación Lógica de Preferencias.

Uno de los objetivos de la metodología propuesta Web-site QEM, era el de utilizar para esta fase, y en consideración de la cantidad de características y atributos intervinientes en el proceso de evaluación de sitios Web, un modelo de agregación y cálculo existente que favoreciera a un proceso de evaluación flexible, estructurado, y con fundamentación y objetividad científica, de manera que proveyera indicadores de calidad cuantitativos que pudieran ser usados como base y justificación de las decisiones más óptimas.

El modelo de agregación de atributos, subcaracterísticas y características (y el procedimiento de cálculo) basado en LSP es precisamente bien organizado, estructurado, cuantitativo y robusto. Básicamente, entre las características generales de LSP, podemos enumerar:

- ✓ Es un modelo de agregación y puntaje para evaluar sistemas complejos (el artefacto consiste de un buen número de subsistemas o componentes, los que a su vez se descomponen en varios elementos; asimismo, existen diversos tipos de relaciones entre elementos y subsistemas)
- ✓ Sus resultados representan el grado de satisfacción de los usuarios conforme a los requerimientos de calidad establecidos (la preferencia de calidad del usuario respecto del producto)
- ✓ Es una generalización de los modelos y técnicas de puntaje aditivos y lineales
- ✓ Tiene sus fundamentos en principios y modelos matemáticos y de lógica

En las siguientes subsecciones, trataremos más bien de los aspectos de utilidad práctica de LSP para la estructuración, evaluación y comparación de atributos y características de los artefactos Web, sin ahondar en los fundamentos matemáticos y lógicos. El lector interesado puede remitirse a la siguiente literatura [Dujmovic 74; Dujmovic et al 82]. Además, en [Dujmovic 91] el autor discute cómo el modelo LSP de agregación y evaluación puede ser eficientemente organizado como redes neuronales preferenciales. Las unidades de procesamiento de tales redes, son neuronas preferenciales que implementan a los conectores u operadores lógicos (que veremos en 7.2.2).

7.2.1 Principios del Modelo LSP

Aunque poco conocido en la comunidad científica de informática, particularmente en Ingeniería de Software, el modelo LSP cuenta con más de 25 años. Este modelo es útil para la agregación, evaluación y comparación de sistemas complejos o componentes del mismo, en donde intervienen generalmente más de cuarenta características y atributos para cada sistema a evaluar, como por ejemplo, puede ser el caso en la evaluación y selección de un Sistema de Control, de Sistemas Operativos, de la

Ubicación Óptima de un Edificio (un aeropuerto, una escuela), o en la evaluación de un Hardware y Software específico [Dujmovic 96].

Para introducir los principios del modelo LSP, recurrimos a un ejemplo simple, como se muestra en la figura 7.2. El criterio de agregación tiene como entrada dos indicadores elementales a partir de sendos criterios elementales (según lo comentado en la sección 6.1.1 y representado en las figuras 6.1 y 6.3), a saber: el criterio elemental $CrE(X_1)$, y el $CrE(X_2)$. Esto es, para cada par de valores de las variables X_1 y X_2 se generan dos preferencias o indicadores elementales correspondientes: IE_1 e IE_2 . El problema consiste ahora en expresar la preferencia o indicador global IG_1 como una función de agregación de los indicadores elementales.

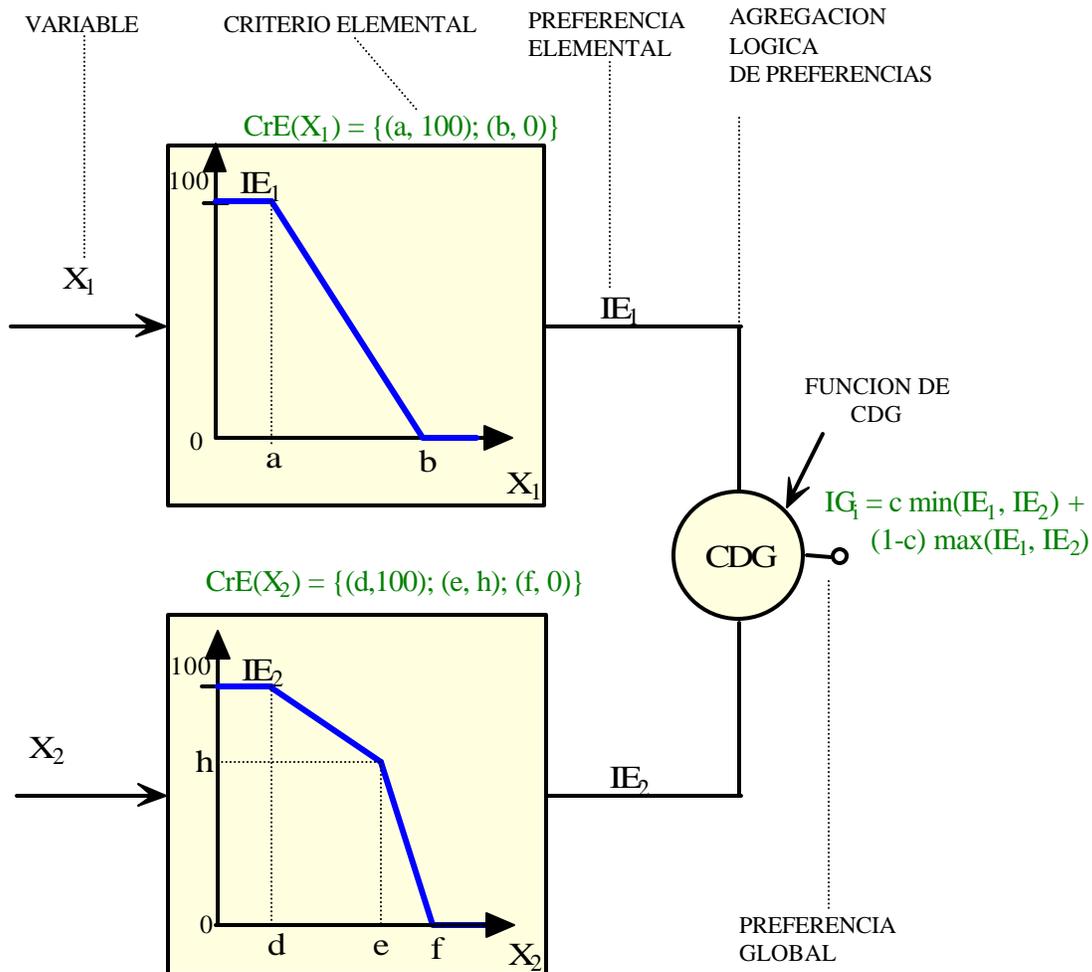


Figura 7.2 Panorama del proceso de determinación de la preferencia de calidad global a partir de preferencias elementales.

A seguir, expresaremos a la función de agregación asumiendo que IE_1 e IE_2 tienen igual importancia (o peso). Una propiedad básica de dicha función es la siguiente desigualdad:

$$\text{Min} (IE_1 , IE_2) \leq IG_i (IE_1 , IE_2) \leq \text{Max} (IE_1 , IE_2) \quad (1)$$

En donde,

$\text{Min} (IE_1 , IE_2) = IE_1 \wedge IE_2$; esto es, la conjunción ;

$\text{Max} (IE_1 , IE_2) = IE_1 \vee IE_2$; esto es, la disyunción

Como indicamos en el capítulo anterior, la interpretación aproximada de IE_1 es que denota el porcentaje del requerimiento elemental satisfecho conforme al valor de X_1 y así para IE_2 . Consecuentemente, la desigualdad (1) puede verbalmente ser expresada como: a) La satisfacción global de un conjunto de requerimientos no puede ser más grande que la satisfacción del requerimiento elemental más satisfecho; y b) La satisfacción global de un conjunto de requerimientos no puede ser menor que la satisfacción del requerimiento elemental menos satisfecho. O dicho en otras palabras: la preferencia global no puede ser más preferida (o mejor) que su parte más preferida (la mejor), ni puede ser menos preferida (o peor) que su parte menos preferida (la peor).

No obstante, los casos extremos son posibles. El caso conjuntivo, esto es, cuando $IG_i = \text{Min} (IE_1 , IE_2)$ representa la situación en donde todos los grados de satisfacción (o preferencias elementales) pueden ser menos considerados, excepto el mínimo. Lo que significa que “*una cadena es tan resistente como lo es su eslabón más débil*”; es decir, se desea la satisfacción simultánea de todos los requerimientos. Por lo tanto, no hay modo de compensar la parte más débil de un elemento de un componente, mejorando el resto de los elementos interrelacionados del componente.

Por el contrario, el caso disyuntivo, esto es, cuando $IG_i = \text{Max} (IE_1 , IE_2)$ representa la situación en donde todos los grados de satisfacción (o preferencias) pueden ser menos considerados, excepto el máximo. Lo que significa que el componente de un sistema es considerado tan bueno como su mejor parte (o elemento del componente). Por lo tanto, no hay modo de mejorar la preferencia global de un componente (o sistema) mejorando su parte más débil.

Sin embargo, los casos extremos son raramente aplicados en la evaluación pero sí grados de conjunción y disyunción.

7.2.1.1 Grados de Conjunción y Disyunción. En casi todas las situaciones prácticas el criterio de evaluación debe tener la propiedad que al mejorar una parte del sistema debe producir una mejora en el sistema como un todo. Esto significa, conforme a la discusión anterior, que el indicador resultante IG_i debe estar entre alguno de los valores extremos de IE_1 e IE_2 . Si el criterio de agregación fuera uno con propiedades conjuntivas, la distancia entre IG_i y el $\text{Min}(IE_1 , IE_2)$, debe ser menor que la distancia entre IG_i y el $\text{Max}(IE_1 , IE_2)$. De igual modo, si el criterio de agregación fuera uno con propiedades disyuntivas, la distancia entre IG_i y el $\text{Max}(IE_1 , IE_2)$, debe ser menor que

la distancia entre IG_i y el $\text{Min}(IE_1, IE_2)$. Este nivel de intensidad o influencia del mínimo y del máximo sobre el punto que representa al indicador global, debe ser ajustable.

Con el fin de ajustar las propiedades lógicas deseadas de la función de agregación de preferencias, se introduce el concepto de un parámetro ajustable c denominado el grado de conjunción (como así también el parámetro ajustable d denominado el grado de disyunción). El mismo especifica la posición de IG_i con respecto del $\text{Min}(IE_1, IE_2)$ y el $\text{Max}(IE_1, IE_2)$.

Si $c=1$ (y $d=0$), entonces IG_i representa el valor resultante de una función de conjunción pura; si $c=0$ (y $d=1$), entonces IG_i representa el valor resultante de una función de disyunción pura ($IG_i = \text{Max}(IE_1, IE_2)$). Finalmente, si $0 < c < 1$ entonces $\text{Min}(IE_1, IE_2) < IG_i(IE_1, IE_2) < \text{Max}(IE_1, IE_2)$.

La función de agregación ajustable que tiene esas propiedades se denomina “*Función de Conjunción-Disyunción Generalizada*” (CDG). Este nombre refleja el hecho que tanto la conjunción como la disyunción son casos especiales de la función CDG. Esta función debe ser capaz de expresar un espectro de funciones que son similares a la conjunción, y del mismo modo, un rango similar a la disyunción. Las funciones que tienen la propiedad en la que IG_i está afectada predominantemente por el $\text{Min}(IE_1, IE_2)$ se llaman *cuasi-conjunción* (CC). Del mismo modo, las funciones que tienen la propiedad en la que IG_i está afectada predominantemente por el $\text{Max}(IE_1, IE_2)$ se llaman funciones de *cuasi-disyunción* (CD).

Una de las formas más simples de representar a la función CDG es mediante la combinación lineal de ambos componentes:

$$IG_i = c \text{Min}(IE_1, IE_2) + (1 - c) \text{Max}(IE_1, IE_2); \quad (2)$$

o, considerando que $c + d = 1$ la ecuación queda así:

$$IG_i = c \text{Min}(IE_1, IE_2) + d \text{Max}(IE_1, IE_2)$$

En el caso particular en que $c=d=1/2$ se obtiene la media aritmética, es decir, la media aritmética está en el medio entre la conjunción y la disyunción. Para representar a las funciones de CC y CD es conveniente definir varios niveles discretos de valores de grados de conjunción y disyunción. Las funciones de CC se caracterizan por tener los parámetros c y d entre $1/2 < c < 1$ y $0 < d < 1/2$; similarmente, las de CD entre $1/2 < d < 1$ y $0 < c < 1/2$.

La figura 7.3 representa geoméricamente dicha situación. Se pueden observar 9 funciones teniendo en cuenta incrementos de los parámetros d y c en valores de 0,125.

La cuasi-conjunción es representada por tres funciones de agregación (C- ; CA ; C+); y la cuasi-disyunción es representada por tres funciones de agregación D- ; DA ; D+ respectivamente (en la sección 7.2.2, veremos un modelo con 17 funciones).

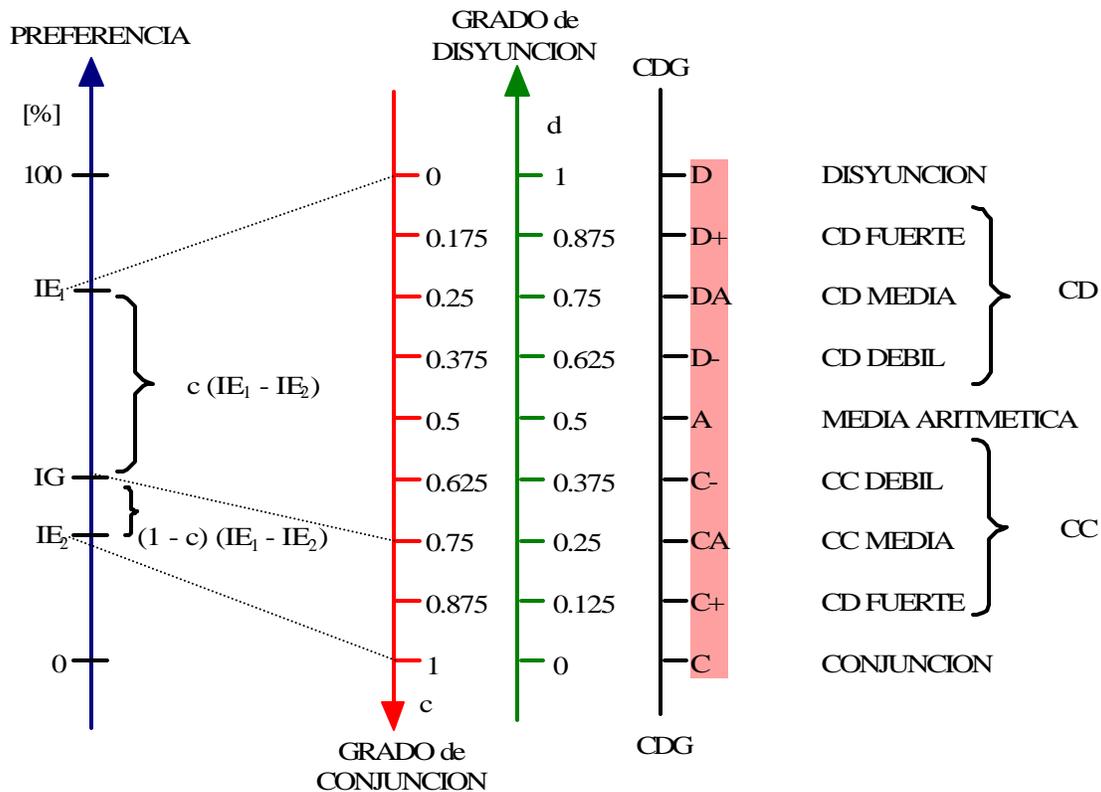


Figura 7.3 Interpretación geométrica de la función CDG para el caso más simple de dos preferencias elementales con igual peso.

7.2.1.2 Algunas Propiedades Deseables de la Función CDG. La versión anterior de la función CDG (2) está destinada esencialmente a introducir los conceptos de conjunción, disyunción, y grados de conjunción y disyunción, pero no es una ecuación de mucho valor práctico para la resolución de problemas de evaluación reales. Se requiere un conjunto de propiedades de la misma para poder agrupar atributos y características que representan requerimientos obligatorios (mandatorios), como así también aspectos de la importancia relativa de dichos requerimientos. Algunas propiedades deseables de la función CDG deben ser, a saber:

- ✓ Debe ser posible agrupar más de dos indicadores o preferencias elementales
- ✓ Debe proveer un modelo para representar requerimientos mandatorios (si el requerimiento obligatorio no es satisfecho, la función debe producir una preferencia parcial, o global de cero)
- ✓ Debe proveer un nivel ajustable de importancia relativa (o peso) de cada preferencia de entrada

Estas son sólo algunas de las propiedades. Dujmovic ha demostrado en sus investigaciones, hace más de veinte años, que la función denominada *media de potencia pesada* (o, en inglés, weighted power mean), satisface esas condiciones. Supongamos que tenemos que producir la preferencia parcial o global, a partir de m preferencias elementales. La función de agregación debe satisfacer:

- 1) Cada indicador elemental IE_i debe tener asociado un peso P_i
- 2) La preferencia resultante tiene un valor entre
 $\text{Min}(IE_1, \dots, IE_m) \leq IG_i \leq \text{Max}(IE_1, \dots, IE_m)$

Las propiedades antes mencionadas de la función CDG se pueden satisfacer por medio de la media de potencia pesada, que se expresa del siguiente modo:

$$IG(r) = (P_1 IE_1^r + P_2 IE_2^r + \dots + P_m IE_m^r)^{1/r}; \quad (3)$$

$$-\infty \leq r \leq +\infty; 0 \leq IE_i \leq 1;$$

$$(P_1 + P_2 + \dots + P_m) = 1; P_i > 0; i = 1 \dots m; \quad (4)$$

$$IG(-\infty) = \text{Min}(IE_1, IE_2, \dots, IE_m);$$

$$IG(+\infty) = \text{Max}(IE_1, IE_2, \dots, IE_m);$$

La ubicación de $IG(r)$ entre el valor mínimo y máximo se puede ajustar seleccionando el valor de r . Si el valor de $IG(r)$ está más cercano al mínimo entonces tal criterio especifica el requerimiento para la simultaneidad de entradas (solamente altos valores en las entradas producen un alto valor en la salida). Si el valor de $IG(r)$ está más cercano al máximo entonces tal criterio especifica el requerimiento para la reemplazabilidad de entradas (un alto valor en una de las entradas produce un alto valor en la salida).

Por lo tanto, ahora los parámetros ajustables son el valor de r (que ya están tabulados, y que se determinan a partir de la cantidad de entradas y como una función de c y d , antes introducidos), y los pesos P . La elección del valor de r permite seleccionar un operador que va desde la pura conjunción, la cuasi-conjunción hasta la cuasi-disyunción y la pura disyunción. El punto medio representa a la aditividad (para $r = 1$) la cual no está ni conjuntiva ni disyuntivamente polarizada, es decir, representa una función de relaciones de neutralidad o independencia entre las entradas. Para valores de r menores o iguales a cero se modela mediante (3) requerimientos mandatorios.

Por último, los detalles matemáticos y su relación con la lógica de preferencia continua se encuentran en la literatura previamente citada. A seguir, discutiremos el empleo de los operadores de LSP para modelar distintas relaciones entre características, subcaracterísticas y atributos. (En la sección 7.4 daremos algunos detalles del procedimiento y herramientas de soporte para el cálculo de las preferencias).

7.2.2 Empleo de los Operadores de LSP para modelar Relaciones Lógicas entre Características y Atributos

Una de las principales fortalezas del modelo LSP con respecto al modelo meramente aditivo y lineal citado, reside en la potencialidad de modelar diferentes relaciones lógicas entre atributos y subcaracterísticas de manera que reflejen las necesidades de los diferentes participantes en el proceso de evaluación.

Las principales relaciones lógicas se definen del siguiente modo:

- ✓ *Simultaneidad* (o relación de conjuntividad): cuando los participantes en el proceso de evaluación perciben que dos o más entradas deben estar presentes simultáneamente
- ✓ *Reemplazabilidad* (o relación de disyuntividad): cuando los participantes en el proceso de evaluación perciben que dos o más entradas puede estar presentes alternativamente (por ej., la presencia de un atributo puede reemplazar a la ausencia de otro)
- ✓ *Neutralidad* (o relación ni de conjuntividad ni de disyuntividad): cuando se percibe que dos o más preferencias de entrada pueden agruparse de un modo independiente
- ✓ *Relación Simétrica*: cuando se percibe que dos o más preferencias de entrada afectan de la misma manera lógica aunque con diferentes grados de importancia
- ✓ *Relación Asimétrica*: cuando se requiere modelar requerimientos mandatorios combinados con requerimientos no-mandatorios (atributos obligatorios se combinan con otros deseables y/u opcionales), o cuando condiciones necesarias se combinan con condiciones suficientes.

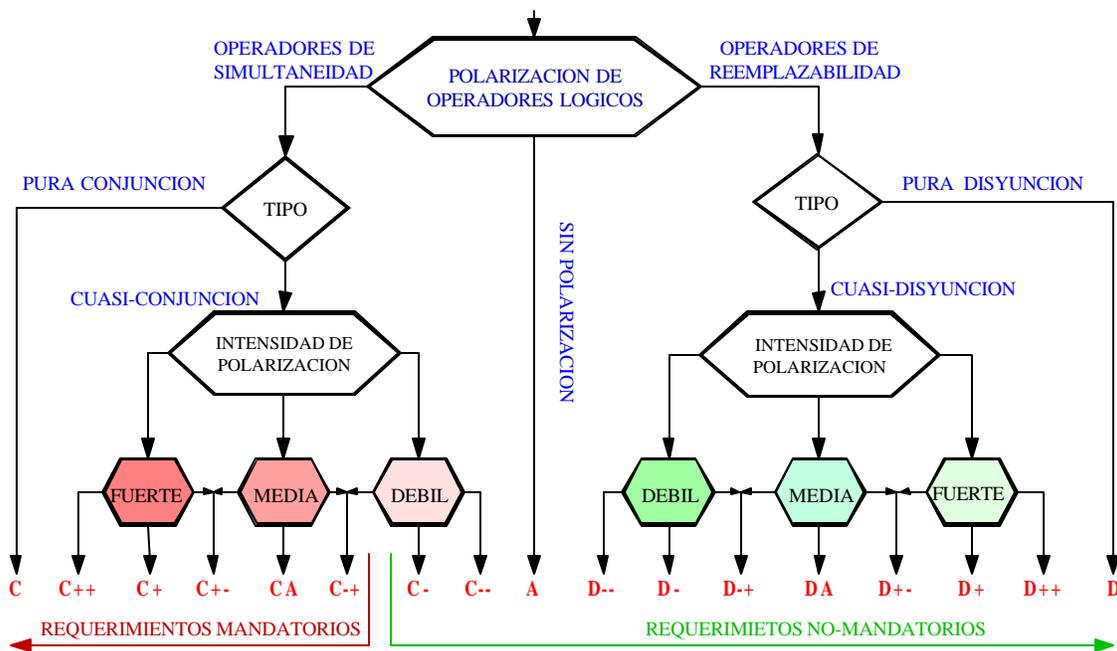


Figura 7.4 Operadores lógicos Conjuntivos y Disyuntivos de LSP y niveles de polarización

La figura 7.4 muestra un modelo de 17 niveles de operadores o conectores lógicos que representan a la función CDG (aunque se podría utilizar un modelo de 25 niveles si se requiriera mayor precisión).

Los principales operadores de LSP son la media aritmética (A) que modela la relación de neutralidad. Dentro de los operadores conjuntivos encontramos el operador (C) que modela a la conjunción pura, y dentro de las funciones de la cuasi-conjunción se encuentran tres niveles de intensidad: débil (C-), medio (CA), y fuerte (C+). Es importante destacar, que los operadores lógicos de cuasi-conjunción representan conectores “y” flexibles. Además de los conectores de cuasi-conjunción descriptos, se cuenta con operadores de valores intermedios; por ejemplo, el operador C-- se posiciona entre A y C-; y el operador C++ está entre CA y C-, y así sucesivamente, como se aprecia en el esquema de la figura 7.4.

Los operadores anteriores (excepto el operador A) significan que dado un bajo valor de una preferencia de entrada nunca puede ser bien compensada por un alto valor de alguna otra entrada para producir una preferencia de salida alta. Sin embargo, dado el nivel de intensidad en la polarización “y” no todos los operadores castigan con la misma fuerza la preferencia de salida. En la figura anterior se puede apreciar que varios operadores de la CC modelan requerimientos mandatorios, y dos de ellos no (los operadores C- y C--); es decir, un cero en una de las entradas no producirá un cero en la salida.

Igualmente a los operadores conjuntivos, también podemos utilizar a los operadores de cuasi-disyunción para modelar relaciones entre entradas, en rangos de intensidad de polarización como el mostrado en la figura 7.4. Estos operadores modelan relaciones de reemplazabilidad, en donde pueden existir entradas alternativas; es decir, un bajo valor de una preferencia de entrada siempre puede ser bien compensado por un alto valor de alguna otra preferencia de entrada para producir una preferencia de calidad alta.

7.2.2.1 Tipos de Funciones de Agregación. A partir de la combinación de los operadores lógicos descriptos en la sección previa, las funciones de agregación de preferencia se pueden clasificar en simples y compuestas. Las funciones de agregación simples modelan relaciones de entradas simétricas, en tanto que las funciones de agregación compuestas modelan relaciones de entrada asimétricas.

En la figura 7.5 observamos tres ejemplos del primer tipo. El diagrama de la izquierda representa a una función de tres entradas a un conector de cuasi-conjunción. Se utiliza el operador C++ (que modela requerimientos mandatorios); asimismo cada entrada tiene su respectivo peso, según la restricción (4). En el diagrama del centro se muestra a una función de neutralidad con tres entradas, utilizando el operador A, que modela requerimientos independientes de un grupo. Por último, el dibujo de la derecha

representa una función de tres entradas, con sus respectivos pesos, a una función de cuasi-disyunción utilizando el operador DA, que modela requerimientos alternativos de un grupo de preferencias.

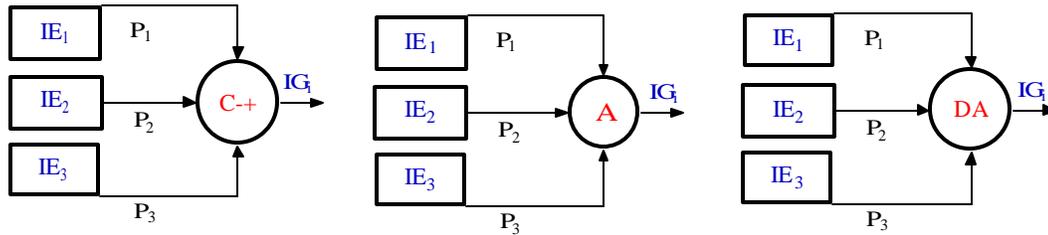


Figura 7.5 Tres funciones simples de agregación de preferencias.

En la figura 7.6 observamos dos ejemplos de funciones compuestas o asimétricas. El diagrama de la izquierda representa una función *conjuntiva de absorción parcial*. En este caso tenemos una entrada con un requerimiento mandatorio (la entrada IE_1); y otra no-mandatoria (la entrada IE_2 representa un requerimiento deseable). Si $IE_1 = 0$ entonces la salida será cero, independientemente del valor de la entrada IE_2 debido al operador CA. No obstante, si $IE_2 = 0$ e $IE_1 > 0$ entonces $0 < IG < IE_1$. Además, una buena calidad en el indicador elemental IE_2 (es decir, $IE_2 > IE_1$) produce una salida tal que $0 < IE_1 < IG < IE_2$.

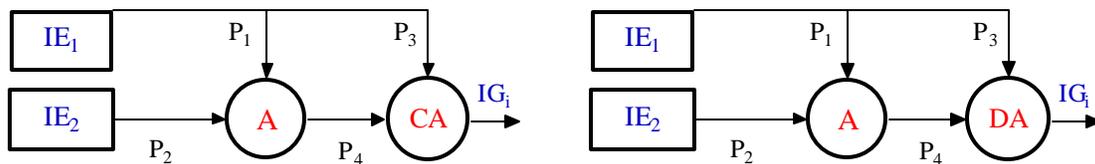


Figura 7.6 Dos funciones compuestas de agregación de preferencias.

El diagrama de la derecha representa una función *disyuntiva de absorción parcial*. En este caso, una buena calidad del indicador elemental IE_1 es suficiente para producir un alto valor en la salida IG , sin considerar el valor de la entrada del indicador IE_2 . Si $IE_1 = 0$ entonces la salida no será cero siempre y cuando el valor de la entrada $IE_2 > 0$, es decir, la entrada IE_2 puede parcialmente compensar a la de IE_1 .

Finalmente, es posible una combinación de funciones anidadas para modelar relaciones de mayor complejidad. Sin embargo, para los casos de estudio realizados utilizamos funciones de agregación simple, como veremos en las próximas sesiones.

7.3 Agregación de las Preferencias Elementales para Producir un Indicador de Calidad Global.

En esta tarea del proceso, los evaluadores deben definir, preparar e implementar la

evaluación global con el fin de obtener un indicador de calidad para cada sistema seleccionado. Aplicando un mecanismo de agregación paso a paso, las preferencias de calidad elementales deben estructurarse y agregarse de un modo de abajo hacia arriba (estrategia bottom-up) para permitir el ulterior cálculo de las preferencias parciales respectivas. A su vez, repitiendo el proceso de agregación recursivo al final puede obtenerse la estructura de agregación de todo el sistema.

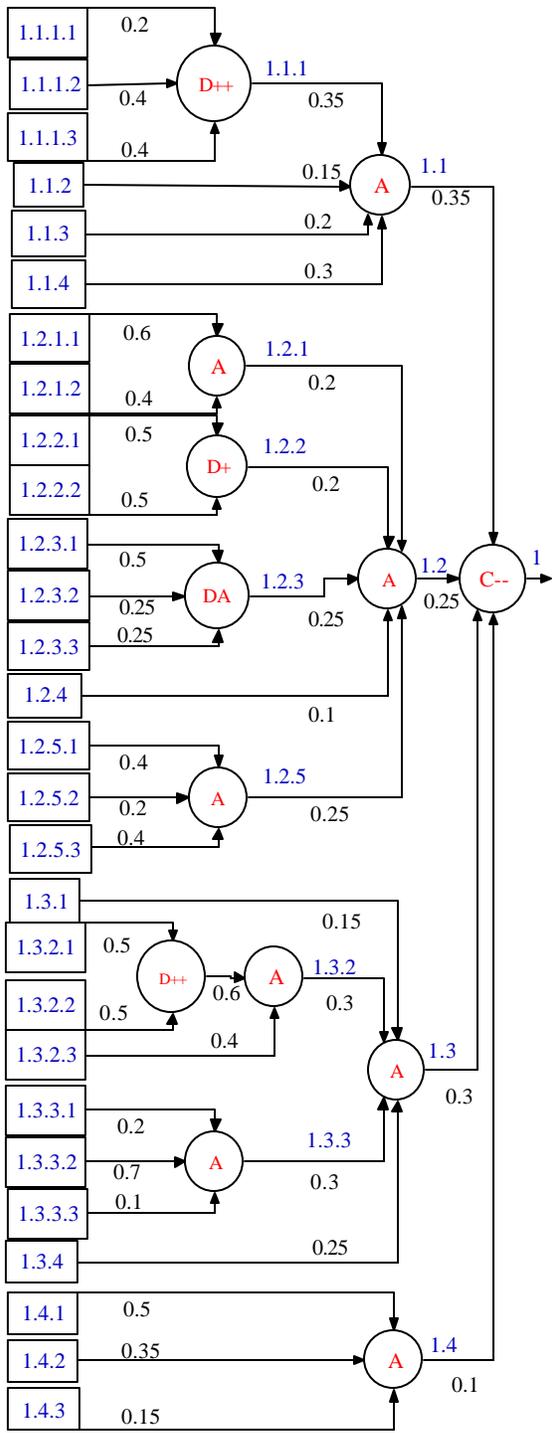
Finalmente, una vez establecida la estructura, se debe llevar a cabo el proceso de cómputo de las preferencias de calidad. Para esta actividad se hace uso de herramientas de cálculo automáticas. La preferencia de calidad global representa el grado de satisfacción de todos los requisitos involucrados.

7.3.1 Consideraciones sobre las Actividades en el Proceso de Agregación

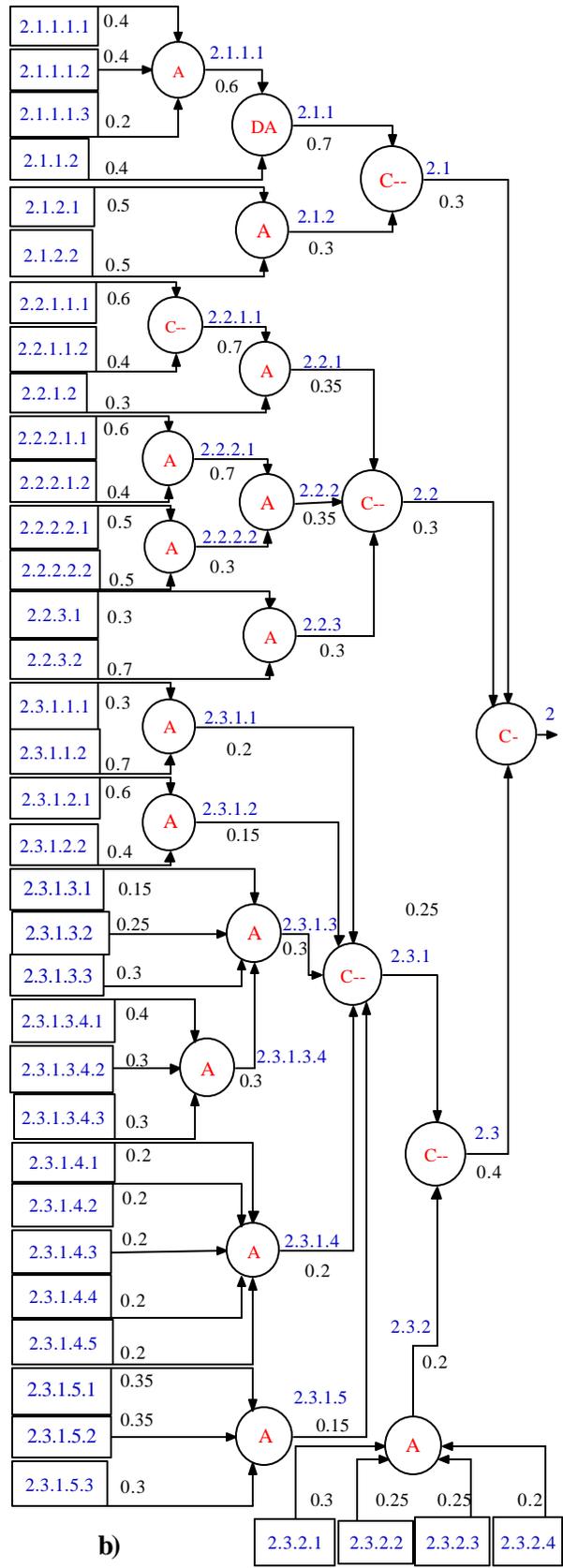
En el proceso de agregación de las preferencias elementales, parciales y global al menos se identifican las siguientes actividades básicas:

- ✓ *Selección del tipo de relación lógica entre elementos y/o componentes:* es decir, la cuestión básica a considerar consiste en preguntarse: ¿es la relación entre las entradas con una polarización lógica conjuntiva, disyuntiva o neutra?
- ✓ *Selección del tipo de función conforme a la relación de entradas entre elementos y/o componentes:* es decir, la cuestión básica a considerar consiste en decidir ¿es la relación de entradas simétrica o asimétrica?
- ✓ *Selección del operador conforme al nivel de intensidad de la polarización lógica*
- ✓ *Selección de la importancia relativa de cada entrada conforme a los requerimientos* (esto es, la selección de los pesos)

Las estrategias y mecanismos para realizar estas actividades puede ser más o menos intuitivo, conforme a la experticia de los evaluadores y el nivel de criticidad del proyecto de evaluación. Por ejemplo, un evaluador puede valorar la relativa importancia de los elementos de entrada a las funciones de acuerdo a su intuición y experticia, o puede utilizar mecanismos como encuestas y establecer fórmulas de relativa importancia para computar los pesos. (En el punto 5.4.4 se describió el mecanismo empleado para el caso de estudio de sitios académicos, principalmente para las características de más alto nivel. Para muchos de los atributos se utilizó una estrategia intuitiva, dado el esfuerzo que implicaba realizar la encuesta para todos los elementos intervinientes). Del mismo modo que para los pesos, se puede utilizar una estrategia intuitiva para la determinación de los tipos de funciones y los niveles de polarización lógica, o cuando la evaluación requiere mayor precisión, se pueden realizar estudios de análisis y sensibilidad.



a)



b)

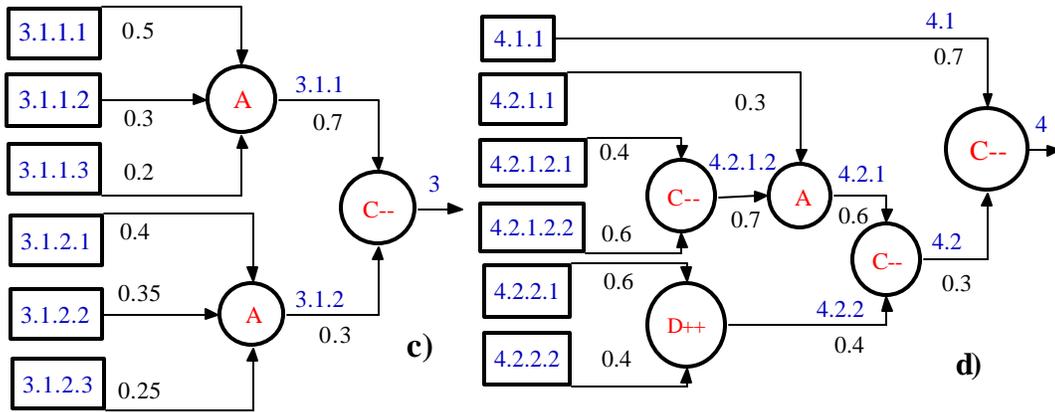


Figura 7.7 Estructura de agregación de preferencias parciales usando el modelo LSP para el caso de estudio de sitios académicos. En la parte (a) se muestra la estructura de agregación para la característica de alto nivel denominada Usabilidad; en la parte (b) se muestra la estructura de agregación para la característica Funcionalidad; en la parte (c) se muestra la estructura para la característica Confiabilidad y, en (d), para Eficiencia

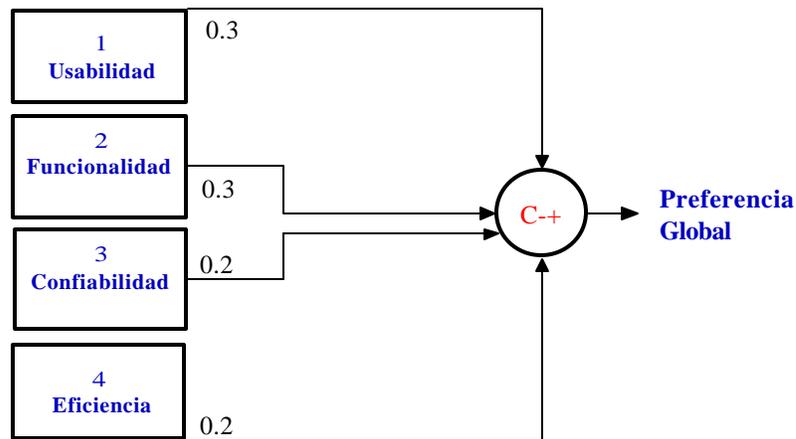


Figura 7.8 Estructura de agregación de preferencias parciales para las características de más alto nivel, útil para computar el indicador de calidad global IG para cada sitio Web.

En la figura 7.7 mostramos sendas estructuras de agregación de preferencias para la característica de alto nivel denominada *Usabilidad* (parte a); la estructura de agregación para la característica *Funcionalidad* (en la parte b); en (c) para la característica *Confiabilidad*, y por último, en (d), para *Eficiencia*. Finalmente, en la figura 7.8 se puede apreciar la estructura de agregación de preferencias parciales para las características de más alto nivel (codificadas 1, 2, 3, y 4 respectivamente, según el árbol de requerimientos de las figuras 5.6 y 5.7). Esta estructura producida es útil para cuantificar el indicador de calidad global IG para cada sitio académico en la Web.

Por ejemplo, en la figura 7.7 (a), al final del proceso de agregación el lector puede observar la subcaracterística codificada 1.1 (a partir del árbol de requerimientos, y denominada *Comprensibilidad Global del Sitio*), con una importancia relativa o peso de 0.35; y la subcaracterística 1.2 (*Mecanismos de Ayuda y Retroalimentación en línea*), con una importancia relativa de 0.25; y la subcaracterística 1.3 (*Aspectos de Interface y Estéticos*), con un peso de 0.3; y, por último, la subcaracterística 1.4 (*Misceláneas*) con un peso de 0.1. Todos estas preferencias de calidad de las subcaracterísticas sirven de entrada a la función lógica C-- la que produce como salida a la preferencia global parcial codificada 1, (*Usabilidad*). El operador de polarización conjuntiva de baja intensidad C-- no modela requerimientos obligatorios, es decir, un cero en una de las entradas no genera un cero en la salida aunque castiga el resultado.

En el ejemplo anterior se ha modelado una relación de conjunción dado que se percibe que las cuatro entradas deben estar presentes simultáneamente aunque con distinta importancia relativa. Tomemos un caso de relación de disyuntividad, dado que los evaluadores consideran (a partir de estudios) que no necesariamente todos los atributos de un grupo deben estar simultáneamente, es decir; pueden ser reemplazables unos con otros. Como discutimos en la sección 6.7 podía ser suficiente o un *Mapa del Sitio* (ver 1.1.1.1 en la figura 7.7 a), o una *Tabla de Contenidos* (1.1.1.2), o un *Índice Global* (1.1.1.3). Estas tres preferencias elementales de calidad sirven de entrada a la función disyuntiva fuertemente polarizada D++ la que produce como salida, a la preferencia parcial 1.1.1 (*Esquema de Organización Global*). En cuanto a la importancia relativa de los atributos, un *Mapa del Sitio*¹ resultó tener menos peso, dado que puede ser generado por herramientas automáticas, o, en sitios muy grandes, podría ser engorroso mostrar visualmente toda la estructura arquitectónica del sitio.

Por otro lado, para ejemplificar una relación de neutralidad (ni conjuntiva ni disyuntivamente polarizada) podemos observar la salida 1.1 del operador lógico A (en la figura 7.7a) que modela a la subcaracterística *Comprensibilidad Global del Sitio*. Observamos una estructura de agregación compuesta por la salida de 1.1.1 (analizada en el párrafo anterior), y que ahora sirve de entrada al conector respectivo, junto con las preferencias de los atributos 1.1.2 (*Calidad del Sistema de Etiquetado*), 1.1.3 (*Visita Guiada Orientada al Estudiante*), y 1.1.4 (*Mapa de Imagen del Campus*). En este caso se modela una relación de independencia entre las entradas.

Finalmente, en la figura 7.8 se puede apreciar la estructura de agregación de las preferencias de calidad de las características de más alto nivel para producir la preferencia de calidad para cada sitio a evaluar. En consideración del perfil de usuario y dominio de evaluación, las características *Usabilidad* y *Funcionalidad* son relativamente más importantes que las características *Confiabilidad* y *Eficiencia* según arrojaron los estudios de campo realizados con estudiantes (sección 5.4.4).

¹ Un mapa del sitio es una representación con componentes gráficos, que muestra la estructura o arquitectura global (a menudo jerárquica) del sitio Web como un todo

7.3.2 Cómputo de las Preferencias Parciales y Globales

Una vez que fueron estructurados y acordados todos los criterios y funciones de agregación, los tomadores de decisión deben ejecutar el programa que calcula las preferencias de calidad parciales y globales para cada sistema participante. La tabla 7.1 muestra los resultados detallados de las preferencias de calidad parciales y globales (para subcaracterísticas y características), para cada sitio académico seleccionado, a partir de los atributos medidos y traducidos en unidades de preferencia, en la fase previa de evaluación elemental.

El indicador de calidad global IG para cada sitio se encuentra al final de la tabla. El proceso de evaluación ha ranqueado primero al sitio de la Universidad de Stanford con un 79.76% de la preferencia de calidad global (entrando en la barra de calidad verde), y ha ranqueado en último lugar al sitio de la Universidad Nacional de Singapur con un 54.46% de la preferencia de calidad global (entrando en la barra de calidad gris, o, en otras palabras, en un nivel de aceptabilidad marginal).

En el capítulo 8 profundizaremos la discusión respecto de los resultados obtenidos con el fin de analizar el estado del arte en la calidad de los sitios Web evaluados y dar recomendaciones y proponer potenciales mejoras.

Tabla 7.1 Resultados detallados de los valores de las preferencias parciales y globales de calidad, para los seis sitios Web de las universidades evaluadas

Características y Subcaracterísticas	UPC	UChile	UTS	NUS	Stanford	UQAM
1. Usabilidad	76.18	51.01	80.08	57.71	71.93	60.94
1.1 Comprensibilidad Global del Sitio	77.76	43.5	98.17	75.70	83.17	42
1.1.1 Esquema de Organización Global	97.91	0	99.08	96.29	99.08	0
1.2 Mecanismos de Ayuda y Retroalimentación en línea	70.15	44.6	73.8	75.8	69.4	68.79
1.2.1 Calidad de la Ayuda	70	48	94	54	82	76
1.2.2 Indicador de Última Actualización	55.78	0	100	100	100	92.97
1.2.3 Directorio de Direcciones	100	100	100	100	100	100
1.2.5 Retroalimentación	40	40	40	80	0	40
1.3 Aspectos de Interfaces y Estéticos	80.6	69.72	83.11	34.9	75.71	64.10
1.3.2 Permanencia y Estabilidad en la Presentación de los Controles Principales	88	78.41	98.04	30	90.04	90.01
1.3.3 Aspectos de Estilo	64	54	54	18	54	62
1.4 Misceláneas	73	43	35	35	35	80
2. Funcionalidad	61.84	50.39	48.89	38.99	77.29	71.12
2.1 Aspectos de Búsqueda y Recuperación	48.59	42.37	29.32	19.73	100	80.55
2.1.1 Mecanismo de Búsqueda en el Sitio	48	50.84	52.17	35.12	100	82.98
2.1.1.1 Búsqueda Restringida	24	40	44	40	100	60
2.1.2 Mecanismos de Recuperación	50	25	0	0	100	75
2.2 Aspectos de Navegación y Exploración	62.01	50.09	57.66	38.80	74.02	50.89
2.2.1 Navegabilidad	26.15	23.15	29.15	29.15	79	45.92
2.2.1.1 Orientación	15.93	15.93	15.93	15.93	70	22.75

2.2.2 Objetos de Control Navegacional	81.1	60.1	88.6	19.5	65.2	51.7
2.2.2.1 Permanencia y Estabilidad en la Presentación de los Controles Contextuales (Subsitio)	88	58	88	0	76	46
2.2.2.2 Nivel de Desplazamiento	65	65	90	65	40	65
2.2.3 Predicción Navegacional	93	78	63	84	93	56
2.3 Aspectos del Dominio orientados al Estudiante	73.46	57.43	61.80	61.87	73.13	82.37
2.3.1 Relevancia de Contenido	81.38	82.36	71.05	63.64	77.97	78.17
2.3.1.1 Información de Unidad Académica	100	100	100	100	100	69.99
2.3.1.2 Información de Inscripción	60	60	0	100	60	60
2.3.1.3 Información de Carreras	88	91	91	15	52	82
2.3.1.3.4 Descripción de Cursos	60	70	70	0	40	40
2.3.1.4 Información de Servicios al Estudiante	80	60	100	80	100	80
2.3.1.5 Información de Infraestructura Académica	70	100	100	100	100	100
2.3.2 Servicios On-line	45	0	30	55	55	100
3. Confiabilidad	60.40	87.62	90.86	88.70	83.05	63.61
3.1 No Deficiencia	60.40	87.62	90.86	88.70	83.05	63.61
3.1.1 Errores de Enlaces	50	82.51	87.05	84.03	76.16	50
3.1.2 Errores o Deficiencias Varias	87.5	100	100	100	100	100
4. Eficiencia	69.09	52.03	76.11	53.12	85.99	69.47
4.1 Performancia	75.3	50.46	82	51.46	100	83.44
4.2 Accesibilidad	55.41	52.69	66.11	57.10	59.52	40.84
4.2.1 Accesibilidad de Información	42.02	29.47	47.84	34.95	38.33	35.87
4.2.1.2 Legibilidad al desactivar la Propiedad Imagen del Browser	60.03	42.11	68.35	49.94	54.77	51.25
4.2.2 Accesibilidad de Ventanas	78.04	96.38	97.55	97.55	97.55	48.77
Preferencia de Calidad Global	66.91	56.55	69.61	54.46	79.76	66.05

7.4 Procedimientos y Herramientas de Soporte

7.4.1 Algunos Detalles del Procedimiento para el Cómputo de las Preferencias de Calidad

En esta sección, presentamos algunos detalles de la interface de la función para el cálculo de las preferencias parciales y globales en el proceso de evaluación, conforme al modelo LSP.

En el Apéndice D.1, se muestra un módulo codificado en C++, que es parte del programa para obtener las preferencias parciales y globales. Particularmente, el módulo mostrado sirve para computar las funciones lógicas simétricas, a partir de los operadores lógicos de LSP introducidos en la sección 7.2.2. La siguiente función denominada WPM (media de potencia pesada):

```
double WPM (int n, double P[], double r, double x[])
```

tiene como interface los siguientes parámetros formales: **n** representa a la cantidad de

entradas; el parámetro $x[]$ ($x[0 \dots n-1]$) que representa al arreglo que guarda los valores de las preferencias de entrada, previamente ingresados, al igual que el arreglo $P[]$ ($P[0 \dots n-1]$) que representa a la importancia relativa de cada entrada, conforme a la ecuación (4). Finalmente, el parámetro r que es el exponente de la ecuación (ver (3)), guarda el valor real conforme al operador lógico y a la cantidad de entradas seleccionadas para una función de agregación dada. La siguiente tabla muestra los valores parametrizados de r para una función CDG de 17 niveles, para 2, 3, 4 y 5 entradas respectivamente.

Tabla 7.2 *Función de Conjunción-Disyunción Generalizada de 17 Niveles y valores del parámetro r para 2,3,4 y 5 entradas [Dujmovic 74]*

Operador LSP	Abrev	c	d	r(2)	r(3)	r(4)	r(5)	Mandat
Disyunción	D	0.0000	1.0000	$+\infty$	$+\infty$	$+\infty$	$+\infty$	No
CD Fuerte (+)	D++	0.0625	0.9375	20.63	24.30	27.11	30.09	No
CD Fuerte	D+	0.1250	0.8750	9.521	11.095	12.27	13.235	No
CD Fuerte (-)	D+-	0.1875	0.8125	5.802	6.675	7.316	7.819	No
CD Media	DA	0.2500	0.7500	3.929	4.450	4.825	5.111	No
CD Débil (+)	D-+	0.3125	0.6875	2.792	3.101	3.318	3.479	No
CD Débil	D-	0.3750	0.6250	2.018	2.187	2.302	2.384	No
CD Débil (-)	D--	0.4375	0.5625	1.449	1.519	1.565	1.596	No
Media Aritmét	A	0.5000	0.5000	1.000	1.000	1.000	1.000	No
CC Débil (-)	C--	0.5625	0.4375	0.619	0.573	0.546	0.526	No
CC Débil	C-	0.6250	0.3750	0.261	0.192	0.153	0.129	No
CC Débil (+)	C-+	0.6875	0.3125	-0.148	-0.208	-0.235	-0.251	Si
CC Media	CA	0.7500	0.2500	-0.720	-0.732	-0.721	-0.707	Si
CC Fuerte (-)	C+-	0.8125	0.1875	-1.655	-1.550	-1.455	-1.380	Si
CC Fuerte	C+	0.8750	0.1250	-3.510	-3.114	-2.823	-2.606	Si
CC Fuerte (+)	C++	0.9375	0.0625	-9.060	-7.639	-6.689	-6.013	Si
Conjunción	C	1.0000	0.0000	$-\infty$	$-\infty$	$-\infty$	$-\infty$	Si

7.4.2 Ambiente integrado WebQEM_Tool para el soporte de cálculos y documentación

En cuanto al ambiente para el soporte de los cálculos y documentación hemos pensado en dos tipos de herramientas. La primera, es un hiperdocumento que integra el árbol de requerimientos, las plantillas presentadas en el capítulo 6, y la tabla de resultados parciales y globales en una planilla de cálculo hiperenlazada (Excel del paquete Office97™). La funcionalidad de cálculo, particularmente los procedimientos para el cálculo de las preferencias parciales y globales fueron realizados en VBasic.

La ventaja es la disponibilidad de este paquete y el navegador IE en PCs y computadoras portátiles, y la flexibilidad para generar e integrar informes textuales,

tabulares y gráficos. Además, la hoja de cálculo contiene funciones preestablecidas de máximos, mínimos, etc. que facilitan la determinación y ordenamiento del ranquin y el análisis. No obstante, una desventaja es que no editamos las funciones de agregación visualmente para luego calcular la estructura de un modo automático, sino que se van ingresando los valores de las celdas a la función embebida por medio de selección de las mismas, e ingresando el tipo de operador lógico. El resultado del cálculo parcial es automático, dejando el valor en las celdas marcadas.

Una herramienta más potente es el empleo de las facilidades de programación Java (del paquete JDK 2.0™). Actualmente estamos desarrollando un ambiente integrado, denominado WebQEM_Tool, para el soporte de cálculos y documentación. Un aspecto interesante, es la edición visual de las funciones de agregación (como las mostradas en la figura 7.7). Esto permite agregar o modificar y reusar componentes, y editar, ingresar y guardar los parámetros de entrada en cualquier momento. Los valores parciales y finales se calculan y recalculan automáticamente, cuando los evaluadores ejecutan el comando correspondiente.

Además, tenemos planificado integrar el ambiente para edición, cálculo y documentación WebQEM_Tool, con la herramienta para automatización de métricas en la Web, denominado Web-siteMA.

Capítulo 8

Fase de Análisis de Resultados, Recomendaciones y Documentación

En la fase de *Análisis de Resultados, Recomendaciones y Documentación* los evaluadores realizan actividades de análisis y comparación de las preferencias de calidad elementales, parciales y globales, y, asimismo, la justificación de los resultados. A partir de las metas establecidas y el punto de vista de usuario a evaluar, el proceso culmina con las conclusiones y recomendaciones del caso. Por otra parte, se utilizan herramientas y mecanismos de análisis y documentación para facilitar la interpretación de los datos, su seguimiento y registración.

Una de las actividades de relevancia de Web-site QEM es la de “*Análisis y evaluación de los resultados parciales y globales*”. Los datos numéricos e información volcada en tablas, gráficos y otros documentos (e hiperdocumentos), pueden ser eficientemente empleados en actividades de análisis y toma de decisión. En este capítulo, analizaremos los resultados del caso de estudio de sitios académicos, y, en el capítulo 9, presentaremos los aspectos salientes del caso de estudio de museos en la Web.

8.1 Empleo de Tablas, Gráficos, y Herramientas

En la tabla 8.1, se muestran los resultados de los valores de las preferencias para las características de calidad de más alto nivel, y los valores finales para los seis sitios de las universidades evaluadas (extractada de la tabla 7.1), y que nos servirá de punto de partida para el análisis y recomendaciones.

Las filas resaltadas en verde corresponden a los valores máximos de las preferencias de calidad para la característica respectiva. Es importante recordar que un valor de preferencia de calidad parcial o global puede caer en la región de aceptabilidad denominada *satisfactoria, marginal, o insatisfactoria*.

Tabla 8.1 Resultados de los valores de las preferencias para las características de más alto nivel, y valores finales para los seis sitios de las universidades evaluadas. (Las filas resaltadas en verde corresponden a los valores máximos)

Características de Alto Nivel	UPC	UChile	UTS	NUS	Stanford	UQAM
1. Usabilidad	76.18	51.01	80.08	57.71	71.93	60.94
2. Funcionalidad	61.84	50.39	48.89	38.99	77.29	71.12
3. Confiabilidad	60.40	87.62	90.86	88.70	83.05	63.61
4. Eficiencia	69.09	52.03	76.11	53.12	85.99	69.47
Calidad Global	66.91	56.55	69.61	54.46	79.76	66.05

Por lo tanto, un nivel de aceptabilidad *satisfactorio* (también representado por una barra de calidad verde como en la figura 8.1, o los números en verde como en la tabla 8.1), indica que los valores están en el rango de 60 a 100 %. Un nivel de aceptabilidad *marginal* indica que los valores de la preferencia caen en el rango de 40 a 60 % (barra y números en gris). Finalmente, un valor de preferencia es *insatisfactorio* si cae dentro de la barra roja (es decir, cumple entre el 0 % y el 40 % de los requerimientos de calidad acordados). En la figura 8.1, se puede apreciar el ranquin final para los seis sitios académicos intervinientes en el proceso de evaluación.

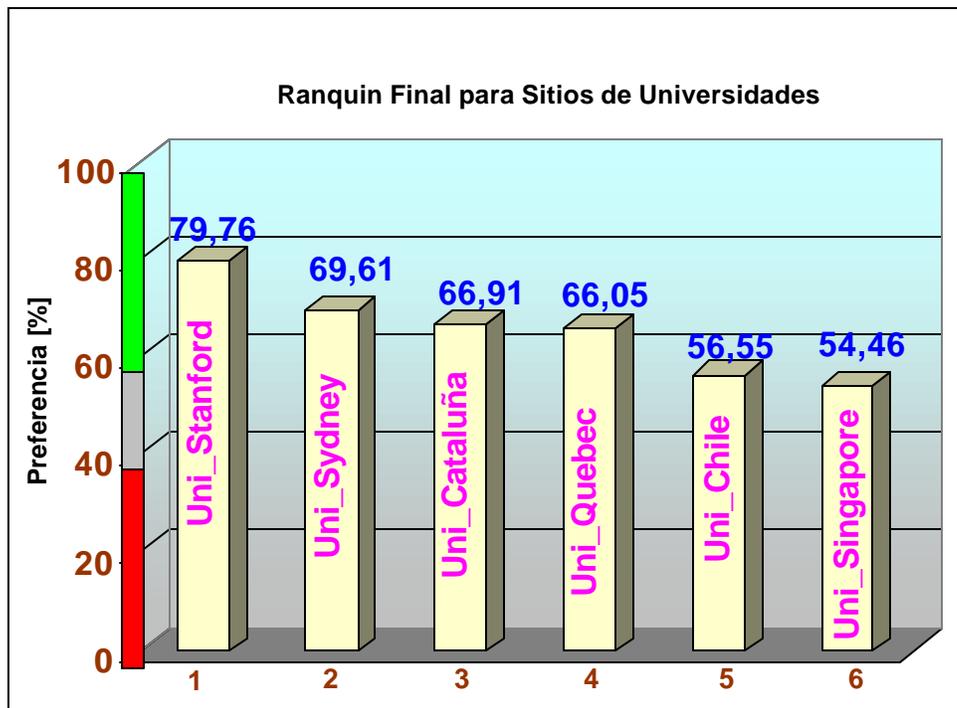


Figura 8.1 Ranquin final para los seis sitios académicos en la Web.

Para la actividad de análisis, evaluación y justificación de resultados parciales y finales el empleo de tablas, gráficos y otros mecanismos de representación, son de fundamental importancia para los evaluadores como fuentes de datos e información: por ejemplo, las tablas 8.1, 8.2, 7.1 y 6.3; los gráficos como los mostrados en las figuras 8.1, 8.2, 8.3 y 9.10; los diagramas o modelos de agregación representados en las figuras 7.7 y 7.8, y asimismo, las plantillas de características y atributos como las discutidas en la sección 6.4.3.

A seguir, utilizaremos estas fuentes de información para analizar y sacar conclusiones sobre la calidad de los artefactos bajo estudio.

8.2 Análisis y Comparación de Resultados Parciales y Globales

La preferencia de calidad global para cada sitio académico se puede observar al final de la tabla 8.1 y un ranquin de los mismos en la figura 8.1. El proceso de evaluación ha

ranqueado primero al sitio de la Universidad de Stanford con un 79.76% de la preferencia de calidad global (dentro de la barra de calidad verde), luego al sitio de la Universidad Tecnológica de Sydney (UTS) con el 69.61%; tercero, al sitio de la Universidad Politécnica de Cataluña (UPC) con el 66.91%; cuarto, casi con igual puntaje, al sitio de la Universidad de Quebec en Montreal (UQAM) con el 66.05%; quinto, al sitio de la Universidad de Chile (UChile) con el 56.55% (cayendo en la barra de calidad gris), y, finalmente, sexto en el ranquin, el sitio de la Universidad Nacional de Singapur (NUS) con el 54.46% de la preferencia global.

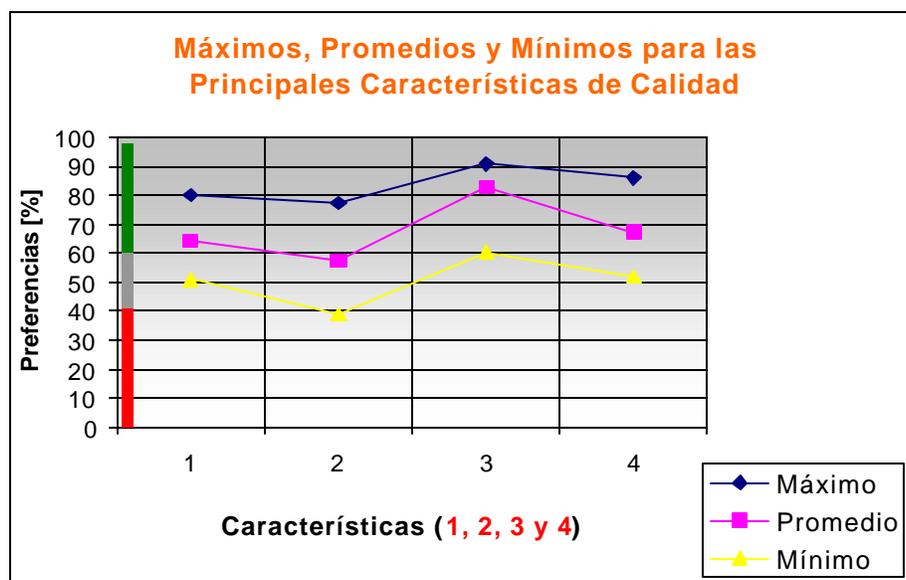


Figura 8.2 Valores máximos, promedio y mínimo para las características Usabilidad (codificada 1 en el árbol de requerimientos); Funcionalidad (2); Confiabilidad (3), y Eficiencia de los seis sitios académicos en la Web.

La tabla 8.2 muestra los valores máximos, mínimos y promedio para las preferencias de calidad de las características de más alto nivel como Usabilidad, Funcionalidad, Confiabilidad, y Eficiencia de los seis sitios académicos en la Web. (Además, la figura 8.2 muestra el gráfico de líneas respectivo).

Tabla 8.2 Valores máximos, mínimos, promedio y desvío estándar para las preferencias de calidad para las características de más alto nivel, y para la preferencia global.

Características	Máximo	Promedio	Mínimo	Desvío
1. Usabilidad	80,08	64,33	51,01	11,45
2. Funcionalidad	77,29	57,34	38,99	14,57
3. Confiabilidad	90,86	82,77	60,40	13,48
4. Eficiencia	85,99	67,34	52,03	13,18
Calidad	79,76	65,29	54,46	9,21

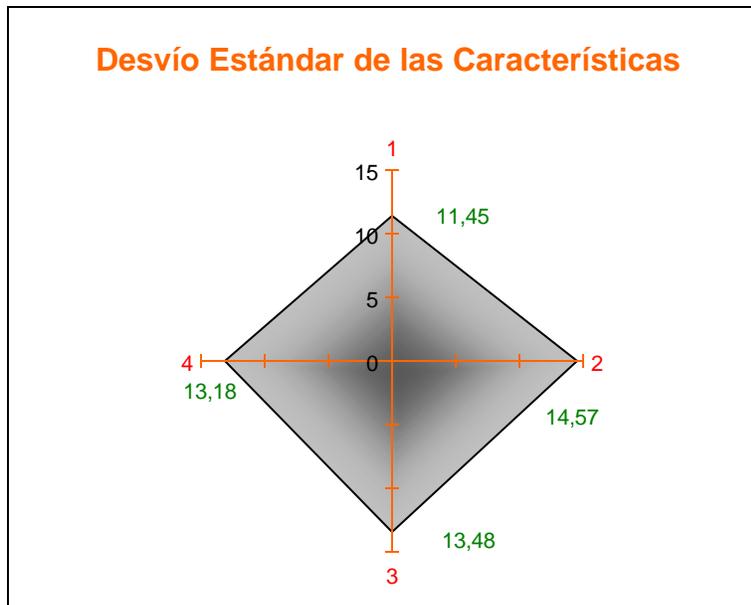


Figura 8.3 *Desvío Estándar de las características Usabilidad; Funcionalidad; Confiabilidad y Eficiencia para los seis sitios.*

8.2.1 Análisis para Usabilidad

Con respecto a la característica *Usabilidad*, el puntaje más alto fue para UTS (obteniendo el 80.08%), y el más bajo fue para UChile (51.01%). Esto es debido en parte a que el sitio de la UTS ha alcanzado un 98.17% de la preferencia en la subcaracterística 1.1 (*Comprensibilidad Global del Sitio*) en tanto que el sitio de la UChile ha alcanzado un 43.50% de la preferencia en dicha subcaracterística (ver tabla 7.1). Por ejemplo, se puede observar que tanto UChile como UQAM no tienen resuelta la subcaracterística 1.1.1, es decir es totalmente inaceptable el *Esquema de Organización Global* implementado. Estos sitios, no tienen disponibles atributos como *Mapa del Sitio* (ver 1.1.1.1 en la tabla 6.3), o *Tabla de Contenidos* (1.1.1.2), o *Índice Global* (1.1.1.3). Una simple recomendación, es que debieran implementar algunos de estos atributos, principalmente, uno de los de mayor importancia relativa como *Tabla de Contenidos*, o *Índice Global* de algún tipo (ver los pesos en el esquema de la figura 7.7a). Por otra parte, el puntaje del 98.17% de UTS para la subcaracterística 1.1 se explica por la alta preferencia para atributos como 1.1.2, 1.1.3 y 1.1.4 (según comentamos en la sección 6.7, y como además se puede observar en la tabla 6.3)

Además, el más bajo puntaje para UChile respecto de *Usabilidad* (51.01%) también se explica en parte por la marca de 44.6% para *Mecanismos de Ayuda* y *Retroalimentación en línea*. Por ejemplo, dicho sitio carecía al momento de la

evaluación del *Indicador de Última Actualización* tanto global¹ como por subsitio; es decir, esta subcaracterística (1.2.2) permite conocer al visitante si se ha realizado alguna modificación o agregado en el sitio Web o en alguna página del subsitio (la que está visitando).

Sin embargo, considerado la subcaracterística *Aspectos de Interfaces y Estéticos* UChile se posiciona mejor que el sitio de NUS que cumplimentó el 34,9% de los requerimientos establecidos. Claramente, la tabla 7.1 muestra que este sitio tiene serios problemas con la *Permanencia y Estabilidad en la Presentación de los Controles Principales*, es decir, por lo general, no permanecen los controles para acceder a las opciones del menú principal, ni directa o indirectamente. Asimismo, tiene problemas de estilo y de estética (analizando cada atributo que compone estas subcaracterísticas, se puede formular recomendaciones precisas para una mejora).

Por último, y con respecto a la característica de *Usabilidad*, vemos que el promedio de los seis sitios académicos es del 64,33% con el menor desvío estándar dentro de las características restantes de más alto nivel, como se aprecia en la figura 8.3. Además, como veremos para el caso de estudio de museos en la Web (tabla 9.4 y figura 9.10), el promedio de cuatro sitios arrojó el 58,40 % para *Usabilidad* con un desvío estándar de 10,14 que es ligeramente menor que para sitios universitarios.

8.2.2 Análisis para *Funcionalidad*

Con respecto a la característica de alto nivel denominada *Funcionalidad*, el puntaje más alto fue para Stanford (77,29%), y el más bajo, para NUS (38,99%) cayendo en el nivel insatisfactorio de aceptabilidad. Si se analiza la fila correspondiente de la tabla 8.1, se observa que los valores caen en los tres niveles de aceptabilidad. Además, en promedio, esta característica ha obtenido el valor más bajo (ver tabla y figura 8.2) y con la mayor dispersión² de puntos con respecto a la media (figura 8.3)

Como comentamos en capítulos previos, la característica *Funcionalidad* se descompone en tres subcaracterísticas; dos de estas, bastantes independientes del dominio como son *Aspectos de Búsqueda y Recuperación*, y *Aspectos de Navegación y Exploración*; y una tercera, más dependiente, denominada *Aspectos del Dominio orientados al Estudiante* (ver el árbol de requerimientos de la figura 5.7)

El mecanismo de búsqueda y recuperación de información representa una funcionalidad imprescindible en consideración del tamaño de los sitios evaluados. En la tabla 6.3 se encuentran todos los indicadores elementales para 2.1 (*Aspectos de Búsqueda y*

¹ Para ser precisos, la UChile poseía en su página principal una fecha, que en vez de ser la de última actualización, era la fecha del día que se actualizaba automáticamente con el cambio de fecha (de modo que al atributo lo computamos como no disponible).

² Una menor dispersión de los puntos respecto de la media se puede interpretar como que la característica es más estable y viceversa.

Recuperación) y en la tabla 7.1 las marcas de preferencia parcial. El sitio de Stanford ha alcanzado la marca sobresaliente de 100% de la preferencia de calidad para la subcaracterística 2.1 debido a que tiene una excelente implementación de búsqueda global y restringida, y excelentes niveles de personalización de la recuperación de las ocurrencias (ver figura 6.10). No obstante, las marcas para NUS y UTS caen en la barra roja para 2.1.

A seguir, concentraremos el análisis en esta subcaracterística. La figura 8.4 muestra el gráfico de líneas de los valores de preferencia computados para la subcaracterística 2.1, y la tabla 8.3 los valores máximos, mínimos, promedio y desvío estándar.

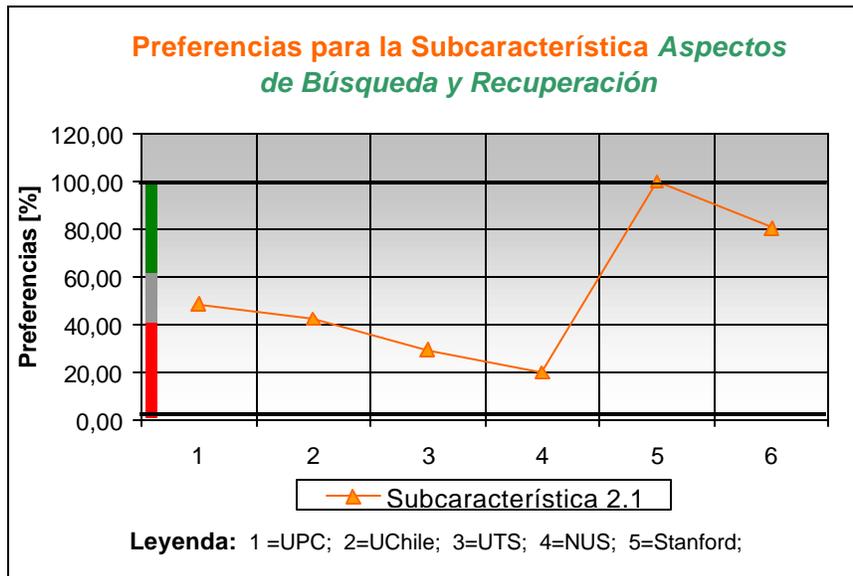


Figura 8.4 Valores de las preferencias de calidad de Aspectos de Búsqueda y Recuperación para los seis sitios (tomados de la tabla 7.1)

Tabla 8.3 Valores máximos, mínimos, promedio y desvío estándar para las preferencias de calidad de la subcaracterística Aspectos de Búsqueda y Recuperación

Subcaracterística	Máximo	Promedio	Mínimo	Desvío Estándar
2.1 Aspectos de Búsqueda y Recuperación	100	54,39	19,73	30,88

Si se analizan los resultados para los mecanismos de búsqueda y recuperación de los seis sitios típicos, fácilmente se puede concluir que mayoritariamente presentan niveles de aceptabilidad insatisfactorios y marginales (cuatro de seis sitios). Además, es importante resaltar que para la subcaracterística 2.1, en consideración de la tabla 8.3, se observa el mayor desvío estándar de todas las subcaracterísticas de segundo nivel presentadas en la tabla 7.1.

Sin embargo, para producir una recomendación, es importante profundizar el análisis para ver de dónde provienen los resultados, para que, en definitiva, podamos valorar si

el problema es mayor en la búsqueda que en la recuperación de información, o en ambas subcaracterísticas. En la figura 8.5 mostramos el esquema de subcaracterísticas y atributos para 2.1. Podemos ver, al tercer nivel del árbol dos subcaracterísticas 2.1.1 (*Mecanismo de Búsqueda en el Sitio Web*) y 2.1.2 (*Mecanismos de Recuperación*).

2. Funcionalidad

2.1 Aspectos de Búsqueda y Recuperación

2.1.1 Mecanismo de Búsqueda en el Sitio Web

2.1.1.1 Búsqueda Restringida

2.1.1.1.1 de Personas

2.1.1.1.2 de Cursos

2.1.1.1.3 de Unidades Académicas

2.1.1.2 Búsqueda Global

2.1.2 Mecanismos de Recuperación

2.1.2.1 Nivel de Personalización

2.1.2.2 Nivel de Retroalimentación en la Recuperación

2.2 Aspectos de Navegación y Exploración

Figura 8.5 Atributos de la subcaracterística 2.1 (tomados de la figura 5.7)

Ambas subcaracterísticas son agregadas mediante un operador lógico de simultaneidad C-- (no mandatorio), y con una relativa importancia de 0.7 para el *Mecanismo de Búsqueda* y de 0.3 para el *Mecanismo de Recuperación* (ver fig 7.7 b). Con respecto a 2.1.1, hemos agrupado a la *Búsqueda Restringida* (2.1.1.1) y a la *Búsqueda Global* (2.1.1.2) con un criterio de intensidad media de reemplazabilidad; esto es, la presencia de un tipo de búsqueda puede reemplazar en parte la ausencia de la otra.

Esto último se observa en la tabla 8.4, por ejemplo para el sitio de NUS, en donde no tiene búsqueda global, pero sí algún tipo de búsqueda restringida, obteniendo un **35.12%** para 2.1.1. Pero el hecho de no tener implementado ningún mecanismo de recuperación personalizable por el usuario, ni retroalimentación en la recuperación (como por ejemplo, qué opciones da el sitio al visitante cuando no obtuvo ninguna ocurrencia), resulta en definitiva en la más baja preferencia parcial para 2.1.

Tabla 8.4 Valores de las preferencias para la subcaracterística 2.1, de los seis sitios.

Subcaracterística 2.1	Pesos	UPC	UChile	UTS	NUS	Stanford	UQAM
2.1 Aspectos de Búsqueda y Recuperación	0.3	48.59	42.37	29.32	19.73	100	80.55
2.1.1 Mecanismo de Búsqueda en el Sitio	0.7	48	50.84	52.17	35.12	100	82.98
2.1.1.1 Búsqueda Restringida	0.6	24	40	44	40	100	60
2.1.1.2 Búsqueda Global	0.4	60	60	60	0	100	100
2.1.2 Mecanismos de Recuperación	0.3	50	25	0	0	100	75

La tabla muestra claramente para la subcaracterística 2.1 que, excepto para los dos sitios en el rango verde, los cuatro sitios restantes tienen problemas tanto en los mecanismos de búsqueda como en los de recuperación de información. Por ejemplo, la UPC poseía búsqueda global básica (no avanzada, dando una preferencia elemental del 60%), y dentro de las tres búsquedas restringidas, sólo disponía de la búsqueda básica de

personas por apellido. Estos resultados parciales agregados dio como resultado una preferencia parcial de 48 % para 2.1.1. Una sugerencia de mejora sería que, por un lado se rediseñe el mecanismo de búsqueda global, con aspectos de búsquedas avanzadas (como filtros, etc.), y que por otro lado, se incorpore búsquedas de cursos y unidades académicas. Un modelo a seguir (u observar) podría ser el mecanismo de búsquedas globales y restringidas de la Universidad de Stanford. En esta dirección se pueden formular varias recomendaciones y sugerencias para mejorar.

Continuando con el análisis de los datos para *Funcionalidad*, y en consideración de la subcaracterística 2.2 (*Aspectos de Navegación y Exploración*), una vez más, el sitio de la Universidad Nacional de Singapur es el más débil. Por ejemplo, asombrosamente el sitio no posee controles de navegación dentro de un subsitio, ni tampoco controles para retornar a la página principal. Por lo tanto, la preferencia fue del 0% para la subcaracterística 2.2.2.1.

Sin embargo, todos los sitios (excepto Stanford) arrojaron una preferencia muy baja o insatisfactoria en la subcaracterística contenida en *Navegabilidad*, denominada *Orientación* (2.2.1.1). Particularmente, los atributos componentes de *Orientación* son el *Indicador del Camino*, y la *Etiqueta de la Posición Actual*. Respecto al atributo *Indicador del Camino* (2.2.1.1.1) significa que los usuarios de la Web, al navegar, deben tener pistas visuales (con elementos de diseño consistentes), que les indique con precisión en dónde se encuentran posicionados, dentro de la estructura del espacio de información del sitio. Este atributo trata con la orientación del usuario en tanto navegan el espacio de información de un sitio Web. Cada página del sitio debiera contener el indicador del camino (ver la sección 9.2.2 y figura 9.4).

Por último, considerando la subcaracterística 2.3 (*Aspectos del Dominio orientados al Estudiante*), tan importante en cuanto al contenido y servicios ofrecidos para el perfil evaluado, el rango de preferencias parciales va entre 57.43% y 82.37% lo que son puntajes en general satisfactorios; la mejor marca fue para el sitio de UQAM (ver tabla 7.1).

8.2.3 Análisis para *Confiabilidad* y *Eficiencia*

Con respecto a la característica de *Confiabilidad*, el valor de preferencia más alto fue para UTS (90.86%), y el más bajo fue para la UPC (60.40%) arrojando valores de preferencia dentro del nivel satisfactorio de aceptabilidad, como se aprecia en la tabla 8.1. En cuanto a la característica 4 (*Eficiencia*), la preferencia más alta fue para Stanford (85.99%), y el valor más bajo fue para el sitio de la UChile (52.03%). En la tabla 8.2 y en las figuras 8.2 y 8.3 se muestran los promedios y desvíos estándar.

Para la característica *Confiabilidad*, el lector puede apreciar en la tabla 7.1, los resultados de las preferencias parciales para las subcaracterísticas 3.1.1 (*Errores de*

Enlaces) y 3.1.2 (*Errores o Deficiencias Varias*). Los indicadores de calidad elementales para 3.1.1 provienen de la agregación de los atributos *Enlaces Rotos* (3.1.1.1), *Enlaces Inválidos* (3.1.1.2) y *Enlaces no Implementados* (3.1.1.3). Analizando los valores elementales se observa que la mayor fuente de errores de enlaces proviene de *Enlaces Rotos*. Evidentemente, una recomendación es que se debe correr con mayor frecuencia programas analizadores de enlaces en la medida en que el sitio, ya en la fase operativa, evoluciona y se mantiene.

La tabla 7.1 indica para la subcaracterística *Errores de Enlaces*, que el sitio de la UPC y de la UQAM, deberían mejorarse debido a que el resultado de la evaluación ha arrojado valores dentro del nivel de aceptabilidad marginal (el 50% respectivamente). Los restantes cuatro sitios respecto de esta subcaracterística (3.1.1) están en el rango del 76 al 87%.

La no-deficiencia en cuanto a enlaces, y la dificultad en lograrla, plantea un problema de mantenibilidad y control de cambios de sitios Web. Por ejemplo, el sitio de la Universidad de Stanford (en donde no se analizó cada Facultad y Escuela), tenía al momento de la evaluación, 18.872 páginas con una cantidad de 244.168 enlaces.

Por último, la característica de alto nivel denominada *Eficiencia* está compuesta por dos subcaracterísticas de segundo nivel denominadas *Performancia* (4.1) y *Accesibilidad* (4.2). En cuanto a *Performancia* el sitio de Stanford ha alcanzado la marca completamente satisfactoria de 100% (como se vio en la sección 6.4.3); sin embargo, los sitios de NUS y UChile arrojaron valores marginales (51.46 y 50.46% respectivamente). En estos casos una recomendación sería planear alguna mejora en el tamaño de los componentes gráficos y/o en la cantidad de objetos en cada página de manera que la mayoría de las mismas no supere el valor umbral de tamaño máximo de página aceptable, previendo líneas de comunicación de baja velocidad (como las de 14.400 bps).

Finalmente, con respecto al subcaracterística *Accesibilidad*, la preferencia más baja fue para el sitio de UQAM (40.84%) y el más alto para el sitio de UTS (63.04%). Desafortunadamente, ningún sitio dispone del atributo de la versión sólo texto (4.2.1.1). Esto sería útil porque los usuarios tendrían accesibilidad total de información en las páginas, principalmente para las personas con discapacidades, o cuando la velocidad es un problema [W3C 99].

Sin embargo, si se observa el esquema de agregación de preferencias de la figura 7.7 d, el operador lógico A tiene dos entradas independientes la 4.2.1.1 (con peso 0.3) y la 4.2.1.2 (con peso 0.7). Esta última entrada y su relativa importancia, refleja que para el usuario estudiante, en general, es más importante contar con una buena legibilidad de las páginas cuando se desactiva la característica de “ver imagen” del navegador. No obstante, esta presencia no garantiza la calidad del texto alternativo. Una tendencia que

se observó es que los desarrolladores no siempre han modificado el texto alternativo de las imágenes generado automáticamente con algún editor. De esta forma, si bien el texto alternativo está presente en el objeto, su significado en muchos casos es irrelevante (no garantiza la calidad del mismo). Por esta razón, se evaluó mediante un criterio de preferencia directa, a la *Legibilidad Global* (atributo 4.2.1.2.2) según lo especificamos en la plantilla respectiva de la sección 6.4.3.

8.3 Consideraciones finales sobre el caso de estudio

Como hemos mencionado previamente (en la sección 4.1.2), esta investigación tiene por objetivo principal desarrollar y mostrar la estrategia propuesta, denominada Metodología de Evaluación de Calidad de Sitios Web (Web-site QEM), la que pretende realizar un aporte innovador e ingenieril al proponer un enfoque sistemático, disciplinado y cuantitativo que se adecue a la evaluación, comparación y análisis de calidad de aplicaciones centrados en la Web más o menos complejas.

El proceso de evaluación genera preferencias elementales, parciales, y globales que pueden ser fácilmente analizables, justificables, con seguimiento hacia atrás y hacia adelante, y pueden emplearse eficazmente en actividades de toma de decisión.

Por otra parte, para mostrar la aplicabilidad de la metodología, hemos realizado el caso de estudio de sitios académicos, el de museos y un survey sobre este último dominio (Apéndice A.3).

En la sección 4.2.1 formulamos la hipótesis que declara: *“que en sitios Web típicos, de un dominio determinado, la calidad de los artefactos satisface en general los requerimientos de calidad en consideración de un perfil de usuario”*. Luego, de un modo particular, declaramos: *“cada sitio Web satisface al menos el punto crítico de aceptabilidad del 60% de la preferencia global, conforme a los requerimientos de calidad acordados y en consideración de un perfil de usuario”*.

En este contexto, nosotros observamos que el estado del arte en la calidad de los sitios académicos típicos, desde el punto de vista del estudiante, es en la mayoría de los casos satisfactorio (en conformidad con los requerimientos). Sin embargo, como hemos podido corroborar, la lista de atributos pobremente diseñados o atributos ausentes no está vacía. Los sitios evaluados obtuvieron una preferencia global que está en el rango que va desde el 54% a 80% de los requisitos especificados (ver tabla y figura 8.1), y en promedio, aproximadamente el 65%. Por lo tanto, la hipótesis nula anterior no se confirma estrictamente conforme a la preferencia global obtenida para cada sitio evaluado. Una parte del análisis de las causas fueron expresadas en la sección 8.2.

Por último, es importante indicar que el ranquin final refleja estos requisitos específicos para esta audiencia específica y, por lo tanto, no debe considerarse obviamente como un puntaje general del sitio para todas las audiencias. De todas maneras, esto no invalida los mecanismos y componentes objetivos utilizados de la metodología Web-site QEM en el proceso de evaluación si no que, en todo caso, se debiera hacer la evaluación más extensiva de manera de abarcar todas las audiencias previstas para el sitio.

Capítulo 9

Ejemplos: Empleo de Web-site QEM.

9.1 Estudios de Campo Realizados

Usamos a la metodología Web-site QEM en diferentes fases del proceso de desarrollo, sin embargo, los estudios más amplios que nos permitieron validar empíricamente sus principales actividades, métodos, modelos y herramientas, fueron realizados en la fase operativa del ciclo de vida de los sistemas de información centrados en la Web.

En este contexto, realizamos dos casos de estudio sobre sitios típicos en el dominio de museos y en el de sistemas de información académicos en la Web. En la sección 4.2.1.2 introdujimos el estudio para el dominio de museos y realizamos algunas consideraciones sobre la edad, el carácter evolutivo de los sitios, entre otros asuntos. Además, en la sección 4.2.1.3 presentamos el caso de estudio para el dominio de sitios académicos y que fuimos desarrollando a lo largo de los Capítulos 5, 6, 7, 8 y Apéndice A.1 de esta tesis.

Por otra parte, hemos realizado un survey en el dominio de aplicaciones de museos en la Web. El estudio se realizó sobre 29 atributos directamente mensurables derivados de algunas de las características como usabilidad, funcionalidad y eficiencia, sobre una muestra de 24 museos internacionales. Si bien este estudio no fue central en el empleo de Web-site QEM (dado que solamente se utilizaron unas pocas actividades de la metodología, como definición de características y atributos y criterios de evaluación elemental), sin embargo, el mismo arrojó interesantes resultados para el dominio. El survey se describe en el Apéndice A.3 de un modo informativo.

Finalmente, hemos realizado un proyecto de desarrollo de sitio Web bajo contrato con un ente público (el Ministerio de la Producción de la Provincia de La Pampa – Decreto Prov. 1942/98), cubriendo las fases de exploración y desarrollo. En este proyecto especificamos los requerimientos de calidad en etapas tempranas y evaluamos la calidad para el perfil de visitante general -la principal audiencia-, al momento de su distribución. Es decir, empleamos la metodología Web-site QEM como herramienta de soporte al proceso de aseguramiento y control de calidad. (El sitio se encuentra temporariamente en la dirección <http://zflp.ing.unlpam.edu.ar>).

En este capítulo nos concentraremos principalmente en los aspectos salientes del caso de estudio de museos (sección 9.2). En la sección 9.3 comentaremos algunos detalles del estudio que hemos comenzado sobre sitios en el dominio de comercio electrónico.

El empleo de Web-site QEM como herramienta para la evaluación y aseguramiento de calidad de artefactos ha sido una tarea compensadora, dado que nos ha ayudado a reconocer patrones de diseño, en varios casos, independientes del dominio, como *Esquema de Organización Global del Sitio*, *Indicador de Última Actualización*, *Retroalimentación Personalizada*, *Estrategias de Búsqueda*, por citar algunos. Vale destacar que recientemente han sido documentados patrones de diseño para la Web como los catalogados en [Rossi et al 99].

9.2 Caso de Estudio sobre Museos Típicos en la Web

Este caso de estudio fue el primero realizado, iniciando la planificación y preparación del mismo en agosto del 98. Remitimos al lector a la sección 4.2.1.2 en donde introducimos y realizamos algunas consideraciones sobre el mismo para el dominio de museos en la Web [Olsina 99]. En dicha sección, en la figura 4.2, se muestran las páginas principales de los cuatro museos evaluados.

9.2.1 Arbol de Características y Atributos

En cuanto al enfoque de modelo de calidad empleado sigue el mismo modelo de calidad mixto, que el discutido en la subsección 5.4.3 y al especificado para el caso de estudio de sitios académicos. Se parte de un conjunto prescripto de características y se define el mecanismo de descomposición restante, se acuerdan las subcaracterísticas de niveles inferiores, los atributos, (los criterios de medición) y las relaciones entre atributos, subcaracterísticas y características.

Por otra parte, el perfil de usuario seleccionado para este estudio fue el del visitante, particularmente el del visitante general (esto es, visitantes casuales o intencionales). Según estudios realizados sobre el dominio de museos [Furano et al 97, Garzotto et al 97] los visitantes generales pueden ser descompuestos en dos audiencias más específicas: visitantes casuales y visitantes intencionales.

Garzotto et al., afirma que un *visitante casual* se define como a la audiencia que ingresa al sitio por azar, y permanece probablemente un corto lapso de tiempo, en donde la aplicación está pensada para ser usada probablemente una sola vez. Esta audiencia puede diferir en los niveles de experticia en el uso de tecnologías Web. Por otra parte, un *visitante intencional* se define como a la audiencia que tiene al menos algún conocimiento o manifiesta algún interés en el dominio de museos, y desea probablemente informarse o aprender más acerca del mismo. Su permanencia en el sitio es generalmente mayor que la de una audiencia casual.

- 1. Usabilidad**
 - 1.1 Comprensibilidad Global del Sitio**
 - 1.1.1 Esquema de Organización Global
 - 1.1.1.1 *Mapa del Sitio*
 - 1.1.1.2 *Índice Global (por Temas, etc.)*
 - 1.1.1.3 *Tabla de Contenidos*
 - 1.1.2 *Calidad en el Sistema de Etiquetado*
 - 1.1.2.1 *Etiquetado Textual*
 - 1.1.2.2 *Etiquetado con Iconos*
 - 1.1.3 *Visita Guiada*
 - 1.1.3.1 *Visita Convencional*
 - 1.1.3.2 *Visita Virtual (con Tecnología VR)**
 - 1.1.4 *Mapa de Imagen (de Pisos y Salas de Exhibición)*
 - 1.2 Mecanismos de Ayuda y Retroalimentación en línea**
 - 1.2.1 *Calidad de la Ayuda*
 - 1.2.1.1 *Ayuda Explicatoria acerca del sitio*
 - 1.2.1.2 *Ayuda de la Búsqueda*
 - 1.2.2 *Indicador de Última Actualización*
 - 1.2.2.1 *Global (de todo el sitio Web)*
 - 1.2.2.2 *Restringido subsitio o página)*
 - 1.2.3 *Directorio de Direcciones*
 - 1.2.3.1 *Directorio E-mail*
 - 1.2.3.2 *Directorio TE-Fax*
 - 1.2.3.3 *Directorio Correo Postal*
 - 1.2.4 *Facilidad FAQ*
 - 1.2.5 *Cuestionario/Survey*
 - 1.3 Aspectos de Interfaces y Estéticos**
 - 1.3.1 *Cohesividad al Agrupar los Objetos de Control Principales*
 - 1.3.2 *Permanencia y Estabilidad en la Presentación de los Controles Principales*
 - 1.3.2.1 *Permanencia de Controles Directos*
 - 1.3.2.2 *Permanencia de Controles Indirectos*
 - 1.3.2.3 *Estabilidad*
 - 1.3.3 *Preferencia Estética*
 - 1.3.4 *Uniformidad en el Estilo del sitio*
 - 1.4 Misceláneas**
 - 1.4.1 *Soporte a Lenguaje Extranjero*
 - 1.4.2 *Característica de Download*
 - 2. Funcionalidad**
 - 2.1 Aspectos de Búsqueda**
 - 2.1.1 *Mecanismo de Búsqueda en el Sitio*
 - 2.1.1.1 *Búsqueda Restringida (Colecciones)*
 - 2.1.1.2 *Búsqueda Global*
 - 2.2 Aspectos de Navegación y Exploración**
 - 2.2.1 *Navegabilidad Local (de subsitio)*
 - 2.2.1.1 *Nivel de Interconexión (para el subsitio Colecciones)*
 - 2.2.1.2 *Orientación*
 - 2.2.1.2.1 *Indicador del Camino*
 - 2.2.1.2.2 *Etiqueta de la Posición Actual*
 - 2.2.2 *Navegabilidad Global*
 - 2.2.2.1 *Acoplamiento entre Subsitios*
 - 2.2.3 *Objetos de Control Navegacional*
 - 2.2.3.1 *Permanencia y Estabilidad en la Presentación de los Controles Contextuales (Subsitio)*
 - 2.2.3.1.1 *Permanencia de los Controles Contextuales*
 - 2.2.3.1.2 *Estabilidad*
 - 2.2.3.2 *Nivel de Desplazamiento*
 - 2.2.3.2.1 *Desplazamiento Vertical*
 - 2.2.3.2.2 *Desplazamiento Horizontal*
 - 2.2.4 *Predicción Navegacional*
 - 2.2.4.1 *Enlace con Título (enlace con texto explicatorio)*
 - 2.2.4.2 *Calidad de la Frase del Enlace*
 - 2.3 Funciones Misceláneas y Específicas del Dominio**
 - 2.3.1 *Relevancia de Contenido*
 - 2.3.2 *Relevancia de Enlaces*
 - 2.3.3 *Aspectos de Comercio Electrónico*
 - 2.3.3.1 *Características de Compra*
 - 2.3.3.1.1 *Carrito de Compras*
 - 2.3.3.1.2 *Catálogo de Productos*
 - 2.3.3.2 *Compra (Transacción) Segura*
 - 2.3.4 *Aspectos de las Imágenes*
 - 2.3.4.1 *Indicador del Tamaño*
 - 2.3.4.2 *Zooming*
- 3. Confiabilidad**
 - 3.1 No Deficiencia**
 - 3.1.1 *Errores de Enlaces*
 - 3.1.1.1 *Enlaces Rotos*
 - 3.1.1.2 *Enlaces Inválidos*
 - 3.1.1.3 *Enlaces no Implementados*
 - 3.1.2 *Errores o Deficiencias Varias*
 - 3.1.2.1 *Número de deficiencias o cualidades ausentes debido a diferentes navegadores (browsers)*
 - 3.1.2.2 *Número de deficiencias o resultados inesperados independientes de browsers (p.ej. errores de búsqueda imprevistos, deficiencias con marcos (frames), etc.)*
 - 3.1.2.3 *Número de nodos Web Muertos (sin enlaces de retorno)*
 - 3.1.2.4 *Número de nodos Destinos (inesperadamente) en Construcción*
- 4. Eficiencia**
 - 4.1 Accesibilidad de Información**
 - 4.1.1 *Soporte a Versión sólo Texto*
 - 4.1.2 *Legibilidad al desactivar la Propiedad Imagen del Browser*
 - 4.1.2.1 *Imagen con Título*
 - 4.1.2.2 *Legibilidad Global*
 - 4.2 Performancia**
 - 4.2.1 *Páginas Rápidas ***

Figura 9.1 *Arbol de requerimientos de calidad para el dominio de museos en consideración de un visitante general*

Si bien el árbol de requerimientos de calidad (el de la figura 9.1) no considera a un *visitante experto*, esta audiencia se define como a los usuarios especialistas en el dominio como investigadores de la historia del arte, museólogos, entre otros. Los requerimientos de información, entre otros, son orientados a contenidos más especializados y ricos. Según Furano et al. un *visitante experto* es aquél que se encuentra interesado generalmente en todo el contenido del museo (en la Web) y en todos los itinerarios provistos tanto en la página principal como en las restantes. En consecuencia, la subcaracterística *Relevancia de Contenido* debiera ser descompuesta para especificar aspectos de relevancia del texto descriptivo de las obras, la calidad de las imágenes que representan a las obras, nivel y variedad en el tipo de ordenamiento de las colecciones y otros atributos relacionados destinados a satisfacer las necesidades de este tipo de audiencia.

En la figura 9.1 se especifican las características, subcaracterísticas y atributos para un visitante general; varios de los componentes y elementos ya fueron tratados en capítulos anteriores.

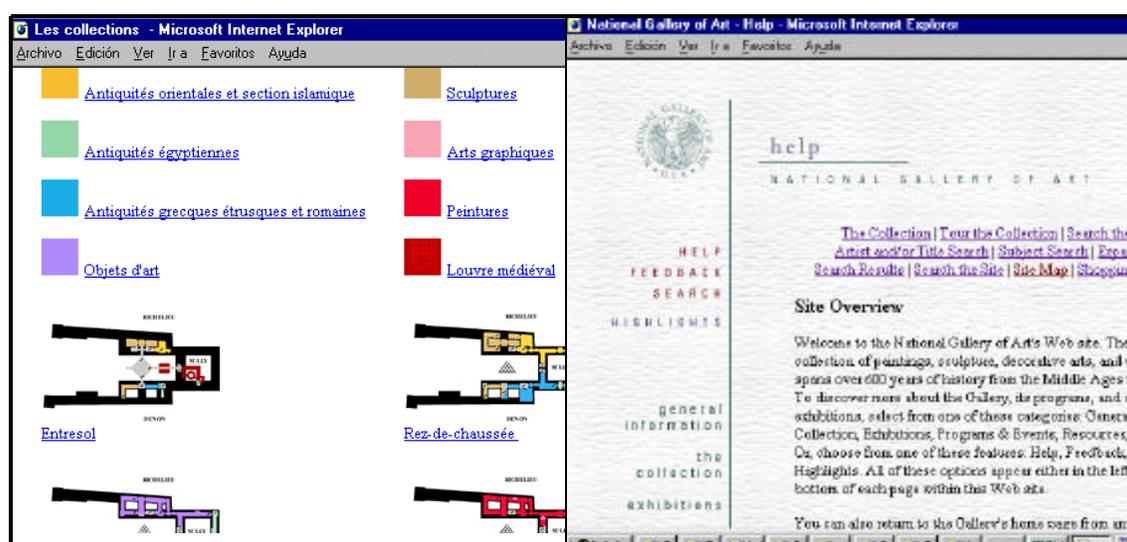


Figura 9.2 Vista parcial del mapa de imagen del museo Louvre, en la parte izquierda de la figura. A la derecha, se muestra una sección de la página de Ayuda Explicatoria acerca del Sitio del museo Galería Nacional del Arte

Como se comentó previamente (4.2.1.2), si bien se consideró dentro de los requerimientos el atributo de *Visita Guiada Virtual*¹ (usando tecnología VR), al momento de la evaluación no estaba presente en ninguno de los sitios evaluados, ya que era una tecnología emergente. Por tal motivo decidimos no incorporarlo. Sin embargo, prueba del creciente interés por implementar visitas virtuales en museos, se manifiesta en la tendencia observada en el survey que fue realizado con posterioridad al caso de

¹ La disponibilidad de este atributo no sólo posibilita mostrar las obras, sino que también se puede pasear dinámicamente por la sala en donde se exponen las obras, permitiendo no sólo apreciar a las mismas sino que también se puede apreciar la ubicación dentro de la sala o galería, entre otros detalles.

estudio (el análisis de los porcentajes de disponibilidad de ambos mecanismos se puede observar en A.3.2.1). Además, por ejemplo, el museo del Louvre introdujo a su sitio rediseñado, en noviembre de 1998, recorridos virtuales muy atractivos.

En cuanto al atributo 1.1.4, *Mapa de Imagen*, es semejante tanto en definición como en tipo de criterio elemental al discutido para sitios académicos (6.4.3). En este caso el atributo representa a un mapa de imagen (con algún modo de navegación) de pisos y salas de exhibición. En la figura 9.2 (en la parte izquierda), se muestra una interesante disposición de pisos y salas del museo Louvre, en donde los colores identifican diferentes clases de obras y exhibiciones dispuestas en los mismos. En la parte derecha de la figura se aprecia la disponibilidad en el sitio del museo NGA, del atributo 1.2.1.1 (*Ayuda Explicativa acerca del sitio*).

En la figura 9.3 (en la parte izquierda), se muestra una vista parcial de la implementación del atributo 1.2.5 (*Cuestionario/Survey*) en el museo Louvre. El cuestionario está estructurado en tres secciones (USTED –visible en la figura–, NUESTRO SERVIDOR –que es fundamental para determinar las preferencias del visitante respecto del sitio y sus sugerencias–, y PREGUNTAS TÉCNICAS –referidas a cómo se ven las imágenes en la computadora del visitante, aspectos de velocidad y tipo de navegador).

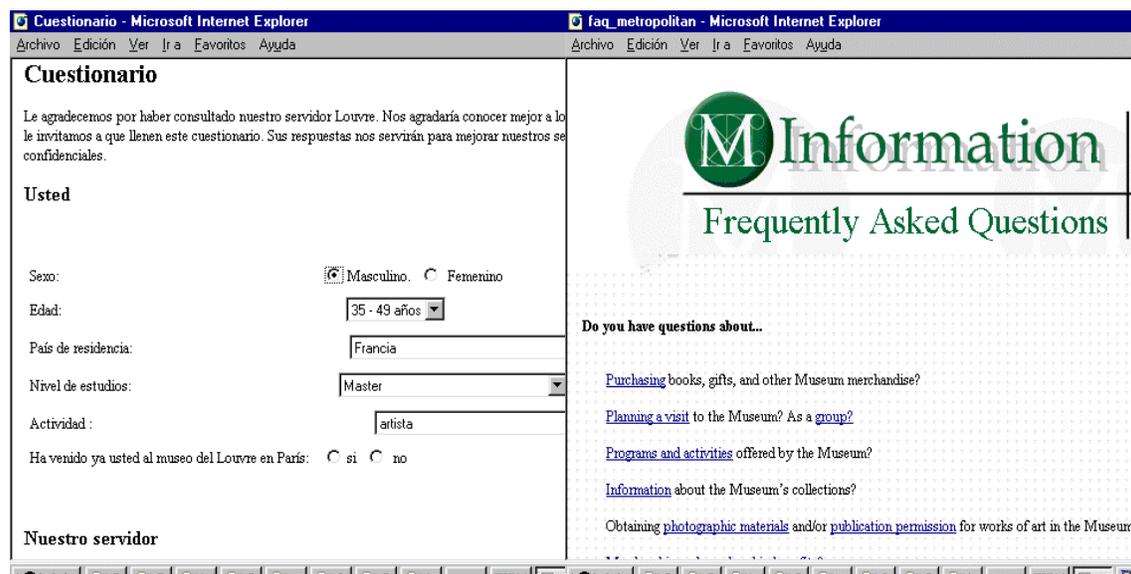


Figura 9.3 Vista parcial del cuestionario del museo Louvre, en la parte izquierda de la figura. A la derecha, se muestra una sección de la página de las Preguntas más Frecuentes (FAQ) del museo Metropolitano

En la parte derecha de la figura, se aprecia la disponibilidad en el sitio del museo Metropolitano, del atributo 1.2.5. El mecanismo FAQ en este museo, además de contar obviamente para cada pregunta con la respuesta respectiva, tiene enlazada en ésta puntos de partida para ampliar la respuesta, favoreciendo el mecanismo de aprendizaje.

Por ejemplo, la primer pregunta acerca de Compra... (“Purchasing...”) de algún libro, publicación, etc. conduce a una página catalogada por tipos de objetos, y luego a cuestiones que el usuario frecuentemente pregunta. En esta última página se permite acceder directamente, por ejemplo, al catálogo de productos y al mecanismo de compra.

A propósito del ejemplo anterior, un aspecto a destacar (aunque no pareciera propio del dominio de museos en la Web) es la funcionalidad avanzada de comercio electrónico que disponían la mayoría de los sitios de los museos evaluados (catálogos de productos, carrito de compras, transacciones seguras, etc.). El único sitio que no tenía implementado este mecanismo en línea era el museo Louvre.

En la siguiente subsección veremos algunos criterios y resultados de preferencia elementales.

9.2.2 Evaluación Elemental

A partir del árbol de calidad antes esquematizado, y para cada atributo cuantificable A_i debemos asociar y determinar la variable X_i , que tomará un valor real a partir de un proceso de medición. Además, para el rango de valores acordados para la variable X_i , por medio de un criterio elemental, se deberá hacer corresponder en una preferencia elemental IE_i (remitirse a lo discutido en 6.1 y 6.2).

A seguir, especificaremos una docena de atributos y sus criterios elementales, que no hayamos visto en 6.4.3. Principalmente, se describen atributos con distintos tipos de criterios de preferencia elemental o mecanismos de recolección de datos. (En el Apéndice A el lector puede encontrar información adicional)

Título: *Tabla de Contenido*; **Código:** (Figura 9.1) 1.1.1.3; **Tipo:** Atributo

Característica de más Alto Nivel: Usabilidad

Super-característica: Esquema de Organización Global

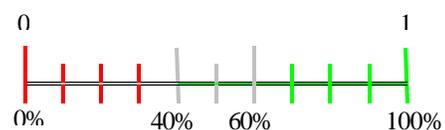
Definición / Comentarios: Es un mecanismo que permite estructurar el contenido de todo el sitio Web, permitiendo navegación directa. Está principalmente disponible en la página principal y resalta (parcialmente) la estructura jerárquica de la información de manera que los usuarios se familiaricen rápidamente de cómo el contenido está organizado en subsitios permitiendo además navegación directa a partir de sus elementos enlazados [Rosenfeld et al 98].

Tipo de Criterio Elemental: es un criterio binario, discreto y absoluto: sólo se pregunta si está disponible (1) o si no está disponible (0).

Escala de Preferencia:

Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s:



- 1) En el caso de estudio de museos, sólo el museo NGA disponía de tabla de contenidos
- 2) En el caso de estudio de sitios académicos, ejemplos de disponibilidad de tabla de contenido son los sitios NUS, UTS, Stanford, y UPC. Por otra parte, para la tabla de contenido del sitio de UTS, se ha establecido una clara organización orientada a la audiencia; esto es, para estudiantes, para personal, y para patrocinadores a la investigación.

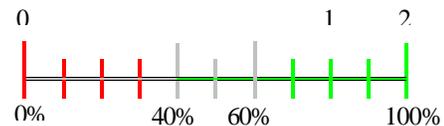
Título: *Facilidad FAQ*; **Código:** 1.2.4; **Tipo:** Atributo

Característica de más Alto Nivel: Usabilidad

Super-característica: Mecanismos de Ayuda y Retroalimentación en línea

Definición / Comentarios: Este atributo representa a un conjunto de preguntas (agrupadas y enlazadas) que se realizan con mayor frecuencia, y que están ya publicadas en el sitio con sus respectivas respuestas. A su vez, las respuestas pueden estar enlazadas a otros contenidos. Esto favorece al mecanismo de aprendizaje y/o ayuda, evitando potencialmente la demora cognitiva de los visitantes.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto, definido como subconjunto; en donde si se evalúa la disponibilidad y tipo de mecanismo, entonces: 0 = no disponible; 1 = FAQ básico, permite acceder a las respuestas, generalmente no enlazadas a otros contenidos; 2 = 1 + permite acceder a las respuestas, generalmente enlazadas a otros contenidos.



Escala de Preferencia:

Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s: Ver figura 9.3 (parte derecha), en donde el museo Metropolitano obtuvo el 100% de la preferencia. El museo NGA obtuvo el 80% de la preferencia, y los dos museos restantes no disponían de este atributo.

Título: *Indicador del Camino*; **Código:** 2.2.1.2.1; **Tipo:** Atributo

Característica de más Alto Nivel: Funcionalidad

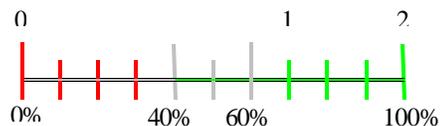
Super-característica: Orientación

Definición / Comentarios: Los usuarios de la Web, al navegar, deben tener pistas visuales (con elementos de diseño consistentes), que les indique con precisión en dónde se encuentran posicionados, dentro de la estructura del espacio de información del sitio. Este atributo trata con la orientación del usuario en tanto navegan el espacio de información de un sitio Web. Cada página del sitio debe contener el indicador del camino.

Estudios indican que el diseñador debe tener en cuenta: “*Provide visual clues with consistent design elements ...every page should have a link up to your home page as well as some indication of where they fit within the structure of your information space.*”, [IEEE WPG], sección NAVIGATION

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde si se observa la presencia del camino (en las páginas del sitio), ya sea en modo gráfico o en modo textual, entonces: 0 = no disponible; 1= parcialmente disponible (sugerido por controles sombreados o texto); 2= totalmente disponible.

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s:

- 1) Los museos del Prado, Met y NGA, disponían parcialmente de este atributo
- 2) Una disponibilidad total se puede observar en el sitio **GIDIS** (<http://gidis.ing.unlpam.edu.ar/>). Este sitio no fue parte del caso de estudio.

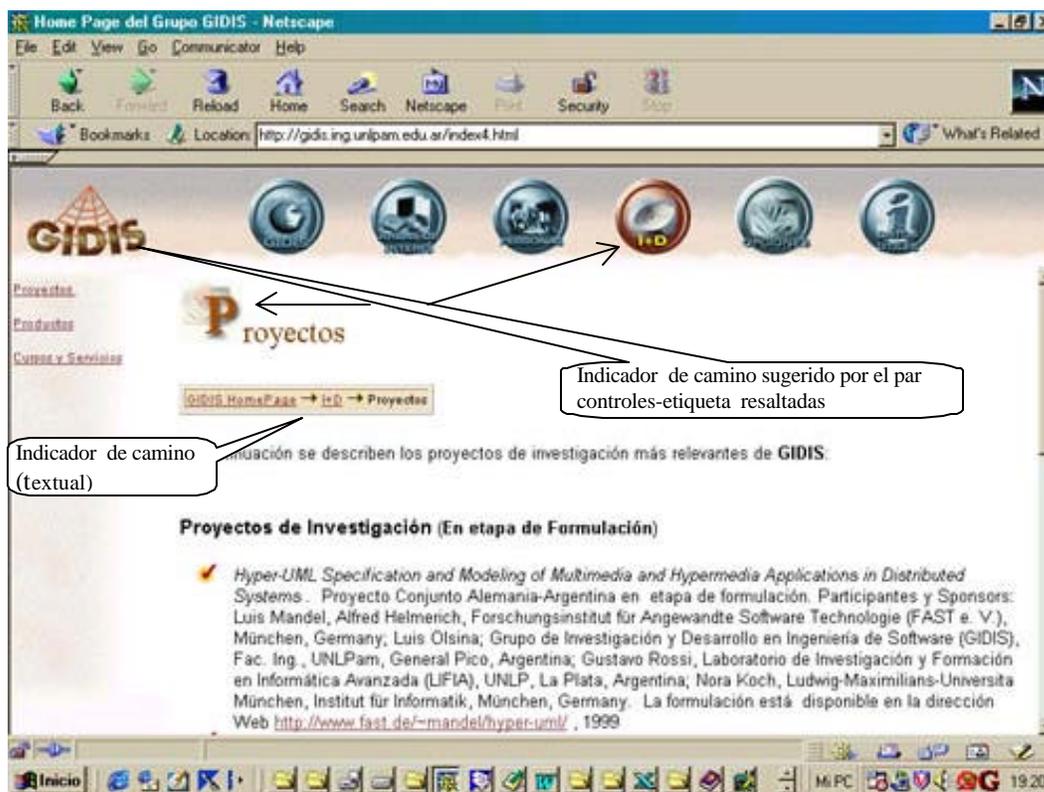


Figura 9.4 Indicador de camino (path) en el sitio **GIDIS**, sugerida tanto por el par control-etiqueta resaltada, como por la tabla con texto en la parte superior del marco de contenido (**GIDIS's Homepage-> I+D-> Proyectos**)

Título: **Soporte a Versión sólo Texto**; Código: 4.1.1 ; Tipo: **Atributo**

Característica de más Alto Nivel: **Eficiencia**

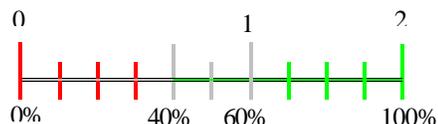
Super-característica: **Accesibilidad de Información**

Definición / Comentarios: Este atributo representa la accesibilidad a la información que está en las páginas, principalmente para las personas con discapacidad visual o cuando la velocidad es un problema [W3C 99]. Es de relevancia que el sitio entero sea editado

en una versión sólo texto; sin embargo, a veces, una disponibilidad parcial también puede ser deseable.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde: 0 = no disponible; 1 = disponible parcialmente; 2 = disponible totalmente

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s: Ningún sitio disponía de este atributo al momento de la recolección de datos.

Por otra parte, si bien la utilización de las plantillas para especificar características, subcaracterísticas y atributos es el mecanismo más adecuado para volcar y comunicar la información (y documentarla en hipertexto), sin embargo, generalmente ocupa más espacio y puede ser menos visual. Por ello, una forma resumida de lo anterior y que lleva suficiente información para comprender el criterio de preferencia de calidad elemental, podría ser como el que se muestra en la figura 9.5. Aquí se especifican cinco atributos con elementos como código, título, somera descripción, definición de la variable (unidad, si fuera necesario) y la escala de preferencia asociada.

<p>1.1.1.2 Índice Global</p> <p>Mecanismo que permite estructurar a través de un ordenamiento ya alfabético, ya numérico, o de otro tipo, el contenido de todo el sitio Web permitiendo navegación, principalmente desde sus componentes textuales.</p> <p>0 = No disponible (IE_i = 0 %) 1 = Disponible (IE_i = 100 %)</p>		<p>2.1.1.1 Búsqueda Restringida (Colecciones)</p> <p>0=No disponible 1=Búsqueda Básica: por Autor y/o Título 2= incluye 1 y Búsqueda Expandida: por Escuela y/o Estilo y/o Siglo (o fecha) y/o Pintura y/o Medio</p> <p>Ej: ver el museo NGA en figura 9.6</p>	
<p>2.2.3.2.2 Desplazamiento Horizontal</p> <p>HS = Nivel de desplaz. horizontal que el visitante debe realizar para ajustar la interface (considerando 640x480 como la mínima resolución) : 0 -> Alta; 0,5 -> Baja; 1 -> Sin desplazamiento</p> <p>La fórmula es: $X = 100 * HS$</p>		<p>2.2.4.1 Enlace con Título</p> <p>Enlace que al pasar el mouse por encima, muestra en un menú tipo pop-up texto explicativo que ayuda a predecir la navegación antes que el visitante la seleccione. 0 = No disponible; 1 = Parcialmente disponible ; 2 = Totalmente disponible</p>	
<p>2.2.4.2 Calidad de la Frase de Enlace</p> <p>Valoración de Preferencia Directa de la calidad de la frase del enlace. Cuando los enlaces no son suficientemente descriptivos, el visitante podría no tener una buena pista respecto de lo que dichos enlaces significan, principalmente cuando no hay un buen contexto. Por lo tanto, se tiene una buena frase de enlace si es auto-descriptiva (significativa cuando se lee fuera de contexto), es concisa y única. "When links are not descriptive enough, do not make sense when read out of context, or are not unique, the auditory user must stop to read the text surrounding each link to identify it" [W3C 99]</p>			

Figura 9.5 Conjunto de cinco atributos y criterios elementales, representados por la escala de preferencia.

Esta forma reducida puede ser útil para documentar en papel gran cantidad de atributos, como puede ser el caso en proyectos de evaluación entre 80 a 500 elementos. Por ejemplo, si observamos la figura 9.5, el atributo 2.1.1.1 representa a la *Búsqueda Restringida*, en donde claramente se ve en una escala ordinal los valores que toma la variable X, y cómo se corresponden con el indicador de calidad elemental IE, por medio de la escala de preferencia.

La medida del atributo 2.2.1.1 es válida porque, por una parte el artefacto exhibe el atributo que se intenta medir (ver figura 9.6); además es válida porque se utiliza una escala ordinal y la unidad es apropiada. (Respecto a la validez de las mediciones, le dedicaremos el próximo capítulo).

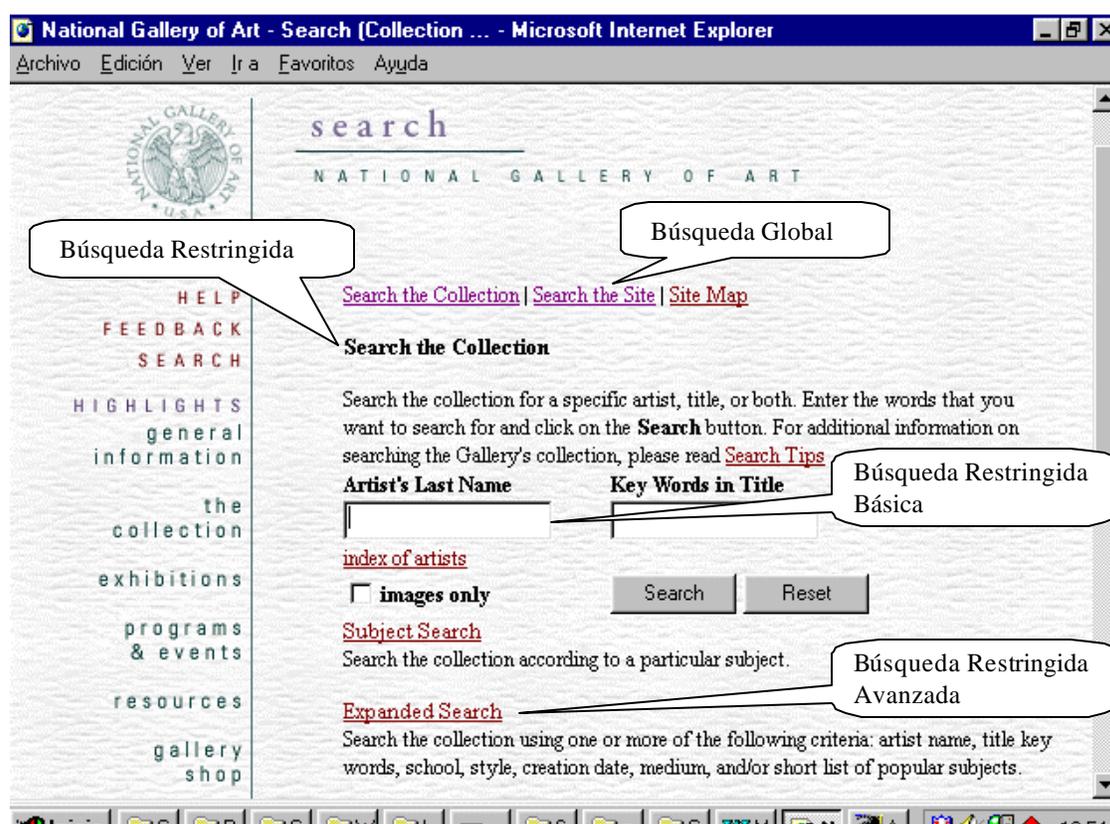


Figura 9.6 Pantalla bajada del museo NGA, en donde se muestra los dos tipos de búsquedas: Global y Restringida. A su vez, la búsqueda restringida se categoriza en Básica y Avanzada

Por último, un comentario respecto al atributo 4.2.1 (*Páginas Rápidas*), el cual tenía una observación (**) en la figura 9.1. Particularmente, este atributo no fue incluido en la evaluación de *Eficiencia*, a causa de que todos los museos incluyen para la mayoría de las imágenes sus respectivas ampliaciones (zooming), cuyos tamaños varían entre 50 y 90Kb aproximadamente. Esto es una buena característica para los visitantes aunque la performance pueda verse afectada. Es más, con la incorporación de las visitas guiadas

virtuales (p. ej. como la del museo Louvre, a partir de Nov-98), cada objeto de este tipo (formato mov) tiene un tamaño que puede rondar entre 150 y 400Kb aproximadamente. Esto afectaría el criterio generalmente usado para medir este atributo, según lo empleamos en 6.4.3.

9.2.2.1 Mediciones Elementales. La siguiente tabla muestra algunos de los valores obtenidos para las preferencias de calidad elemental, para cada uno de los cuatro sitios de museos en la Web que integraron el caso de estudio (en el Apéndice A –sección A.2-, se encuentran registrados los valores para todos los atributos).

Tabla 9.1 Resultados parciales de las preferencias de calidad elemental, para los cuatro sitios de museos en la Web

1. Usabilidad A _i	Louvre IE _i	Prado IE _i	Metropolitano IE _i	Gallería de Arte IE _i
1.1.1.1	0	0	0	0
1.1.1.2	0	0	0	100
1.1.1.3	0	0	0	100
1.1.3.1	50	100	50	100
1.1.4	100 A=1	80 A=0.8	80 A=0.8	0 A=0
1.2.1.1	60	60	60	100
1.2.1.2	0	100	0	100
1.2.2.1	100	100	0	0
1.2.2.2	0	0	0	0
1.4.1	90 N=3; S=1	60 N=1; S=2	0 N=0	24 N=4; S=0.2
2. Funcionalidad	Louvre	Prado	Metropolitano	Gallería de Arte
2.1.1.1	0	100	0	100
2.1.1.2	0	0	0	70
2.2.1.2.1	0 A=0	70 A=0,7	70 A=0,7	70 A=0,7
2.2.3.2.1	70 VS =0,7	70 VS =0,7	70 VS =0,7	70 VS =0,7
2.2.3.2.2	100 HS =1	100 HS =1	100 HS =1	100 HS =1
2.3.1	80	80	80	80
2.3.3.1.1	0	100	100	100
2.3.3.1.2	0 A=0	80 A=1	80 A=1	80 A=1
2.3.3.2	0	0	100	100
2.3.4.1	0 I=0	100 I=1	50 I=0,5	0 I=0
2.3.4.2	100 Z=1	100 Z=1	100 Z=1	100 Z=1

9.2.3 Evaluación Global

Las actividades, modelos y procedimientos de este proceso, fueron tratados en el capítulo 7, de manera que presentaremos los resultados del proceso. Vale recordar que los evaluadores deben preparar e implementar la evaluación global con el fin de obtener un indicador de calidad para cada sistema seleccionado. Aplicando un mecanismo de agregación paso a paso, los indicadores de calidad elementales deben agregarse usando una estrategia bottom-up para permitir el ulterior cálculo de las preferencias parciales respectivas. A su vez, repitiendo este proceso de agregación paso a paso, al final se genera la estructura de agregación de todo el sistema a evaluar.

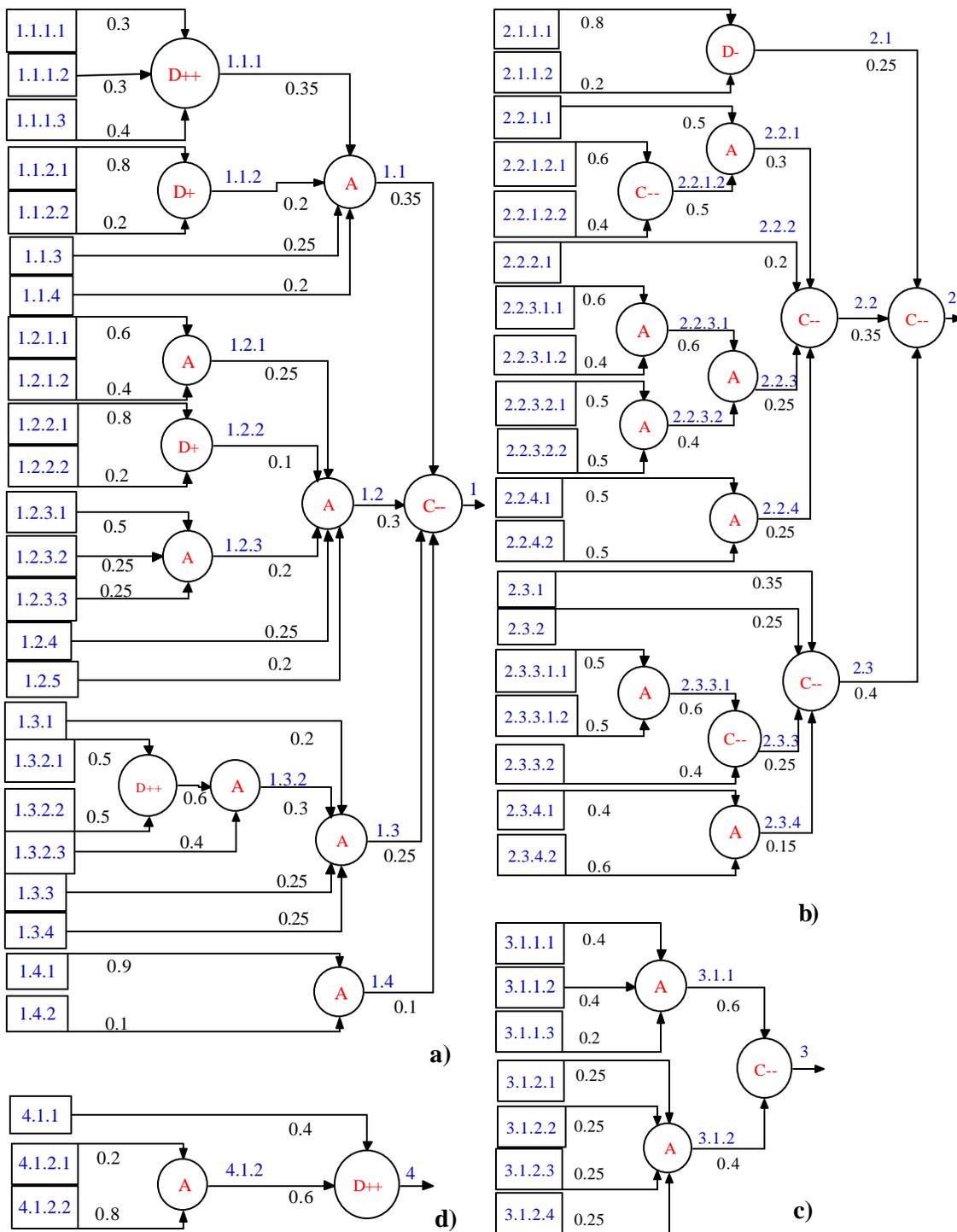


Figura 9.7 Estructura de agregación de preferencias parciales usando el modelo LSP para el caso de estudio de museos en la Web. En la parte (a) se muestra la estructura de agregación para la característica de alto nivel denominada Usabilidad; en la parte (b) se muestra la estructura de agregación para la característica Funcionalidad; en la parte (c) se muestra la estructura para la característica Confiabilidad y, en la parte (d) de la figura, para la característica Eficiencia

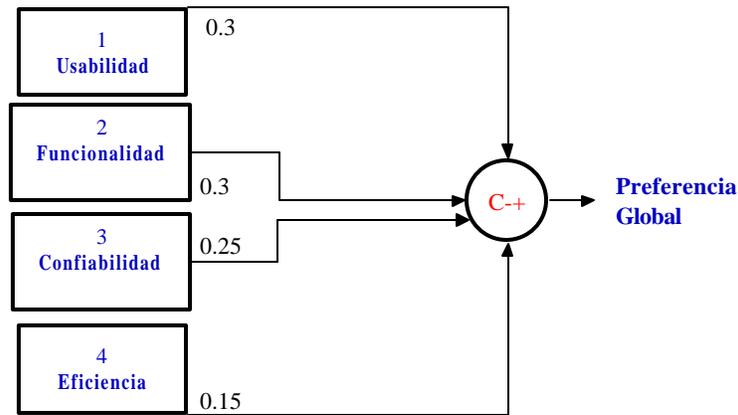


Figura 9.8 Estructura de agregación de preferencias parciales para las características de más alto nivel, conforme a un visitante general.

Por otra parte, una vez establecida la estructura de agregación, se debe llevar a cabo el proceso de cómputo de las preferencias globales de calidad para cada sitio. La preferencia global de calidad representa el grado de satisfacción de todas las necesidades involucradas.

Las figuras 9.7 y 9.8 representan las estructuras o esquemas de agregación de preferencias parciales y globales usando el modelo LSP para el caso de estudio de museos en la Web. La tabla 9.2 muestra los resultados detallados para las preferencias parciales y globales de calidad para los cuatro sitios evaluados. Es decir, los valores calculados para las subcaracterísticas y características fueron obtenidos indirectamente (no por mediciones directas).

9.2.4 Análisis y Comparación de la Calidad de los Sitios Evaluados.

Los valores de las preferencias de calidad global para cada museo Web evaluado se muestran al final de la tabla 9.2, y en la tabla 9.3; además un ranquin de dichos valores se aprecia en el gráfico de barras de la figura 9.9.

El proceso de evaluación, en consideración de un visitante general (casuales e intencionales), ha ranqueado primero al sitio del museo Galería de Arte (Washington) con un 79.26% de la preferencia global de calidad, luego al sitio del museo del Prado (Madrid) con el 68.40%; tercero, al sitio del museo del Louvre (París) con el 51.74% (cayendo en la barra de calidad gris), y, finalmente, cuarto en el ranquin, el sitio del museo de Arte Metropolitano (Nueva York) con el 50.95% de la preferencia global.

Observando la tabla 9.3, podemos decir que el sitio del museo NGA tiene una preferencia de calidad satisfactoria en todas sus características de más alto nivel; que el del Prado, tiene una preferencia de calidad satisfactoria en todas sus características de más alto nivel (excepto para *Usabilidad*, la cual ha caído en el rango marginal). Para los sitios de los museos Louvre y Met los niveles de preferencia para dichas características

por debajo del 60%, son la regla antes que la excepción. Por ejemplo, el museo Metropolitano obtuvo entre el 45 y el 53% para *Usabilidad*, *Funcionalidad* y *Confiabilidad* (observar además los pesos relativos en la figura 9.8; particularmente la característica *Eficiencia* tiene un bajo peso relativo, debido a que no intervino la subcaracterística *Performancia*). Sin embargo, el sitio del museo del Louvre obtuvo una muy baja preferencia para *Funcionalidad*, del 27,94% (y podemos afirmar que el menor o mayor grado de preferencia parcial de todos los sitios evaluados con respecto a *Funcionalidad*, no fue debido a la influencia del atributo *Relevancia de Contenido*, ya que todos recibieron mediante una preferencia directa la marca de 80% - sin embargo, para una evaluación más precisa, o para un perfil de usuario experto este debiera descomponerse).

Tabla 9.2 Resultados detallados de los valores de las preferencias parciales y globales de calidad, para los cuatro sitios Web de los museos evaluados

Características y Subcaracterísticas	Louvre	Prado	Met.	NGA.
1. Usabilidad	59.73	57.81	45.66	70.39
1.1 Comprensibilidad Global del Sitio	48.13	57	44.13	79.03
1.1.1 Esquema de Organización Global	0	0	0	98.54
1.1.2 Calidad en el Sistema de Etiquetado	78.15	80	78.17	97.68
1.2 Mecanismos de Ayuda y Retroalimentación	58.77	48.77	54	65
1.2.1 Calidad de la Ayuda	36	76	36	100
1.2.2 Indicador de Última Actualización	97.68	97.68	0	0
1.2.3 Directorio de Direcciones	100	100	100	100
1.3 Aspectos de Interfaces y Estéticos	70.41	72.53	74.93	90.91
1.3.2 Permanencia y Estabilidad en la Presentación de los Controles Principales	98.02	78.42	86.42	98.02
1.4 Misceláneas	81	54	0	21.6
2. Funcionalidad	27.94	72.67	49.19	80.41
2.1 Aspectos de Búsqueda	0	89.53	0	94.78
2.1.1 Mecanismo de Búsqueda en el Sitio	0	89.53	0	94.78
2.2 Aspectos de Navegación y Exploración	47.79	62.98	78.88	71.04
2.2.1 Navegabilidad Local	47.97	75	75	75
2.2.1.2 Orientación	15.93	70	70	70
2.2.2 Navegabilidad Global	80	80	80	80
2.2.3 Objetos de Control Navegacional	34	61.6	86.8	52
2.2.3.1 Permanencia y Estabilidad en la Presentación de los Controles Contextuales	0	46	88	30
2.2.3.2 Nivel de Desplazamiento	85	85	85	85
2.2.4 Predicción Navegacional	40	40	75	80
2.3 Aspectos específicos del Dominio y Funciones Varias	44.75	71.34	83.39	80.17
2.3.3 Aspectos de Comercio Electrónico	0	39.43	93.95	93.95
2.3.3.1 Características de Compras	0	90	90	90
2.3.4 Aspectos de las Imágenes	60	100	80	60
3. Confiabilidad	89.67	82.97	53	89.67
3.1 No Deficiencia	89.67	92.97	53	89.67
3.1.1 Errores de Enlaces	100	80	40	100
3.1.2 Errores o Deficiencias Varias	75	87.5	75	75
4. Eficiencia	62.44	62.44	64.39	80
4.1 Accesibilidad	62.44	62.44	64.39	79.99
4.1.2 Legibilidad al desactivar la Propiedad Imagen del Browser	64	64	66	82
Preferencia de Calidad Global	51.74	68.40	50.95	79.26

Tabla 9.3 Resultados de los valores de las preferencias de calidad para las características de más alto nivel, y valores finales para los para los cuatro sitios de museos evaluados.

Características	Museo Louvre	Museo Prado	Museo Met	Gallery of Art
1. Usabilidad	59,73	57,81	45,66	70,39
2. Funcionalidad	27,94	72,67	49,19	80,41
3. Confiabilidad	89,67	82,97	53	89,67
4. Eficiencia	62,44	62,44	64,39	80
Preferencia Global	51,74	68,40	50,95	79,26

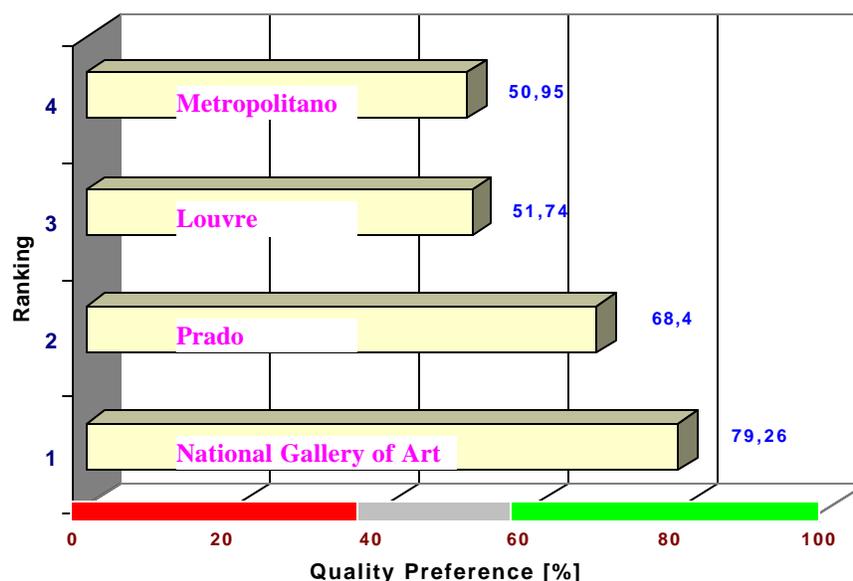


Figura 9.9 Ranquin final para los cuatro sitios de museos en la Web.

Si refinamos el análisis y queremos dar explicaciones de por qué el sitio del museo Louvre obtuvo el 27,94% para *Funcionalidad*, a partir de la tabla 9.2 vemos que para la subcaracterística 2.1 (*Aspectos de Búsqueda*) la marca fue cero, es decir, ¡no contaba ni con búsqueda global ni restringida, ni básica ni avanzada!. Para el tamaño del sitio, cantidad de información contenida, y cantidad de visitantes que tiene por día, esta falta es inadmisibles. Por lo tanto la recomendación es que debe urgentemente incorporar esta funcionalidad. (Esta falta también la tenía el sitio del museo Metropolitano hasta el período de evaluación, sin embargo, en Noviembre de 1998, incorporó búsqueda global básica –ver los comentarios en 4.2.1.2). Por otra parte, los museos del Prado y NGA, tenían excelentes mecanismos de *Búsqueda Restringida* (para Colecciones, -ver valores para el atributo 2.1.1.1 en tabla 9.1, y la pantalla de búsqueda en figura 9.6).

Además, el sitio Louvre presentaba problemas en las subcaracterísticas 2.2.1.2 (*Orientación*), y 2.2.3 (*Exploración Objetos de Control Navegacional*). Por ejemplo, no contaba con controles de navegación dentro del subsitio, el único mecanismo de control

era un botón al final de cada página para ir a la página principal. De manera que si uno arribaba al subsitio de “Colecciones”, y luego de navegar (por medio de los links semánticos) tres o cuatro páginas dentro del subsitio, el único modo de volver al tope del subsitio era por medio de la facilidad “atrás” del navegador, o volviendo al menú principal y de allí a “Colecciones”. Esto se soluciona fácilmente por medio de los objetos de control navegacionales. Finalmente, con respecto a la subcaracterística 2.3 (*Aspectos específicos del Dominio y Funciones Varias*), podemos observar que tres museos proveen al sitio de funcionalidad para comercio electrónico (excepto el Louvre), factor atractivo teniendo en cuenta una audiencia amplia.

En cuanto a *Usabilidad*, sólo el sitio NGA cae en la barra verde de aceptabilidad (con el 70.39%), los restantes tres sitios caen en la barra gris, entre el 45 y el 59%. Esto es debido en parte a que el sitio del museo NGA ha alcanzado un 79.03% de preferencia en la subcaracterística 1.1 (*Comprensibilidad Global del Sitio*) en tanto que el sitio del museo Met ha alcanzado un 44.13% de la preferencia en dicha subcaracterística (el más bajo). Por ejemplo, se puede observar que tres sitios (Prado, Met, Louvre) no tienen resuelta la subcaracterística 1.1.1, es decir es totalmente inaceptable el *Esquema de Organización Global* implementado. Estos sitios, no tienen disponibles atributos como *Mapa del Sitio* (ver 1.1.1.1 en la tabla 9.1), ni *Indice Global* (1.1.1.2), ni *Tabla de Contenidos* (1.1.1.3). Una recomendación elemental, es que debieran implementar algunos de estos atributos, principalmente, uno de los de mayor importancia relativa como *Tabla de Contenidos* (ver los pesos en el esquema de la figura 9.7a).

Respecto de las *Visitas Guiadas Convencionales* (1.1.3.1), en dos casos estaban marginalmente implementadas (sitios del Louvre, y Met -ver tabla 9.1). Por otra parte, el sitio del museo Louvre fue el que mejor soportaba el atributo 1.4.1, debido a que poseía soporte parcial a tres lenguajes extranjeros (español, japonés e inglés), en tanto que el museo del Prado obtuvo una marca menor, debido a que soportaba sólo al lenguaje inglés totalmente. Finalmente, Galería del Arte soportaba mínimamente 4 lenguajes, en tanto que el sitio del museo Metropolitano no tenía soporte a idiomas extranjeros. Sin embargo, NGA y Met, carecían al momento de la evaluación del *Indicador de Ultima Actualización* tanto global como por subsitio y estaban bien implementados en los dos restantes.

En cuanto a la característica de alto nivel *Confiabilidad*, algunas métricas son difíciles de automatizar, como es el caso para el atributo 3.1.1.2, *Enlaces Inválidos*. Este atributo es medido mediante el número de enlaces encontrados que conducen a un nodo destino semánticamente no relacionado o inválido; este tipo de errores es muy poco frecuente, no obstante, el museo del Prado acusó el siguiente error cuando se atravesaba los siguientes enlaces: “Home” – “Visitas” – “Colecciones” - Búsqueda seleccionando todos las escuelas, todos los temas, todos los estilos; se buscó el autor Goya desde el siglo I al XX – se recuperó varios documentos enlazados entre ellos “La Maja Vestida” y cliqueando este enlace, conducía a la “Maja Desnuda”.

Para el atributo 3.1.1.3, *Enlaces no Implementados*, se encontraron varios mapas de imagen (en el sitio del Prado y del Met), con áreas cliqueables que conducen al mismo nodo origen (esto es una deficiencia en la implementación, dado que el usuario tiene la sensación que no cliqueó bien, o que “el sitio no anda bien”; por ejemplo, en el sitio del Prado, en la página principal, en “Imagen Museo del Prado”, fuera del área de los botones dentro de la imagen, se produce esta deficiencia. En el sitio del Met, se encontraron 12 enlaces de este tipo).

Por otra parte, dentro de la subcaracterística 3.1.2, tenemos el atributo 3.1.2.3 (*Número de nodos Web Muertos*), es decir, representa aquellos nodos o páginas sin enlaces (de retorno) al sitio. Esta deficiencia se da si el usuario ingresa el URL absoluto de este nodo, una vez cargado no tendría modo de volver o recorrer el sitio (excepto que ingrese un nuevo URL). Este atributo es fácilmente automatizable mediante el lenguaje WebL y la herramienta que estamos desarrollando.

Tabla 9.4 *Valores máximos, mínimos, promedio y desvío estándar para las preferencias de calidad de las características de más alto nivel, y para la preferencia global en el caso de estudio de museos.*

Características	Máximo	Promedio	Mínimo	Desvío
1. Usabilidad	70,39	58,40	45,66	10,14
2. Funcionalidad	80,41	57,55	27,94	23,79
3. Confiabilidad	89,67	78,83	53,00	17,51
4. Eficiencia	80	67,32	62,44	8,50
Calidad	79,26	62,59	50,95	13,72

Para concluir este análisis, en la tabla 9.4 se muestran los valores máximos, mínimos y promedio para las preferencias de calidad de las características de más alto nivel como *Usabilidad*, *Funcionalidad*, *Confiabilidad*, y *Eficiencia* de los cuatro sitios de museos en la Web. Además, la figura 9.10 muestra el gráfico de líneas respectivo, y la figura 9.11 los desvíos correspondientes.

Es de hacer notar, que con este análisis de promedios y desvíos correspondientes a las características de los sitios típicos, no queremos extraer ninguna conclusión (o tendencia) generalizada. Simplemente queremos observar cómo se comportaron estos puntos, y dado que las características son semejantes ver qué se observa en ambos casos de estudio. (La tabla 9.5 y la figura 9.12 muestra ésta comparación).

Por ejemplo, la figura 9.11 muestra que la característica *Funcionalidad*, tuvo la mayor dispersión de puntos para los cuatro sitios, respecto del promedio comparado con las otras tres características. Ahora bien, si observamos y comparamos la misma característica con respecto al caso de estudio de sitios académicos, vemos que en promedio ambos obtuvieron el 57%, pero con un desvío casi del doble para los sitios de los museos.

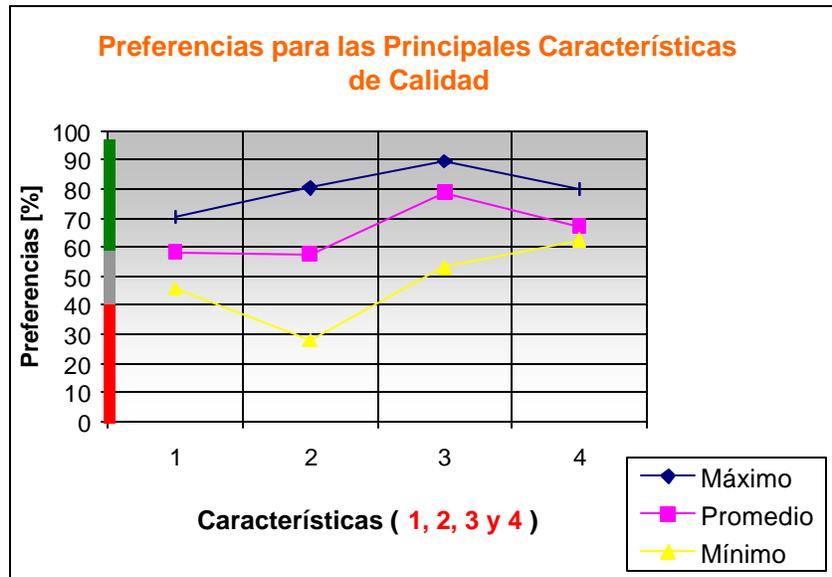


Figura 9.10 Valores máximos, promedio y mínimo para las características Usabilidad (codificada 1 en el árbol de requerimientos); Funcionalidad (2); Confiabilidad (3), y Eficiencia de los cuatro museos en la Web.

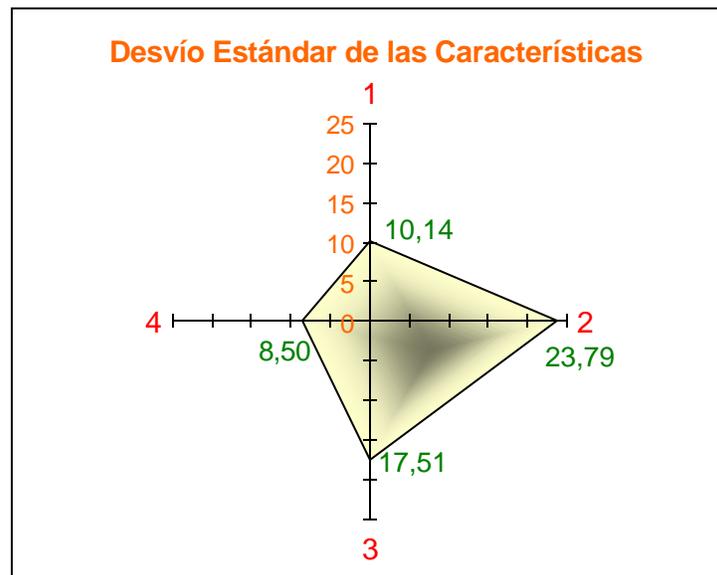


Figura 9.11 Desvío Estándar de las características Usabilidad; Funcionalidad; Confiabilidad y Eficiencia para los sitios de museos

Lo que se observa en este dominio en la Web al momento de las evaluaciones es una carencia en la implementación de mecanismos de búsqueda (principalmente confirmado por los valores arrojados en el survey, -ver A.3.2.2).

En la curva de los desvíos estándar de la figura 9.12 se observa en general que ésta es

más estable para los sitios académicos que para los sitios de museos estudiados. Sin embargo, como mencionamos anteriormente, de estos datos no se puede concluir que en general la calidad de los sitios académicos es mayor que la de los sitios de museos. Sería preciso realizar surveys sobre ambos dominios para analizar tendencias (no obstante, esto sería motivo de otros estudios y que están fuera del alcance y objetivo de esta tesis)

Tabla 9.5 Comparación de los promedios y desvíos estándar para las preferencias de calidad en ambos casos de estudio.

Características	Promedio Museos	Desvío Std	Promedio Universidades	Desvío Std
1. Usabilidad	58,40	10,14	64,33	11,45
2. Funcionalidad	57,55	23,79	57,34	14,57
3. Confiabilidad	78,83	17,51	82,77	13,48
4. Eficiencia	67,32	8,50	67,34	13,18
5. Calidad	62,59	13,72	65,29	9,21

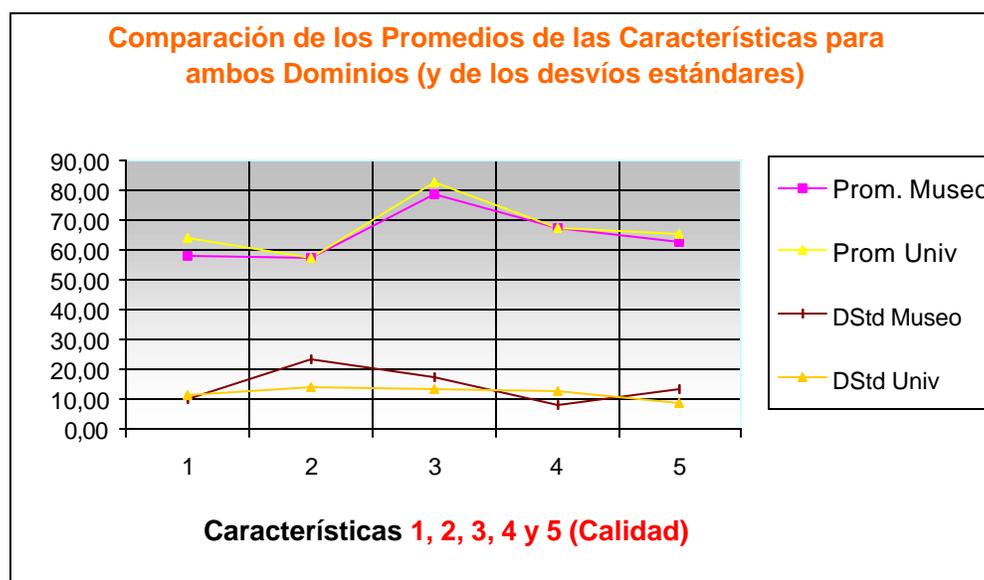


Figura 9.12 Comparación de los promedios y desvíos estándar para las características Usabilidad; Funcionalidad; Confiabilidad; Eficiencia y Calidad en ambos casos de estudio.

9.3 Caso de Estudio en Preparación: Sitios Web de Comercio Electrónico

Actualmente (julio de 1999), estamos preparando un caso de estudio para sitios típicos de comercio electrónico de venta de libros, a saber: Amazon (www.amazon.com), Barnes&Noble (www.BarnesandNoble.com), Cúspide (www.cuspide.com.ar), Borders (www.borders.com), y Díaz de Santos (www.diazdesantos.es). Los sitios con comercio electrónico, incorporan muchas de las mismas características que los comercios físicos.

Estudios de campo en el dominio de comercio electrónico fueron realizados, entre otros investigadores por [GVU 99, Lohse et al 98, Tilson et al 98]. Particularmente, Lohse et al. identificaron y midieron 32 atributos para 28 negocios en la Web, y documentaron características y atributos. Los autores identificaron en el survey, seis categorías para un comercio en la Web, a saber: Mercancías, Servicios, Promoción, Conveniencia, Confirmación de Venta, y Navegación, a partir de clasificaciones previas [Arnold et al 77]. Tomando en cuenta dichas categorías describieron los atributos que influyen al tráfico y a las ventas.

Particularmente, en esta sección nos interesa presentar para la característica *Funcionalidad*, los requerimientos más específicos del dominio para un usuario general, los cuales están en fase de análisis y discusión (no presentando otros aspectos de relevancia que deben ser personalizados como índices, búsquedas, ayudas, facilidad FAQ, etc., como vimos en los casos de estudio discutidos). En la figura 9.13, esbozamos un esquema para la subcaracterística *Funciones Específicas del Dominio*.

2.3 Funciones Específicas del Dominio

2.3.1 Información del Producto

- 2.3.1.1 Descripción del Producto/s
 - 2.3.1.1.1 Descripción Básica (título- autor – edición – tipo – tamaño/peso – ISBN-precio - disponibilidad)
 - 2.3.1.1.2 Descripción del Contenido
 - 2.3.1.1.3 Imagen del Producto
 - 2.3.1.1.3.1 Disponibilidad de Imagen
 - 2.3.1.1.3.2 Zooming
 - 2.3.1.1.4 Catálogo
 - 2.3.1.1.4.1 Disponibilidad de Catálogo/s
 - 2.3.1.1.4.2 Facilidad de Download de Catálogo/s
- 2.3.1.2 Evaluación del Precio
 - 2.3.1.2.1 Disponibilidad de Comparación de Precios
- 2.3.1.3 Disponibilidad de Rating del Producto
- 2.3.1.4 Recomendación de Productos Relacionados

2.3.2 Características de la Compra

- 2.3.2.1 Modos de Compra
 - 2.3.2.1.1 En Línea
 - 2.3.2.1.1.1 Carro de Compras

2.3.2.1.1.2 Facilidad 1-click

2.3.2.1.1.3 Transacción Segura

2.3.2.1.2 Fuera de Línea

2.3.2.2 Políticas de Compras

- 2.3.2.2.1 Información de Políticas de Devolución
- 2.3.2.2.2 Información de Envío
- 2.3.2.2.3 Información de Política de Crédito y Pago
- 2.3.2.2.4 Reenvío de una Compra (gift service)

2.3.3 Personalización del Cliente

- 2.3.3.1 Anuncios de Suscripción
- 2.3.3.2 Facilidad de Cuenta
- 2.3.3.3 Revisión o Valoración del Producto

2.3.4 Características del Almacén

- 2.3.4.1 Taza de Disponibilidad de Libros
- 2.3.4.2 Ranquin de los Productos
 - 2.3.4.2.1 Los Libros Top
 - 2.3.4.2.2 Los Libros Best Seller
- 2.3.4.3 Taza de Disponibilidad de Títulos

2.3.5 Política de Promoción

- 2.3.5.1 Promoción con la Venta
- 2.3.5.2 Anuncios de Promoción (Banners)
- 2.3.5.3 Appetizers (Sorteos, etc.)

Figura 9.13 Arbol parcial de requerimientos de calidad para el dominio de librerías en la Web, en consideración de un visitante general

A la subcaracterística 2.3, la dividimos en cinco componentes básicos: *Información del Producto* (codificada 2.3.1), *Características de la Compra* (2.3.2), *Personalización del Cliente* (2.3.3), *Características del Almacén* (2.3.4), y *Política de Promoción* (2.3.5). Concentrándonos en la *Información del Producto* se observan cuatro componentes o

elementos (subcaracterísticas y atributos) al mismo nivel. Por ejemplo, una subcaracterística es la *Descripción del Producto/s* (2.3.1.1.) que, a su vez, está compuesta de otras subcaracterísticas y atributos. En una descripción del producto intervienen elementos como una *Descripción Básica* del libro por título, autor, edición, tipo, tamaño/peso, ISBN, precio y disponibilidad (en este caso se puede utilizar un criterio de multivariadas discretas, como discutido en 6.3.2.4); la *Descripción de Contenido* del libro (este atributo se puede descomponer en niveles de disponibilidad y riqueza del contenido); la *Imagen* (gráfica) del producto (es deseable, que el usuario tenga una imagen ampliada del libro); un *Catálogo*, tanto para observar y navegar en línea por categorías, como para bajarlo a la PC e imprimirlo.

En el código 2.3.2, se especifica a las *Características de la Compra* en donde intervienen dos subcaracterísticas principales: *Modos de Compra* (2.3.2.1) y *Políticas de Compra* (2.3.2.2). En cuanto a la modalidad de compras en línea, atributos deseables son disponer de un *Carro de Compras*, *Transacción Segura*, y una característica rápida de compra (como por ejemplo, la *Facilidad 1-click*). Por otra parte, para *Políticas de Compra* el visitante (esencialmente para un potencial cliente), se debe especificar claramente las políticas de devolución, de pago y crédito, información de envío (tanto de costos adicionales como de días de tardanza), y además, opcionalmente, el servicio o facilidad de *Reenvío de una Compra* con fines de regalo.

En el código 2.3.3, se especifican atributos para la *Personalización del Cliente*. Un atributo importante es la *Facilidad de Cuenta*, en donde el cliente puede analizar sus compras y ver el estado del saldo y movimientos de su cuenta. Por otra parte, puede suscribirse a anuncios y novedades de su preferencia; y, una manera de personalización, y de sentimiento de comunidad (según estudios recientes), es la de posibilitar la *Revisión o Valoración de un Producto*.

Además, algunos atributos pertenecen a *Características del Almacén* con respecto a los productos que se ofrecen. Como apunta Lohse et al. los consumidores infieren información de la cantidad, variedad y calidad de los productos a partir del nombre y reputación del negocio. Algunos atributos deseables son la *Taza de Disponibilidad de Libros*, la disponibilidad por parte del negocio de un *Ranquin de los Productos*, como *Los Libros Top*, y *Los Libros Best Sellers*, sirven de orientación para varios visitantes generales.

Finalmente, visitas adicionales al sitio Web de comercio electrónico y potenciales ventas, se generan por medio de una *Política de Promoción*. Esta característica involucra la disponibilidad de *Promociones con la Venta* (principalmente en rebajas de precios); *Anuncios de Promoción* materializada en los banners; y *Appetizers*, como sorteos, competencias, entre otros mecanismos.

Capítulo 10

Aspectos de Validación de Métricas para la Web.

10.1 Introducción

Si bien en varias disciplinas tradicionales como la física, la química e ingenierías clásicas, la evolución, empleo y validación de métricas se ha desarrollado a lo largo de siglos y décadas, de modo que hoy muchas de las métricas están incorporadas en la vida cotidiana como medidas de temperatura, velocidad, distancia, entre otras, sin que nadie dude sobre la validez de las mismas; sin embargo, no sucede lo mismo en Ingeniería de Software, en donde todavía se debate si existe comprensión suficiente sobre las no demasiadas métricas existentes y popularmente empleadas como por ejemplo, punto función [Albrecht et al 83]. Así es que algunos autores como [Kitchenham et al 96] refutan la validez de la métrica citada dado que no cumple con los principios básicos de la condición de representación, indicando que viola, por ejemplo, los límites impuestos por el tipo de escala¹.

No obstante, métricas directas para medir el tamaño de un programa fuente procedural, como líneas de código fuente (Métrica LOC, o KLOC), o algunas métricas estructurales como Flujo de Información [Henry et al 81] o para un programa Orientado a Objetos [Chidamber et al 94], o métricas para medir densidad de defectos, entre otras, ya han sido validadas teórica y empíricamente (ver los estudios realizados en [Fenton 91, Fenton et al 97, Zuse 98]). Además, han sido validados modelos predictivos como el modelo COCOMO [Boehm 81], empleado en la determinación del esfuerzo de desarrollo de un sistema de software.

Con todo, los procesos y productos en el campo de Hipermedia y de la Ingeniería de Software en la Web, son más bien recientes, de manera que hay mucho por hacer para comprender, medir y validar. Por otra parte, el producto o aplicación basada en la Web posee muchos aspectos distintivos respecto del producto de software tradicional, como indicamos en el capítulo 2.

Sin embargo, no cabe duda que la actividad de *validación* de las métricas en la Web, como por ejemplo muchas de las empleadas en nuestra tarea de evaluación de sitios por medio de Web-site QEM, es un proceso importante, y al menos realizaremos algunos

¹ La métrica PF es independiente del lenguaje de implementación y se puede establecer en etapas tempranas del ciclo de vida de desarrollo con especificaciones de requerimientos completas y apunta a aspectos funcionales [IEEE 830] de los mismos. PF usa una escala ordinal (simple, promedio y compleja) para clasificar a las Entradas, Salidas, Consultas, Interfaces y Archivos, sumando luego los números resultantes (no permitido en la escala ordinal)

aportes iniciales partiendo de los avances en investigaciones recientes. Particularmente, nos concentraremos en la validación de métricas útiles para evaluar atributos de entes existentes (como un sitio) y no trataremos la validación de modelos y sistemas predictivos para la Web. Esta es realmente una línea abierta para futuras investigaciones.

Podemos definir a la validación como al proceso de asegurar que las medidas sean una caracterización numérica apropiada del atributo de un ente, mostrando que se satisfaga la condición de representación [Fenton et al 97]. Esto es, que la correspondencia entre el dominio empírico (objeto del mundo real) y el nuevo dominio numérico o simbólico (objeto del mundo formal) preserve a la relación de manera que estudiando y analizando a los números podamos explicar y conjeturar sobre el ente del mundo real. Dicho con palabras de los autores:

“...the representation condition asserts that a measurement mapping M must map entities into numbers and empirical relations into numerical relations in such a way that the empirical relations preserve and are preserved by the numerical relations” ... “For the (binary) empirical relation ‘taller than’ we can have the numerical relation $x > y$. Then, the representation condition requires that for any measure M , A is taller than B if and only if $M(A) > M(B)$...” (Fenton et al 97, pp. 31-32)

Del mismo modo, se puede tener una relación empírica unaria “es alto” que se corresponda con la relación numérica $x > 1,75$ [mts], y la condición de representación requiere que para cada medida M , A es alto sí y sólo sí $M(A) > 1,75$ [mts]. Por lo tanto, se puede ver que puede haber varias medidas (valores) para un atributo: cualquier medida que satisfaga la condición de representación es una medida válida.

Con todo, consideremos el atributo 2.1.1.1 del árbol de requerimientos de la figura 9.1 denominado *Búsqueda Restringida* (para Colecciones). Observando y utilizando un sitio Web comprendemos que existen categorías para el mecanismo de búsqueda, que facilitan al usuario una *menor o mayor funcionalidad* en el proceso de búsqueda (ver por ejemplo la pantalla de la figura 9.6). Por lo tanto, nuestra comprensión inicial del atributo a partir del sistema empírico (experimentando con varios sitios), puede llevarnos naturalmente a categorizar dicha funcionalidad en tres relaciones unarias BR_1 , (ningún mecanismo de búsqueda restringida disponible –mec1, para mayor brevedad); BR_2 , (disponible mecanismo de búsqueda restringida básica, por Autor o Título –mec2) y BR_3 , (disponible mecanismo de búsqueda restringida expandida, por Autor o Título por Escuela y/o Estilo y/o Siglo y/o Pintura y/o Medio –mec3). Entonces asumimos que cada mecanismo observado o es BR_1 , o BR_2 , o BR_3 . Además, supongamos por lo pronto, que no consideramos la relativa mayor o menor funcionalidad que implican los miembros de las clases, pero queremos hacer corresponder los miembros de las clases del mundo empírico, en tres diferentes números reales, en el mundo o representación numérica o matemática. Por ejemplo, se puede asignar una correspondencia o mapeo M ,

tal que:

$$\begin{aligned} M(\text{mec1}) &= 5; \\ M(\text{mec2}) &= 1; \\ M(\text{mec3}) &= 2; \end{aligned} \quad (1)$$

Esta asignación es una representación, debido a que tenemos relaciones numéricas² correspondientes a BR_1 , BR_2 , y BR_3 respectivamente. Esto es, la representación numérica correspondiente a BR_1 es la relación “es 5”, y del mismo modo las restantes.

Ahora bien, si nosotros queremos indicar a la anterior representación para que exprese la relación empírica binaria “*es más funcional que*” la cual permita establecer un ranking, sabiendo que el mec3 es más funcional que el mec2, y que el mec2 es más funcional que el mec1, entonces, tenemos que buscar un mapeo más apropiado. De este modo, “*x es más funcional que z*” si se cumple la siguiente condición de representación:

x pertenece a BR_3 y z pertenece a BR_1 o BR_2 ;
o,
 x pertenece a BR_2 y z pertenece a BR_1

Para encontrar una representación numérica más apropiada que la expresada en (1) para la relación empírica más compleja que la inicialmente planteada, debemos tener más cuidado con la asignación de los números. Una representación numérica más aceptable que la anterior es la siguiente:

$$\begin{aligned} m(\text{mec1}) &= 0; \\ m(\text{mec2}) &= 1; \\ m(\text{mec3}) &= 2; \end{aligned} \quad (2)$$

Esta representación numérica expresa en el mundo formal (matemático) una relación de orden a partir de la relación empírica “*x es más funcional que z*”. Y la misma es la empleada para representar el atributo 2.1.1.1 (*Búsqueda Restringida* para Colecciones) como se muestra en la figura 9.5. Es decir, en dicha figura se encontraba expresado de la siguiente forma:

0 = No disponible el mecanismo de búsqueda
1 = Búsqueda Básica: disponible por Autor y/o Título
2 = incluye 1 y Búsqueda Expandida: disponible por Escuela y/o Estilo y/o Siglo (o fecha) y/o Pintura y/o Medio

Por lo tanto, la métrica del atributo 2.1.1.1 satisface la condición de representación y es,

² que, en este caso, debido a la transformación de equivalencia uno-a-uno (permitido en el tipo de escala nominal) pueden ser interpretados como símbolos o etiquetas.

en primera instancia, una métrica válida.

10.1.1 Sistema Relacional Empírico y Formal

Prosiguiendo con esta introducción, podemos indicar que existe una teoría, denominada teoría de la medición ([Krantz et al 71, Roberts 79], entre otros), la cual declara cómo combinar condiciones empíricas con condiciones numéricas bajo la asunción de equivalencia lógica u homomorfismo. Extensiones de los fundamentos de esta teoría con implicancias para la Ingeniería de Software, están bien analizadas y documentadas en el libro de Zuse [Zuse 98].

Por ejemplo, sea S un conjunto de entes (sitios Web) en donde X es un atributo observable de modo que x_1, x_2 pertenecen a S_1, S_2 respectivamente; así, la relación empírica binaria $x_1 \bullet > x_2$ se mantiene, sí y sólo sí se juzga que x_1 es más funcional que x_2 . Por lo tanto, se desea asignar un número real $m(x_1)$ y $m(x_2)$ para cada x_1, x_2 de modo que para todo par perteneciente a S , se cumple que

$$x_1 \bullet > x_2 \Leftrightarrow m(x_1) > m(x_2) \quad (3)$$

Dicha declaración es la base de toda medición, y puede ser leída de dos formas: primero, si x_1 es más funcional que x_2 implica (\Rightarrow) en el mundo formal, numérico, que $m(x_1)$ es mayor que $m(x_2)$. Y viceversa, si $m(x_1)$ es mayor que $m(x_2)$, entonces implica (\Leftarrow) que x_1 es más funcional que x_2 . Esta doble implicación es denominada homomorfismo. Según Fenton et al.,

“...the representation condition asserts that a measurement mapping M must map entities into numbers and empirical relations into numerical relations in such a way that the empirical relations preserve and are preserved by the numerical relations”

En (3) la declaración empírica es: $x_1 \bullet > x_2$; y la declaración formal, numérica es: $m(x_1) > m(x_2)$; siendo $\bullet >$ y $>$, los operadores relacionales respectivos. La declaración empírica ($x_1 \bullet > x_2$), no puede ser definida y probada sobre un escritorio como un modelo matemático, dice Zuse, sino que proviene de experimentos, de la experiencia real del día a día, y de la intuición. (Además, deja sentado que las características de calidad de alto nivel expresadas en el estándar ISO 9126, son definidas como declaraciones empíricas. Es importante indicar que las declaraciones empíricas no son verdaderas *per se*; pueden ser falseadas por medio de la observación).

De manera semejante, si escribimos la siguiente expresión:

$$x_1 \bullet \geq x_2 \Leftrightarrow m(x_1) \geq m(x_2) \quad (4)$$

Puede ser leída como: si x_1 es igual o más funcional que x_2 implica (\Rightarrow) que $m(x_1)$ es

igual o mayor que $m(x_2)$; y viceversa.

Por otra parte, las expresiones (3, 4) representan a la propiedad de orden de ranquin, u homomorfismo de ranquin [Zuse 98] pp. 106,109. Para la relación empírica $\bullet \geq$ puede haber varias interpretaciones, según el caso. Ejemplos son: igual o más difícil de comprender; igual o mayor nivel de disponibilidad, igual o mayor nivel de errores, etc. El autor define a esta relación como de orden débil (weak order), la cual cumple las propiedades de transitividad y completitud (pp. 113-114), e indica: “*the weak order is a prerequisite for ranking order measurement*”

En dicha literatura, el autor introduce rigurosamente el concepto de sistema relacional empírico, sistema relacional numérico y el de métrica (o medida). Para la estructura de ranquin, sea el sistema relacional empírico definido como:

$$\mathbf{S} = (S, \bullet \geq),$$

y el sistema relacional numérico como:

$$\mathbf{N} = (R, \geq), \text{ donde } R \text{ es el dominio de los números reales,}$$

Entonces, podemos escribir para la métrica m , la siguiente expresión:

$$(\mathbf{S}, \mathbf{N}, m) = ((S, \bullet \geq), (R, \geq), m) \quad (5)$$

En donde la métrica es una correspondencia (mapeo) $m: S \rightarrow R$ en la que se cumple la expresión (4)

Bajo un procedimiento semejante, se define a la estructura de aditividad para una métrica, según la expresión:

$$(\mathbf{S}, \mathbf{N}, m) = ((S, \bullet \geq, o), (R, \geq, +), m), \text{ en donde } o \text{ es el operador de concatenación.}$$

Dice Zuse en la pag. 109: “*we want to have something above poor ranking or comparing of objects. We want to be additive in the sense that the combination of two objects is the sum of their measurement values*”. Luego define la estructura extensiva la cual describe las condiciones empíricas, es decir, el modelo cualitativo que está detrás de las métricas. En breve, el sistema relacional empírico $(S, \bullet \geq, o)$ asume una estructura extensiva (modificada), sí se cumplen los axiomas de orden débil, asociatividad, conmutatividad y monotonicidad débil, y el axioma Arquimideano. Según Zuse, las implicaciones para Ingeniería de Software son que el teorema de la estructura extensiva da criterios para la validación de métricas y de modelos predictivos, entre otros beneficios.

10.1.2 Magnitud, Tipos de Escala y Unidades

Magnitud (scale) y tipo de escala son dos entidades diferentes que con frecuencia son confundidas. El concepto de magnitud queda definido por la tupla $(\mathbf{S}, \mathbf{N}, m)$ (ver la

expresión (5)). Conforme a esa definición se observa que una medición asume la equivalencia lógica u homomorfismo. Claramente se aprecia que necesitamos el sistema relacional empírico S , el sistema relacional numérico N , y la métrica m con el fin de medir y obtener un valor o magnitud. En cambio, un tipo de escala se define por medio de una transformación admisible. Una transformación admisible es una regla de conversión f en la que dadas dos magnitudes m y m' , se mantiene que $m' = f m$. Por ejemplo, transformaciones admisibles son, entre otras, $m' = a m + b$, con $a > 0$ y $b \in \mathbb{R}$; $m' = a m$; etc. El tipo de escala no cambia al realizar una transformación admisible.

Los tipos de escala más comunes son el nominal, ordinal, intervalo, proporción y absoluta (una extensa explicación con ejemplos se puede encontrar en el libro de [Fenton et al 97] en pp. 45-53, y en el de [Zuse 98] en pp. 130-150, entre otros). Los tipos de escala de una medida, afecta el tipo de transformación admisible, las operaciones aritméticas y el análisis estadístico que se puede aplicar sobre los valores. Por ejemplo, para una magnitud expresada en la escala nominal, una media o promedio es inadmisibles, al igual que la suma o resta en la escala ordinal. A seguir, discutiremos estos aspectos de relevancia, que nos serán útiles para comprender un marco conceptual aplicable para validar métricas, en la sección 10.2.

La tabla 10.1 muestra los tipos de escala más empleados en Ingeniería de Software, las transformaciones admisibles, funciones y análisis estadísticos que se pueden aplicar. El tipo de escala debe leerse de un modo jerárquico, es decir, el nivel más bajo es el nominal y el más alto es el absoluto. Así, las transformaciones admisibles para la escala de tipo intervalo, son las propias más las permitidas para la ordinal y nominal.

Tabla 10.1 Tipos de Escala e información adicional

Tipo de Escala	Transformación Admisible	Estadísticas Significativas	Ejemplos
Nominal	Equivalencia uno-a-uno	✓ Moda, Frecuencia ✓ Test no paramétricos	Clasificación, etiquetado (ver (1))
Ordinal	Funciones monotónicas incrementales, (ver la expresión (4))	✓ Mediana, Spearman y Kendall Tau, W ✓ Test no paramétricos ✓ (Todo lo previo)	Grado de disponibilidad, funcionalidad, complejidad, etc.
Intervalo	$m1 = a m + b$	✓ Media Aritmética, Desvío estándar ✓ Coef. de correlación de Pearson ✓ Test no paramétricos ✓ (Todo lo previo)	Temperatura (Celsius, Fahrenheit) Tiempos relativos, etc.
Proporción	$m1 = a m$	✓ Porcentajes, Media Geométrica ✓ Test paramétricos y no paramétricos ✓ (Todo lo previo)	Preferencia de calidad, tamaño (LOC), longitud, etc.
Absoluta	$m1 = m$ (Identidad)	✓ (Todo lo previo)	Conteo

Es importante resaltar, que detrás de los tipos de escalas, se esconden condiciones empíricas y numéricas. Se puede ver a las condiciones empíricas como a una modelización de los hechos empíricos. Zuse les llama axiomas de condiciones empíricas. Los axiomas son asunciones de la realidad.

Por ejemplo, el tipo de escala nominal, implica una condición empírica muy simple: la relación de equivalencia. Supongamos que el sistema relacional empírico es (S, \approx) . Entonces existe una función $m: S \rightarrow R$, y dado un atributo observable de modo que x_1, x_2 pertenecen a S_1, S_2 ,

$x_1 \approx x_2 \Leftrightarrow m(x_1) = m(x_2)$, y
 $(S, N, m) = ((S, \approx), (R, =), m)$, es una magnitud nominal.

Es oportuno decir que Zuse sólo considera un mapeo en R , pero esto no es necesario, pues una representación simbólica única (por ej., etiquetas) para cada clase, es suficiente. Esa representación (numérica o simbólica) no representa un orden entre clases sino una identificación de las mismas.

Por otra parte, según este mismo autor, a un valor nominal y ordinal no se le puede asignar una unidad, debido a que para el primero un número no tiene significancia, y para el segundo, sólo expresa un orden o ranquin. Para Zuse, la unidad tiene significancia para tipos de escala superiores, a saber: intervalo, proporción y absoluta. Esta posición no es coincidente, por ejemplo, con la de [Kitchenham et al 96], quienes consideran que se debe extender el uso de unidades para valores de tipos de escala nominal y ordinal: “*We have extended the use of units in our structure model to allow for the definition of scale points for ordinal scale measures and the categories used for nominal scale measures ... Thus, in the context of nominal and ordinal scale measures where our measures are mappings to arbitrary labels (sólo para nominales), we suggest a “unit” is needed to ensure that such measures are used consistently*”, pp. 930.

Con respecto al tipo de escala ordinal, el sistema relacional empírico del nominal es extendido para reflejar las propiedades de ranquin según se expresa en (4). Así, una magnitud ordinal es:

$((S, \bullet \geq), (R, \geq), m)$

Zuse define a esta relación como de orden débil (weak order), la cual cumple las propiedades de transitividad y completitud (completeness or connectedness), a saber:

$x_1 \bullet \geq x_2$, y $x_2 \bullet \geq x_3 \Rightarrow x_1 \bullet \geq x_3$, (6) Transitividad

En donde la relación empírica $\bullet \geq$ describe la propiedad de ranquin.

$x_1 \bullet \geq x_2$, o $x_2 \bullet \geq x_1$ (7) Completitud

Retomando la métrica expresada en (2) del atributo 2.1.1.1 (presentada en la introducción) observamos que satisface la condición de representación de los sistemas relacionales empírico y numérico. La métrica produce un valor ordinal, por lo tanto cumple las propiedades (6) y (7). Así, para transitividad, combinando ambos sistemas relacionales, para la relación empírica $\bullet>$ (“más funcional que”) se cumple que:

$$x1 \bullet> x2, \text{ y } x2 \bullet> x3 \Rightarrow x1 \bullet> x3 \Leftrightarrow m(x1) > m(x2), \text{ y } m(x2) > m(x3) \Rightarrow m(x1) > m(x3); \text{ donde } m(x1) = 1, m(x2) = 2, \text{ y } m(x3) = 3$$

El tipo de escala limita el tipo de operaciones aritméticas y estadísticas permitidas, según se mostró en la tabla 10.1.

Finalmente, el tipo de escala proporción es bien conocido en física y ciencias tradicionales (por ej., medidas de distancia, longitud, dinero, etc.). La idea de una magnitud proporción está unida a la propiedad de adición (sin embargo, magnitudes proporción no aditivas son también posibles). Una magnitud proporción aditiva queda representada por:

$$((S, \bullet \geq, o), (R, \geq, +), m), \text{ en donde } o \text{ es el operador de concatenación}$$

Por lo tanto, $(S, \bullet \geq, o)$ es una estructura extensiva cerrada, si existe una función m sobre S tal que para todo $x1, x2$ perteneciente a S ,

$$x1 \bullet \geq x2 \Leftrightarrow m(x1) \geq m(x2), \text{ y } \\ m(x1 \text{ o } x2) = m(x1) + m(x2)$$

Además, existe una función m' que es la transformación admisible de la magnitud proporción, a saber: $m'(x) = a \cdot m(x)$. Por ejemplo, la métrica LOC, satisface la regla de combinación aditiva, y es una magnitud proporción. KLOC es una transformación admisible de LOC (donde $a=1000$, o 1024 , en algunos casos)³.

[Fenton et al 97], en pag. 51, identifica las siguientes características para el tipo de escala proporción: 1) es una correspondencia que preserva el orden, el tamaño de intervalos y proporción entre entidades; 2) Hay un elemento cero, representando la falta total del atributo; 3) El mapeo de medición debe comenzar en cero e incrementar a intervalos iguales, conocidas como unidades; 4) Toda aritmética puede ser significativa si se aplica a las clases en el rango de la correspondencia. El modelo de preferencias elementales y de agregación empleado en este trabajo cumple con estas características mencionadas.

³ Remitimos al lector a [Zuse 98], Cap. 5, en donde se discuten los axiomas del teorema de la estructura extensiva (y la modificada), condiciones de independencia, y las reglas de combinación.

En la próxima sección discutiremos un marco conceptual aplicable para validar métricas, complementándolo con las propiedades parcialmente introducidas en este trabajo⁴, conforme a investigaciones recientes [Briand et al 96, 97, Fenton et al 97, Kitchenham et al 96, Zuse 98].

Este marco conceptual y su estrategia asociada nos permitirá validar al menos teóricamente, la mayoría de las métricas empleadas, poniendo en tela de juicio, como ya lo advertimos en la sección 6.3.1.4, a aquellas valores obtenidos mediante una valoración directa y subjetiva. Por tal razón, cuando se utilizaba un criterio de preferencia directa para evaluar un atributo, el valor final debía surgir del consenso de dos o más evaluadores expertos, con el fin de minimizar errores de valoración intencionales y/o involuntarios. De todos modos no se podría validar teóricamente aunque sí empíricamente.

10.2 Un Marco Conceptual para Validar Métricas

10.2.1 Estructura y Modelo Conceptual para Métricas

En esta sección presentamos un modelo conceptual para métricas útil para representar a las principales clases y relaciones intervinientes. Además, describimos a cada clase y vemos distintas contribuciones de las mismas al proceso de medición y validación. La figura 10.1 muestra el esquema para métricas directas. Una métrica directa es la resultante de una correspondencia o mapeo directo entre un atributo de un ente (del dominio empírico) y el valor o magnitud (del dominio formal, numérico), y que sirve como referencia para describir y explicar aspectos o situaciones del mundo empírico. Los atributos se miden generalmente por medio de métricas directas (aunque no siempre) y las características y subcaracterísticas se miden por medio de métricas indirectas (también denominadas métricas derivadas en [Roberts 79, Zuse 98]).

En la figura se observan dos mundos o dominios modelados: el empírico y el formal. En el primer mundo (que tiene un impacto empírico, perceptible sobre los usuarios) se modela y formula el sistema relacional empírico; en el segundo mundo, se modela y formula el sistema relacional formal (numérico y, en algunos casos simbólico) y los mecanismos para ayudar en la determinación de valores. Vale decir que, al igual que Kitchenham et al 96, no asumimos que esta estructura conceptual sea suficiente. Además, la estructura formal axiomática presentada en [Roberts 79, Zuse 98], entre otros, no es siempre aceptada en la comunidad de métricas [Kitchenham et al 97]. Las principales clases y relaciones intervinientes se describen a continuación.

⁴ La investigación sobre validación de métricas para evaluación y de modelos predictivos sería motivo de otro trabajo de tesis *per se*

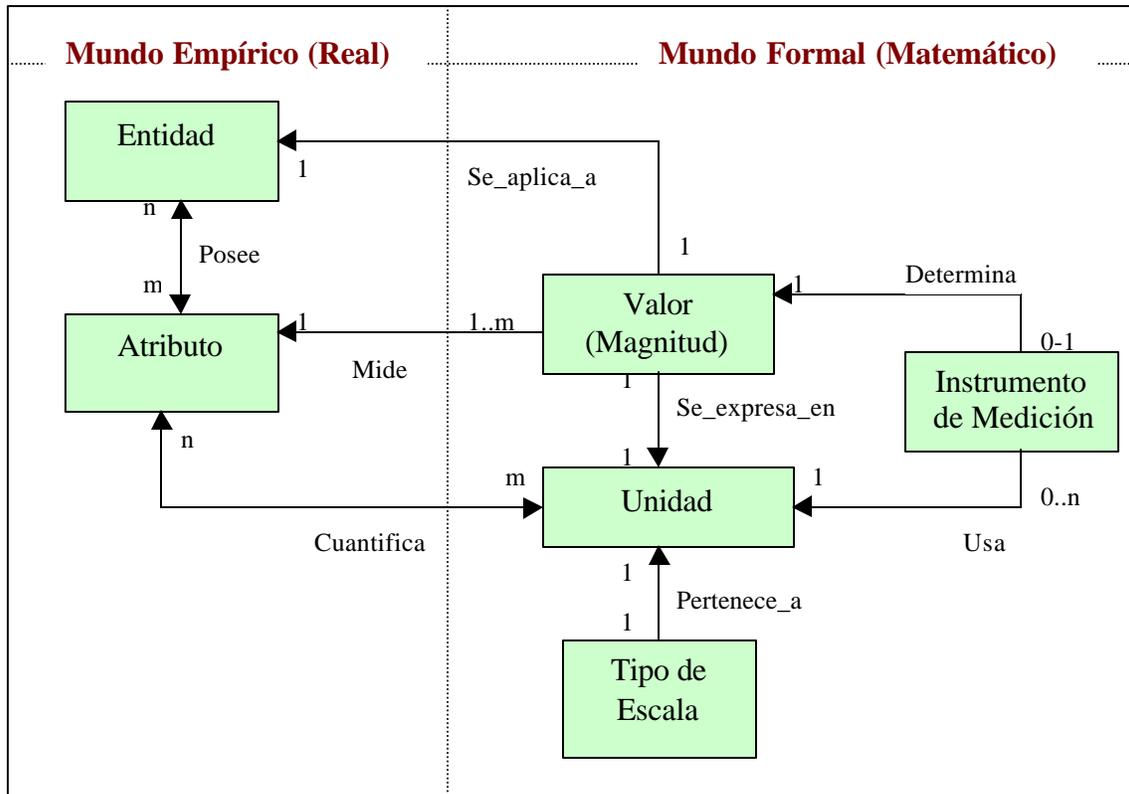


Figura 10.1 Modelo conceptual para métricas directas

10.2.1.1 Entes, Atributos y sus Relaciones. Desde el punto de vista de la evaluación, en el dominio empírico tenemos a la clase *Ente* y a la clase *Atributo*. Como vimos en la sección 5.1 los entes pueden descomponerse básicamente en tres clases principales de interés para los evaluadores, a saber:

- ✓ Proceso: *es el ente compuesto posiblemente de otros subprocesos y actividades, usado para producir artefactos (ver diagrama B.1 en el Apéndice B);*
- ✓ Artefacto: *es el ente temporario o persistente que representa al producto de realizar un proceso,*
- ✓ Recurso: *es un ente requerido como entrada por un proceso para producir alguna salida especificada (recursos de un proyecto son: humanos, monetarios, materiales, tecnológicos, temporales –ver diagrama B.2)*

La clase *Atributo*, representa a lo que se observa y atribuye respecto de lo que es propio de un ente del mundo real, y que es un elemento de interés a ser evaluado. En métricas de software, se suele definir a un atributo como a la característica elemental -que no soporta descomposición- de un ente, la cual se la puede medir generalmente mediante la aplicación de una métrica directa (o en algunos casos indirecta). Para un atributo dado, siempre hay al menos una relación empírica de interés que podemos capturar y representar en el mundo numérico, habilitándonos a explorar la relación matemáticamente (como discutimos en la sección 10.1).

En cuanto a la relación mostrada en la figura entre las clases *Ente* y *Atributo*, observamos una cardinalidad de muchos a muchos. Es decir, un ente puede poseer varios atributos, en tanto que un atributo puede pertenecer a varios entes. Por ejemplo, si hablamos de un atributo compuesto (característica) como calidad, puede pertenecer a un proceso, a un producto o a un recurso. En tanto que un ente, por ejemplo, artefacto o sitio Web, puede tener muchos atributos según vimos en los casos de estudio.

10.2.1.2 Unidades, Tipos de Escalas y sus Relaciones. Una *Unidad* de medida determina cómo se debe medir al atributo del ente. Por ejemplo, para medir el tiempo de respuesta de una base de datos, podría usarse diferentes unidades [segundos | milisegundos]. La figura 10.1 muestra una cardinalidad m a n entre la clase *Atributo* y la clase *Unidad*. Pero además de la unidad de medida se debe considerar los diferentes *Tipos de Escala*. Los tipos más comunes son el nominal, ordinal, intervalo, proporción y absoluta como describimos en la sección 10.1.2.

10.2.1.3 Valor. Cuando medimos un atributo específico de un ente particular, aplicamos una unidad perteneciente a un tipo de escala, para obtener magnitudes de tipo *Valor*. Por lo tanto, el valor medido no puede ser interpretado a menos que sepamos a qué entidad se aplica, a qué atributo se mide y en qué unidad se expresa (es decir, se debe especificar claramente al sistema relacional empírico y numérico). Este valor generalmente es numérico, pero podría ser simbólico como por ejemplo una etiqueta (en un tipo de escala nominal). La importancia de mapear a números es bien conocida, sin embargo, no todos los tipos de escala permiten aplicar las mismas operaciones aritméticas y funciones estadísticas. Según vimos, existen propiedades para las escalas y transformaciones admisibles para los tipos de escala. Por ejemplo, no se pueden sumar números en una escala ordinal ni obtener la media aritmética, pero sí se pueden aplicar funciones estadísticas más avanzadas en una escala de tipo intervalo, proporción o absoluta (ver tabla 10.1).

En rigor se espera que las medidas válidas sean definidas sobre un conjunto o rango de valores permitidos y que se pueda aplicar operaciones matemáticas y estadísticas permitidas para el tipo de escala. Este conjunto numérico puede ser finito o infinito, continuo o discreto (ver sección 6.3).

10.2.1.4 Instrumento de Medición. Para determinar el valor se puede hacerlo de un modo manual, o asistido total o parcialmente por un *Instrumento de Medición* o herramienta de recolección de datos y/o cómputo.

10.2.1.5 Métricas Indirectas. El modelo presentado en la figura 10.1 es apropiado para métricas directas, y la figura 10.2 representa un esquema para métricas indirectas. Es el caso cuando el valor de una medida se obtiene por medio de modelos matemáticos que involucran variables que se deben medir previamente.

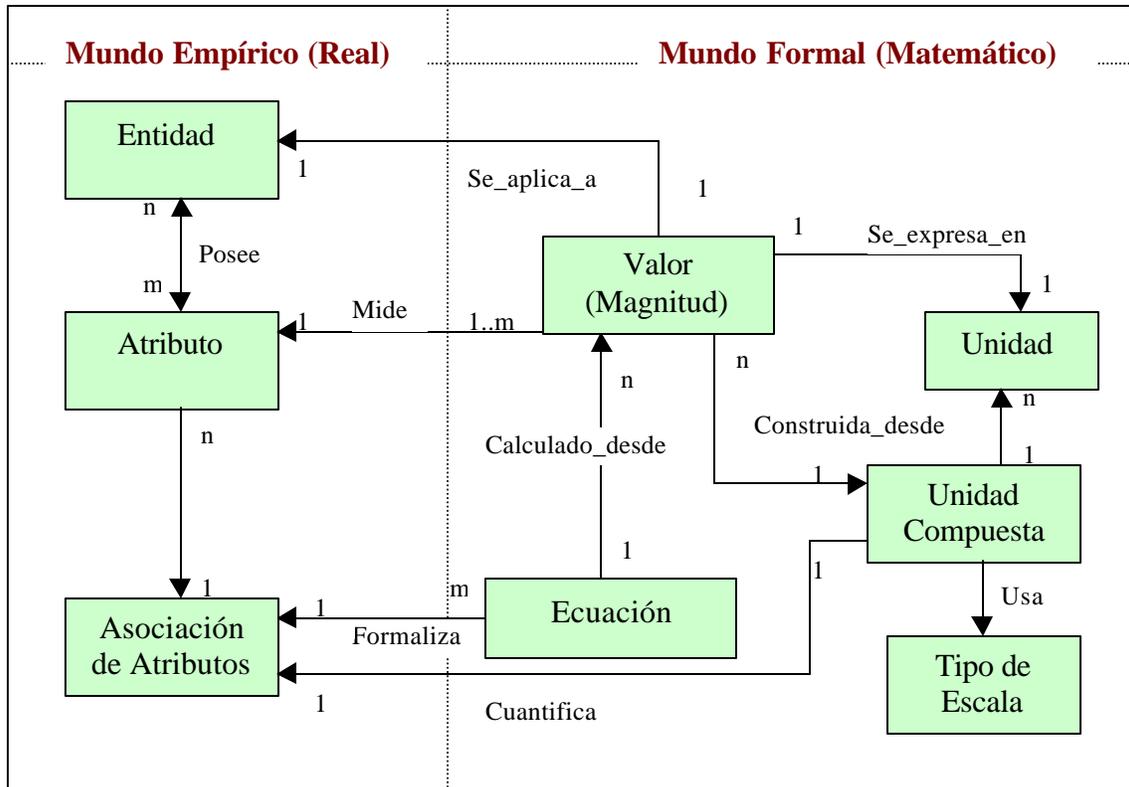


Figura 10.2 Un modelo conceptual para métricas indirectas

10.2.2 Implicaciones del Modelo para Validación de Métricas

Los modelos antes descriptos tienen varias implicaciones para la validación de métricas, y entre ellas se encuentran:

1. Diferentes entes (o subentes) pueden compartir el mismo atributo. En la práctica, cualquier marco conceptual que agrupe atributos como perteneciendo a un solo ente, no debe implicar que la relación sea de muchos a uno, en general.
2. Dado que una medida hace corresponder magnitudes a un atributo (del mundo empírico), entonces si la condición de representación se cumple, el comportamiento del atributo en el dominio empírico se debe ver reflejado en el comportamiento del nuevo dominio numérico. Es decir, se debe observar la equivalencia lógica u homomorfismo y esto puede implicar una validación teórica y/o empírica.
3. Dado que un atributo se puede medir de diferentes maneras (y con diferentes tipos de escalas), la definición de los atributos son independientes de las unidades en que las miden. Por lo tanto es inválida la definición de un atributo implicando a una unidad y tipo de escala específica.
4. Una unidad se puede aplicar a diferentes atributos que pertenecen a diferentes entes. Sin embargo, un valor que no tenga asociada una unidad es un claro signo de que la

medida está carente de significado. Es más, una medida específica, su unidad y tipo de escala, debe estar definida para un atributo específico, en un contexto concreto de evaluación. En este caso, se deben considerar las operaciones y transformaciones admisibles particulares.

5. Dado que métricas directas hacen corresponder valores a un atributo, se debe considerar la característica del dominio y el rango de valores posibles y deseables. Además, se debe determinar si el criterio de medición y, específicamente, el rango es finito o infinito, continuo o discreto.
6. Para una métrica indirecta se debe considerar además el modelo o ecuación para calcular las magnitudes.

10.2.3 Otros Modelos para Validación de Métricas

Además de los componentes de los modelos antes descritos, se debe tener en cuenta otros aspectos como modelos para la instrumentación de la medición, protocolos de medición, y modelos para relacionar atributos.

Por ejemplo, en un modelo de instrumentación de la medición, se debe ser consciente si el modelo que subyace en el instrumento de medición a emplear es válido, y si el instrumento está debidamente calibrado. En cuanto a la definición y uso de un protocolo de medición para un atributo específico de un ente específico, permite a los evaluadores ser consistentes y repetir apropiadamente el proceso. Por ejemplo, la plantilla para el atributo 4.1.1 de la tabla 5.4, favorece la definición del protocolo de medición para dicho atributo.

Finalmente, se emplea un modelo para corresponder valores de atributos a valores de preferencia, debido a que el valor de una subcaracterística y/o característica se deriva a partir de una agregación funcional y/o lógica entre indicadores elementales. Este modelo de agregación se requiere con fines de puntaje, cuando se evalúan sistemas de software en la que intervienen gran cantidad de elementos y componentes. Cabe recordar que para cada variable medida X_i , se producía una preferencia elemental IE_i por medio de un criterio elemental. Esta correspondencia discutida en el capítulo 6, unifica las distintas unidades de los atributos, a la unidad que representa la preferencia de la calidad elemental, definida en el rango unitario. Este resultado final, elemental, se puede interpretar como el grado -o mediante una transformación escalar, el porcentaje- del requerimiento del usuario satisfecho para un rango de valores del atributo A_i .

En el capítulo 7, se presentó y ejemplificó la fase respectiva de Web-site QEM, a partir del modelo de agregación de atributos y características denominado LSP. En dicho modelo, se emplea preferencias de calidad elementales como entradas, para producir una preferencia de calidad parcial o global como salida, por medio de un operador apropiado. Como se discutió previamente, la preferencia de calidad representa el porcentaje del requerimiento de calidad satisfecho; en este caso, se observa la validez

de las unidades entre entradas y salida, debido a que están uniformadas. Para el modelo LSP -ver fórmula (3) en sección 7.2.1.2-, la regla de combinación es aditiva para $r = 1$, para $r > 0$ es supra-aditiva y para $r < 0$ es sub-aditiva. Por lo tanto, supra y sub-aditividad son propiedades significativas para el tipo de escala proporción no-aditiva [Zuse 98] pp.222. La regla de combinación:

$$m(x1 \text{ o } x2) = (m(x1)^r + m(x2)^r)^{1/r}$$

es significativa para el tipo de escala proporción.

10.2.4 Estrategias para Validación de Métricas

Básicamente, hay dos estrategias para corroborar o falsear la validez de las métricas: la teórica y la empírica. A su vez, dos enfoques comúnmente empleados para validación se corresponden con los atributos internos de los entes, en donde se dice que una medida con fines de evaluación es válida internamente, o “*válida en un sentido estrecho*”, o, por otra parte, con las características externas de más alto nivel (por ejemplo, costo, calidad), en donde se dice que una medida es válida externamente, o “*válida en un sentido amplio*”. En esta última categoría entran las métricas empleadas en los modelos predictivos.

La validación teórica permite confirmar que la medida no viola las propiedades de los sistemas relacionales empíricos y numéricos y los modelos de definición de la medición introducidos anteriormente.

La validación empírica permite corroborar una medida por medio de la observación y la planificación de experimentos y surveys, para ver por ejemplo, si los usuarios concuerdan con la existencia de algún atributo, o si el mapeo del mundo real al modelo mental empírico es una representación adecuada del atributo, o si el tipo de escala es el adecuado en el contexto del proyecto de evaluación (entre otros aspectos, como la determinación de si una métrica de un atributo interno puede ser usada para predecir el valor de una variable dependiente -característica externa).

[Kitchenham et al 96] asumen que para que una medida sea válida se debe mantener estas dos condiciones: 1) que la medida no viole cualquier propiedad necesaria de sus elementos; 2) que cada modelo usado en el proceso de medición debe ser válido. Además, de acuerdo al marco conceptual para validación de métricas propuesto por los autores para decidir si una métrica es válida, los mismos afirman que es necesario al menos confirmar:

- a) *La validez del atributo*: esto es, si el atributo en cuestión es realmente exhibido por el ente que se desea evaluar.

- b) *La validez de la unidad*: esto es, si la unidad de medición a ser usada es apropiada para medir al atributo.
- c) *La validez del instrumento*: esto es, si el modelo que subyace al instrumento de medición es válido y el mismo está propiamente calibrado.
- d) *La validez del protocolo*: esto es, si se ha adoptado un protocolo aceptable para la medición de modo que sea repetible y replicable.

Por lo tanto, desde el punto de vista de la validación teórica, habría que probar para cada atributo los puntos anteriores, y, particularmente, para la validación de un atributo y su unidad, tener en cuenta los criterios expresados en la sección 10.2.2.

En la tabla 10.2, se muestra descripciones de la validez teórica para un conjunto de métricas presentadas en los casos de estudio desarrollados. Por lo tanto, se hace referencia a descripciones y cumplimiento de propiedades específicas al caso.

El ente de los atributos evaluados es un sitio Web (artefacto), en la fase operativa del ciclo de vida del producto. La validez de instrumento no es aplicable cuando el tipo de recolección de datos fue realizado manualmente. En el caso en que fue empleado un instrumento, como por ejemplo en el atributo 4.1.1 (ver figura 5.6), los evaluadores confirmaron la validez de las unidades y parámetros que la herramienta permitía establecer. Además, para cada medición se repitió varias veces la misma para observar el comportamiento de la herramienta en cuanto a la tolerancia de error.

Podemos demostrar que para los dos últimos atributos de la Tabla 10.2, el tipo de escala es absoluto.

Sea la métrica $m = a A/B$ (8)

en donde a es la constante 100; A , por ejemplo, representa la cantidad de enlaces rotos encontrados (conteo), y B la cantidad total de enlaces del sitio (conteo).

Siempre se cumple que $A \leq B$, y por lo tanto se mantiene que $A \subseteq B$. La relación entre A y B puede ser descripta mediante la siguiente fórmula:

$A = b B$ con $b > 0$. (9)

Reemplazando (9) en (8), obtenemos:

$m = a b B/B = a.b = c$.

Por lo tanto, $m = c$, es la transformación admisible para el tipo de escala absoluto.

Tabla 10.2 Estrategia de validación teórica y criterios para validar métricas (se muestran algunos ejemplos distintivos)

Atributo	Escala	Unidad	Criterios y Propiedades aplicadas
(2.1.1.1) Búsqueda Restringida para Colecciones –fig. 9.1 (Idem, para 2.1.1.1.1 a 2.1.1.1.3 de la fig. 5.7)	Ordinal	Ranquin de funcionalidad para 2.1.1.1	<ul style="list-style-type: none"> ✓ Métrica directa. El atributo es exhibido en el sitio. ✓ Cumple con la condición de representación (conforme a lo visto en 10.1), es decir, cumple las propiedades de transitividad y completitud ✓ Diferentes sitios pueden tener diferente funcionalidad para el atributo ✓ Diferentes sitios pueden tener igual funcionalidad para el atributo ✓ La unidad y escala está definida y es corroborada
(1.1.1.2) Tabla de Contenido –fig. 5.7	Ordinal	Ranquin de disponibilidad para 1.1.1.2	<ul style="list-style-type: none"> ✓ Métrica directa. El atributo es exhibido en el sitio Web. ✓ Cumple con la condición de representación (conforme a lo visto en 10.1), es decir, cumple las propiedades de transitividad y completitud ✓ Diferentes sitios pueden tener diferente disponibilidad del atributo ✓ Diferentes sitios pueden tener igual disponibilidad del atributo ✓ La unidad y escala está definida y es apropiada
(3.1.1.1) Enlaces Rotos	Absoluto	% de enlaces rotos para 3.1.1.1	<ul style="list-style-type: none"> ✓ Métrica indirecta (la ecuación es $X1 = BL * 100/TL$; BL representa la cantidad de enlaces rotos encontrados; TL representa la cantidad total de enlaces del sitio. ✓ Cumple con la condición de representación (mayor cantidad de enlaces rotos con respecto al total de enlaces del sitio conducen a mayor grado de deficiencia para el atributo) ✓ Diferentes sitios pueden tener diferentes grados de enlaces rotos para 3.1.1.1 ✓ Diferentes sitios pueden tener igual % de enlaces rotos para el atributo 3.1.1.1 ✓ La unidad y escala está definida y es apropiada para las variables y para la ecuación. El conteo de cada variable es de tipo de escala absoluto. Produce un tipo de escala final absoluto
(4.2.1.2.1) Imagen con Título	Absoluto	% de ausencia de la propiedad ALT (en HTML) para 4.2.1.2.1	<ul style="list-style-type: none"> ✓ Métrica indirecta (la ecuación $X1 = AAR * 100/TAR$; AAR representa la cantidad de objetos sin título en la imagen; TAR representa la cantidad total de objetos (imágenes) que deben referenciar la propiedad ALT. ✓ Cumple con la condición de representación (mayor ausencia de título de imagen en los objetos gráficos conduce a menor accesibilidad en la lectura, esto es, cuando se desactiva la característica de “ver imagen” en el browser) ✓ El conteo de cada variable es un tipo de escala absoluto. Produce un tipo de escala final absoluto

Finalmente, mediante la validación empírica podemos realizar experimentos para ver si un grupo de personas acuerdan sobre la existencia de un atributo, o si el mapeo del mundo real al modelo mental empírico es una representación adecuada del atributo, entre otros aspectos.

Un conjunto de estadísticas no-paramétricas se puede emplear en este caso. Por ejemplo, el coeficiente de concordancia W de Kendall mide el grado de acuerdo entre k evaluadores ($k \geq 2$) en el ordenamiento de un conjunto de entidades con respecto a un atributo; o se puede emplear la técnica de correlación de Kendall entre varios evaluadores y un criterio de ranquin; entonces se puede observar cuan bien se compara el orden de ranquin producido por la medida respecto del ranquin producido por expertos humanos. Este tipo de estudios e investigaciones sobre modelos predictivos de esfuerzo y costo (a partir de métricas de tamaño de páginas, cantidad de páginas y enlaces) son parte de nuestras líneas futuras de trabajo.

Capítulo 11

Modelo de Proceso de Evaluación de Artefactos de Software.

El objetivo central de esta tesis fue la de discutir una metodología que cubriera las principales actividades y métodos del proceso de evaluación y comparación de calidad de artefactos Web. De este modo, presentamos un conjunto bien definido y cooperante de estrategias, métodos, modelos, procedimientos y herramientas que, aplicados sistemáticamente a las principales actividades del proceso, produjeran los documentos conteniendo indicadores de calidad globales, parciales y elementales, útiles para efectuar análisis y recomendaciones.

Particularmente, en el capítulo 4, presentamos las principales fases del proceso de evaluación, y en los capítulos subsiguientes fuimos desarrollando distintos aspectos de la metodología, en consideración de las mismas; sin embargo, el foco no estuvo puesto en discutir un modelo de proceso de evaluación con el objeto de especificar todas las actividades, sus entradas, sus salidas, los productos generados, el comportamiento dinámico y los puntos de control, entre otros aspectos.

Por lo tanto, si bien el objetivo de este estudio no es postular y formalizar en principio un modelo de proceso de evaluación particular, no obstante queremos describir el estado actual de las investigaciones realizadas en el área por ISO/IEC, y cómo el proceso implícitamente empleado en Web-site QEM puede adherir a las mismas (en la medida en que estamos tratando de un producto o resultado surgido del consenso de muchos grupos de trabajo y reuniones técnicas).

Específicamente, de interés para nuestro estudio es el recientemente editado estándar del proceso para evaluadores de productos de software [ISO/IEC 14598-5]. Este proceso representa principalmente a la vista funcional y de información de un modelo de proceso [Curtis et al 92, Olsina 98a], y se apoya en el modelo de proceso definido sucintamente en el estándar ISO/IEC 9126, que además, prescribe características de calidad de productos. Asimismo, vale destacar que el proceso para evaluadores ISO/IEC 14598, es aplicable dentro del proceso de ciclo de vida de desarrollo definido en ISO/IEC 12207 [ISO/IEC 12207].

Los modelos de proceso para evaluación de la ISO que pasaremos a describir, son en su especificación original suficientemente genéricos y abstractos, por lo tanto, personalizables a las distintas necesidades del proceso de evaluación de calidad de artefactos de software y artefactos Web. Los modelos de proceso de evaluación no prescriben ni recomiendan métodos ni procedimientos específicos para realizar las

actividades sino que representan un marco conceptual genérico, en donde las mismas se puedan aplicar.

11.1 Modelo de Proceso de Evaluación definido en el Estándar ISO 9126.

La figura 11.1 muestra el esquema del modelo de proceso de evaluación, extraído del estándar ISO ([ISO/IEC 9126] pp. 6-8). Este modelo de proceso de evaluación se puede aplicar en la fase apropiada del ciclo de vida de un producto de software.

En la parte derecha de la figura, se aprecian las tres fases o etapas fundamentales del modelo, a saber:

- ✓ *Definición de Requerimientos de Calidad*
- ✓ *Preparación de la Evaluación*
- ✓ *Procedimiento de Evaluación*

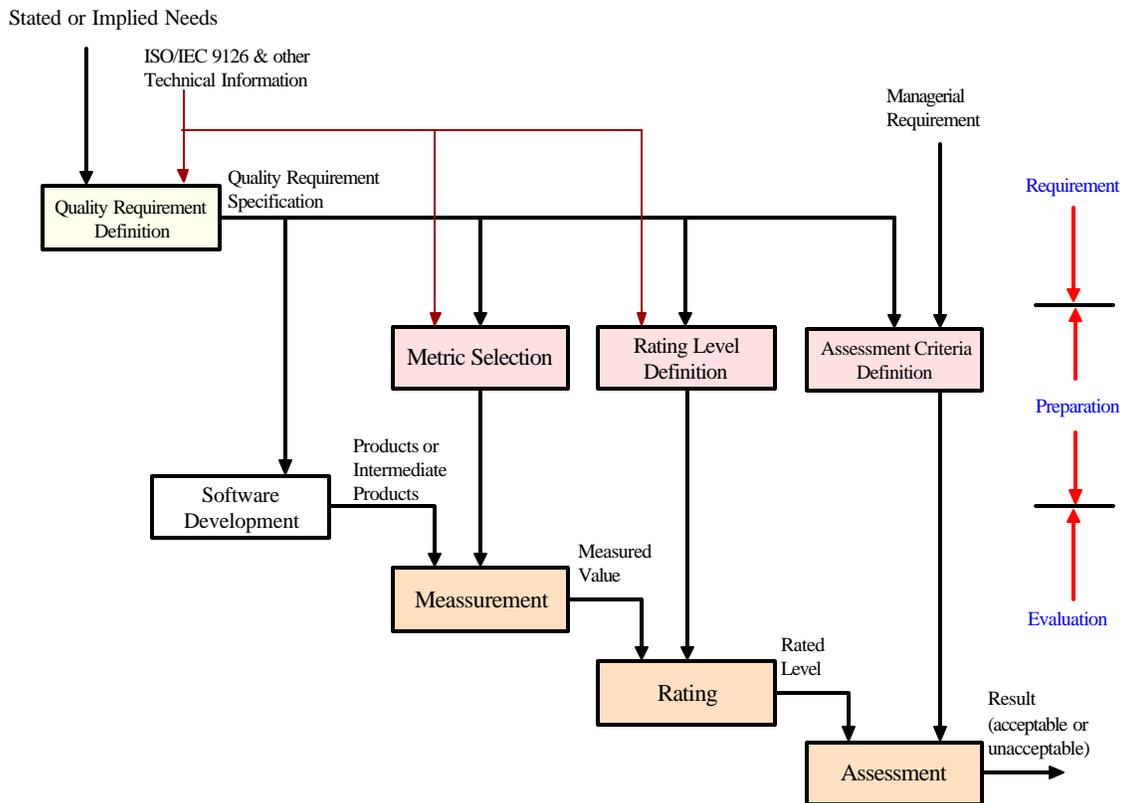


Figura 11.1 Modelo de Proceso de Evaluación (tomado de [ISO/IEC 9126])

En la fase de *Definición de Requerimientos de Calidad* se especifican los requerimientos en términos de las características (y posible subcaracterísticas) prescritas en este estándar. Las mismas sirven de entrada al proceso, al igual que todas

las necesidades explícitas e implícitas del usuario, y produce como salida una especificación de requerimientos de calidad. Los requerimientos expresan la demanda del contexto para el producto de software bajo evaluación, y se deben definir antes del desarrollo.

En la fase de *Preparación de la Evaluación*¹ se especifican tres pasos generales, a saber: 1) *Selección de las Métricas de Calidad*; 2) *Definición de los Niveles de Puntaje*; y 3) *Definición de los Criterios de Valoración*.

En el primer paso (o actividad) los evaluadores deben decidir y seleccionar a las métricas directas (o indirectas) debido a que las características definidas en el estándar (ver sección 5.4.2), son de muy alto nivel. Las métricas seleccionadas, las cuales deben correlacionarse con las características respectivas, difieren del contexto y de la fase del proceso de desarrollo.

En el segundo paso, *Definición de los Niveles de Puntaje*, se debe definir los niveles de satisfacción. Esto es necesario debido a que el valor de una métrica, es mapeado en un rango de una escala, que no expresa el nivel de satisfacción. Por lo tanto, el estándar indica que se deben definir grados (rangos de valores) de satisfacción de los requerimientos, y que son los mostrados en la figura 5.8². El estándar indica que debido a que la calidad se refiere a necesidades dadas y específicas, ningún nivel de puntaje general es posible sino que se debe definir con respecto a cada evaluación en particular.

En el tercer paso (*Definición de los Criterios de Valoración*) de la fase *Preparación de la Evaluación*, los evaluadores deben definir procedimientos para resumir los resultados de las diferentes características. Por ejemplo, dice el estándar, se pueden usar tablas de decisión o promedios pesados. El procedimiento puede incluir otros aspectos como tiempo y costos, en un contexto particular (ver la entrada “managerial requirements”).

En la fase de *Procedimiento de Evaluación*³ se especifican tres pasos generales, a saber: 1) *Medición*; 2) *Puntaje*, y 3) *Valoración*.

En la actividad de *Medición*, se aplican las métricas seleccionadas al producto de software, obteniendo como salida un valor en la escala de medición (ver fig. 5.8).

¹ Como se observa en la figura 4.1, nosotros definimos (preparamos), a la Evaluación Elemental y a la Evaluación Global como actividades separadas.

² Nuestra estrategia presentada en el capítulo 5 y discutida en 6 es semejante, debido a que el valor de una métrica no define la preferencia elemental de calidad del atributo, respecto del punto de vista de usuario seleccionado. En nuestro caso, a diferencia de la sugerida en ISO, la escala de preferencia es fija y está definida en el intervalo unitario de los números reales. Entre esos valores de preferencia definimos tres niveles de aceptabilidad, según lo discutimos en dichos capítulos.

³ Como se observa en la figura 4.1, nosotros implementamos a la Evaluación Elemental y a la Evaluación Global como procesos separados.

En la actividad de *Puntaje*, se determina el nivel de puntaje alcanzado (grado de satisfacción) a partir del valor determinado en la actividad previa.

La *Valoración* es la última actividad del modelo de proceso de evaluación de ISO 9126. El resultado es una declaración de la calidad del producto de software. De este modo, la calidad resultante puede ser comparada con otros aspectos como costo (según los criterios), de manera que la decisión gerencial será la que determinará la aceptación o rechazo, o la distribución de la versión del producto de software, o la decisión que corresponda.

Lo descrito anteriormente representa a todas las etapas y actividades especificadas en el modelo de proceso de evaluación del estándar ISO 9126. Por ejemplo, el modelo no presenta la fase de Planificación, la fase de Análisis y Conclusiones, ni la fase de Validación.

11.2 Proceso para Evaluadores definido en el Estándar ISO/IEC 14598.

Como indicamos en la introducción de este capítulo, es de interés para nuestro estudio el estándar editado recientemente respecto al proceso para evaluadores de productos de software, descrito en el documento [ISO/IEC 14598-5]. Este proceso de evaluación genérico representa a un conjunto de subprocesos, sus entradas y salidas, y se apoya en el modelo de calidad definido en el estándar ISO/IEC 9126. Sin embargo, no prescribe métodos y procedimientos específicos para realizar las tareas, sino que es responsabilidad del evaluador el seleccionar los más apropiados para un proyecto de evaluación dado.

El estándar define los subprocesos necesarios para analizar los requerimientos de evaluación, para especificarlos, diseñarlos (planificarlos), ejecutar las acciones de evaluación, y obtener conclusiones (recomendaciones) para cualquier tipo de artefacto de software. El mismo se puede usar para: 1) evaluar productos existentes, o 2) para evaluar productos en desarrollo (en este caso, el proceso de evaluación debe sincronizarse con el proceso de desarrollo).

Potencialmente, identifica dos partes involucradas en el proceso de evaluación de un producto de software (o artefacto Web), a saber: el solicitante y el evaluador. El primer rol, el de solicitante, puede ser jugado por un desarrollador, un usuario del software, un proveedor o adquirente de software; y el segundo rol, el de evaluador, puede ser asignado, por ejemplo, a un laboratorio u organización destinado a evaluar software, un laboratorio que realiza comparaciones entre productos, entre otros. Por lo tanto, el proceso de evaluación consiste de un conjunto de tareas cooperativas e interacciones

entre el solicitante y el evaluador. Este proceso se realiza sobre la base de los datos, información y estándares provistos por ambas partes, o producidas por otras actividades. Como consecuencia de realizar los subprocesos, se produce artefactos (documentos, especificaciones, etc.) que pueden servir de entrada a otras actividades, o que puede ser el resultado del proceso de evaluación (informe y registros de mediciones, informe con recomendaciones).

El proceso de evaluación según el estándar ISO/IEC 14598, comprende a cinco subprocesos, con sus respectivas entradas y salidas, como se aprecia en la figura 11.2. Los subprocesos son los siguientes:

- ✓ *Establecimiento de los Requerimientos de Evaluación*
- ✓ *Especificación de la Evaluación*
- ✓ *Diseño de la Evaluación*
- ✓ *Ejecución de la Evaluación, y*
- ✓ *Conclusión de la Evaluación*

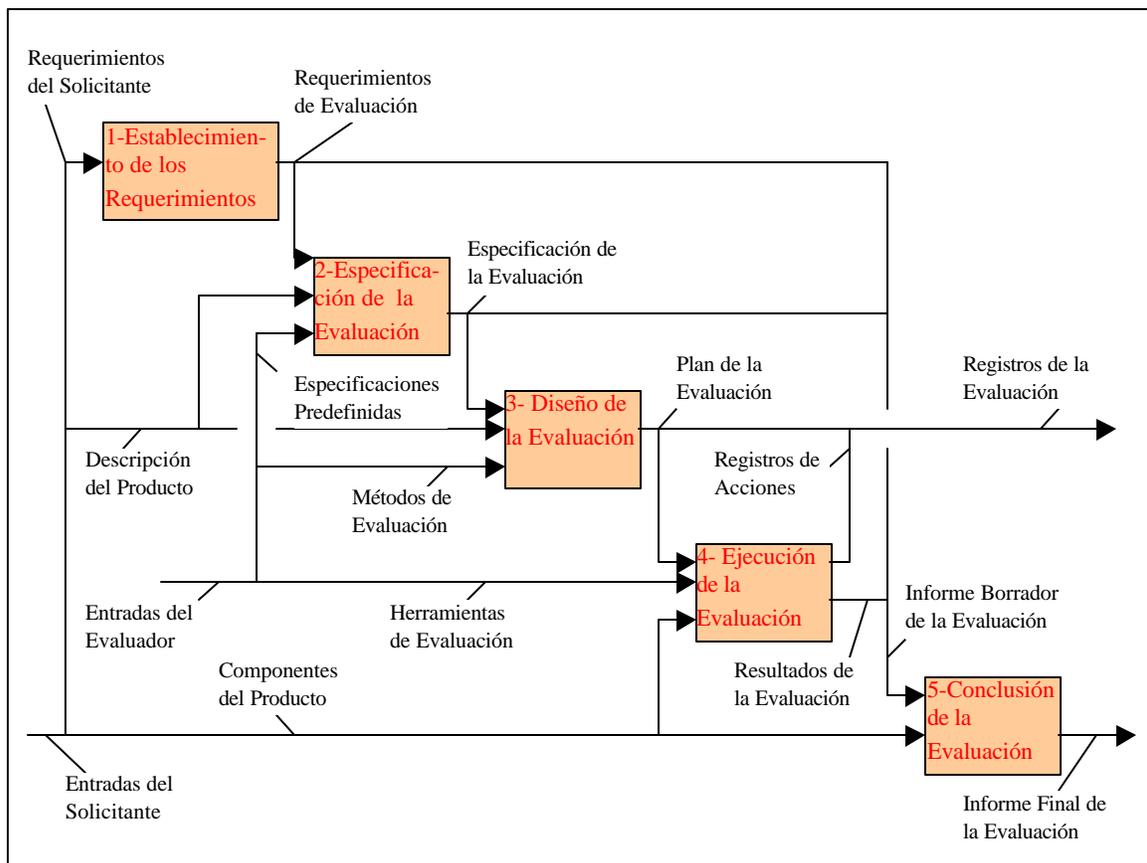


Figura 11.2 *Proceso para Evaluadores* (traducido de [ISO/IEC 14598-5])

En cuanto a las entradas al proceso, el solicitante provee la descripción del producto (y las necesidades), y los componentes del producto. El evaluador potencialmente provee como entradas, especificaciones predefinidas de evaluación, métodos y herramientas de

evaluación.

En cuanto a las salidas al proceso, como se observa en la figura 11.2, hay productos intermedios y productos finales. Entre los primeros se encuentran los documentos de requerimientos, especificación y plan de la evaluación; entre los segundos los registros e informes de evaluación.

El documento de requerimientos describe la meta de la evaluación, el punto de vista y los requerimientos de calidad para el artefacto seleccionado (y descripto). El documento de especificación define los análisis y criterios e identifica las mediciones a realizarse sobre el producto y sus componentes. El plan de la evaluación describe los métodos y procedimientos a implementar. Los registros de evaluación consisten en las acciones específicas implementadas conforme al plan. El informe de evaluación contiene los requerimientos, las especificaciones y resultados de las métricas y análisis realizados, y toda información de manera que el proceso sea repetible y reproducible. Este documento se produce primero como un borrador para revisión y luego, con las recomendaciones finales, se distribuye al solicitante. En cuanto a la plantilla para este último documento, el estándar ISO/IEC 14598 prescribe su estructura y contenido (en el anexo normativo A, pp. 20-21).

Con el fin de satisfacer resultados de un proceso de evaluación repetibles, reproducibles, imparciales y objetivos, el evaluador debe actuar en un contexto organizacional que le provea la necesaria garantía y recursos para obtener calidad en sus actividades. Para satisfacer estas necesidades, el estándar prescribe guías (un marco prescriptivo) que establecen las responsabilidades del solicitante y del evaluador. Entre otras, se establecen responsabilidades contractuales de confidencialidad, responsabilidades técnicas y tecnológicas, legales, de provisión de recursos e información en tiempo y forma.

11.2.1 Desarrollo del Proceso de Evaluación

En esta sección desarrollaremos, conforme al estándar, los cinco procesos antes mencionados, el objetivo de cada uno, los subprocesos, un resumen de la descripción del contenido de los documentos y los puntos de control. Para mayor simplicidad, estructuramos la información para cada proceso como sigue⁴:

Nombre del Proceso: Establecimiento de los Requerimientos de Evaluación **Código:** 1 (según la figura 11.2)

Objetivo: el propósito de este proceso es describir la meta y objetivos de la evaluación. Tales objetivos se relacionan con el uso del artefacto en consideración de uno o varios

⁴ La estructuración de la información de este modo no implica prescribir una plantilla de proceso con sus principales elementos como en [Olsina 98b], sino que la intención es volcar las indicaciones más relevantes del estándar en la misma.

puntos de vista de usuario y los riesgos asociados (es decir, los requerimientos de evaluación pueden especificar niveles de evaluación para las características seleccionadas –los niveles se detallan en el anexo informativo B del estándar). El dominio de la aplicación del producto a evaluar debe ser considerado; aspectos críticos como seguridad, económicos, legales o de contexto deben ser tomados en cuenta.

Nombre de los Subprocesos:

1.1 *Proposición de los requerimientos por parte del solicitante*

1.2 *Declaración del grado de cobertura en la evaluación por parte del solicitante*

1.3 *Soporte del solicitante en analizar el objetivo de la evaluación y en describir los requerimientos con el evaluador*

1.4 *Explicación del grado de confianza y rigor de la evaluación al evaluador*

1.5 *Acordar los requerimientos de evaluación*

Comentarios: El solicitante debe proveer como punto de partida, los requerimientos iniciales. En los mismos se debe expresar cuan extensiva debe ser la cobertura o alcance de la evaluación. Por otra parte, el evaluador debe asegurar el rigor necesario del proceso de evaluación para determinar la calidad del producto. Por lo tanto, ambas partes deben acordar sobre los requerimientos como un prerequisite para la continuación del proceso.

Descripción del Contenido de la Salida: El documento de requerimientos de evaluación debe contener una descripción del dominio de la aplicación del producto sometido y una descripción general del propósito del producto.

El documento de requerimientos contendrá asimismo una lista de los requerimientos de calidad, referidas por ejemplo, a las prescritas en el estándar ISO/IEC 9126; en este contexto, se pueden emplear también las subcaracterísticas. Cuando cierto requerimiento se refiere a una característica no definida en el citado estándar, se deberá referenciar a la literatura autorizada que la define, y ambas partes declararán explícitamente la comprensión y utilización de la misma.

Por otra parte, se deberá acordar y expresar en el documento la importancia relativa de cada característica. Además, se deberá proveer para cada requerimiento la especificación de la información contenida en el producto y los componentes a ser evaluados (el nivel y forma de la información requerida en el documento puede estar relacionada al costo de la evaluación, o a la importancia específica de un requerimiento de calidad).

Otros Aspectos (de Control, etc.): El documento de requerimientos de evaluación deberá ser aprobado en revisión conjunta por el solicitante y el evaluador. Este documento se incluirá en los registros de evaluación y en el informe final de evaluación (ver figura 11.2)

Nombre del Proceso: Especificación de la Evaluación Código: 2

Objetivo: el propósito de este proceso consiste en definir el alcance de la evaluación y las mediciones a realizarse en el artefacto a evaluar y sus componentes. El nivel de detalle de la salida (el documento de especificación de la evaluación) debe ser tal de modo que se asegure la repetitividad y reproducibilidad del proceso.

Nombre de los Subprocesos:

2.1 Analizando la descripción del producto

2.2 Especificando las mediciones

2.3 Verificando las especificaciones producidas en consideración con los requerimientos

Descripción de los Subprocesos:

2.1 Analizando la descripción del producto: El solicitante debe proveer una descripción del producto a ser sometido a evaluación. Esta descripción puede permitir definir el alcance de la evaluación (es decir, puede permitir identificar qué componentes son partes del producto y cuáles no). Definir el alcance de la evaluación es importante cuando el producto a evaluar está embebido en un sistema que puede consistir de hardware, otros productos de software, redes, etc. y no siempre es tan obvio definir los límites. Por otra parte, analizar la descripción del producto y sus componentes, permitirá al evaluador comprender su estructura, funcionalidad y relaciones entre las partes.

Esta descripción debe contener la lista de componentes del producto a evaluar y, en lo posible, referencias a documentos relacionados.

2.2 Especificando las mediciones: El evaluador debe asignar los requerimientos de evaluación al producto y sus componentes identificados en la descripción del producto. Esto debe conducir a una descomposición de los requerimientos de evaluación, por ejemplo, en características y subcaracterísticas. El resultado de la descomposición puede ser diferente para los diferentes componentes sometidos. En consecuencia, el evaluador especificará las distintas métricas destinadas a valorar las características, subcaracterísticas y atributos. Estas especificaciones pueden contener algunas de estas declaraciones:

- una especificación formalizada de una métrica a ser aplicada, junto con las instrucciones de presentación de la misma en el informe de evaluación
- una referencia a la especificación al requerimiento correspondiente que deberá ser verificado, como así también el procedimiento de verificación del mismo
- la especificación de un requerimiento que estaba ausente en el documento o que requiere mayor nivel de detalle y explicación, como así también el procedimiento de verificación del mismo
- una referencia a declaraciones de estándares o normativas en donde se provee información adicional del requerimiento

Para esta tarea el evaluador puede usar especificaciones de evaluación predefinidas.

2.3 Verificando las especificaciones producidas en consideración con los requerimientos: El evaluador debe realizar una verificación de la especificación de la evaluación con respecto a los requerimientos de evaluación. Se debe garantizar que las medidas especificadas son suficientes para alcanzar los objetivos del proceso como declarado en los requerimientos.

Descripción del Contenido de la Salida: El documento de especificación de la evaluación debe contener:

- El alcance de la evaluación referenciando a los componentes del producto tal como estaban identificados en la descripción del mismo.
- Una especificación de las mediciones y verificaciones a ser realizadas y las referencias respectivas a los componentes del producto.
- Una correspondencia entre la especificación de las mediciones y verificaciones, y el documento de especificación de requerimientos (junto con las referencias a documentos, estándares, etc., o justificaciones para cada medida y verificación).

Otros Aspectos (de Control, etc.): El documento de especificación de la evaluación deberá ser aprobado en revisión conjunta por el solicitante y el evaluador. Este documento se incluirá en los registros de evaluación y en el informe final de evaluación (ver figura 11.2). Cualquier cambio al documento de requerimientos surgido de alguna de las actividades de este proceso, será informado en los registros de evaluación.

Nombre del Proceso: Diseño de la Evaluación Código: 3

Objetivo: el propósito de este proceso consiste en documentar los métodos y procedimientos a utilizar por el evaluador para realizar las mediciones y verificaciones contenidas en el documento de especificación de la evaluación. El evaluador producirá como resultado de este proceso el plan de la evaluación que describe los recursos necesarios (humanos, materiales, tecnológicos, etc.) y la distribución y asignación de los mismos a las actividades.

Nombre de los Subprocesos:

3.1 Documentando los métodos y procedimientos de evaluación y produciendo un plan

3.2 Programando las actividades conforme a los recursos disponibles

Descripción de los Subprocesos:

3.1 Documentando los métodos y procedimientos de evaluación y produciendo un borrador del plan: El objetivo de esta actividad es combinar las diferentes métricas y verificaciones con los distintos componentes del producto con el fin de documentar detalladamente los métodos y procedimientos a ser aplicados para implementar dichas mediciones y verificaciones sobre los componentes y sus elementos. El evaluador debe analizar restricciones técnicas como:

- los formalismos usados para los componentes del producto
- el hecho de que los componentes a evaluar sean presentados en formato digital o en papel
- la existencia de métodos de evaluación predefinidos
- la disponibilidad de herramientas que soporten al método o procedimientos específicos
- el tamaño de los componentes del producto

El evaluador debe documentar en el plan, para cada métrica y verificación especificada, el método apropiado (como así también, cuando corresponda, la herramienta a emplear, indicando al menos el nombre, la versión y su origen).

Luego se debe optimizar el plan con el fin de remover las duplicaciones al asignar los métodos y procedimientos a los distintos elementos de los componentes del producto que utilizan las mismas técnicas de evaluación.

3.2 *Programando las actividades conforme a los recursos disponibles:* El evaluador debe tomar en cuenta la disponibilidad de recursos para programar las actividades. Además, debe acordar con el solicitante, la fecha de distribución de los resultados, el formato de los mismos, como así también, por otra parte, los requerimientos para las reuniones durante el curso de la evaluación.

Descripción del Contenido de la Salida: El documento del plan de la evaluación está compuesto de dos partes: 1) la documentación de los métodos de evaluación, y 2) la programación respectiva

Otros Aspectos (de Control, etc.): El plan de la evaluación deberá ser aprobado en revisión conjunta por el solicitante y el evaluador. Este documento se incluirá en los registros de evaluación y la documentación de los métodos de evaluación o referencias a los mismos se incluirán en el informe final de evaluación (ver figura 11.2).

Nombre del Proceso: Ejecución de la Evaluación **Código:** 4

Objetivo: el propósito de este proceso es obtener los resultados al realizar todas las acciones para medir y verificar el producto conforme a los requerimientos de evaluación, según lo especificado y planeado. Al final del proceso se completan los registros de evaluación y el borrador del informe de evaluación.

Nombre de los Subprocesos:

4.1 *Administrando los componentes del producto provistos por el solicitante*

4.2 *Administrando los datos producidos por la evaluación (incluyendo registros e informes)*

4.3 *Administrando las herramientas necesarias por la evaluación*

4.4 *Administrando las acciones de evaluación fuera del sitio acordado*

4.5 *Administrando los requerimientos surgidos por el uso de técnicas específicas*

Descripción de los Subprocesos:

4.1 *Administrando los componentes del producto provistos por el solicitante:* El solicitante debe distribuir al evaluador los componentes de los productos y documentos relacionados, conforme a lo programado. La confidencialidad de todos los componentes de los productos y documentos relacionados deben ser protegidos de acuerdo a lo acordado.

4.2 *Administrando los datos producidos por la evaluación (incluyendo registros e informes):* Realizar el proceso de evaluación consiste generalmente en medir los atributos y características de los componentes de los productos, para obtener datos e interpretación de los mismos con el fin de incluirlos en el informe de evaluación. Los datos intermedios y finales se deberán proteger del mismo modo que los componentes de los productos conforme a lo acordado. Los datos y sus interpretaciones deberán incluirse en los registros de evaluación.

4.3 *Administrando las herramientas necesarias por la evaluación:* Al realizar el proceso de evaluación se podría necesitar herramientas para recolectar datos, o para realizar la interpretación de los mismos. El evaluador debe documentar en el informe de evaluación, la herramienta empleada, indicando al menos el nombre, la versión y su origen. Además, se debe registrar las acciones realizadas para la

validación del instrumento. Finalmente, si fuera necesario, el personal de evaluación deberá ser entrenado en la herramienta.

4.4 Administrando las acciones de evaluación fuera del sitio acordado: Algunas veces, las acciones de evaluación no se podrán llevar a cabo en el sitio acordado. Por ejemplo, se podría realizar en el lugar donde trabajan los desarrolladores, o donde el artefacto está en operación. En estos casos el evaluador deberá asegurar la confidencialidad, y evitar circunstancias que invaliden al proceso.

Descripción del Contenido de la Salida: Las salidas de este proceso son dos documentos: 1) los registros de evaluación, y 2) un borrador del informe de evaluación. Durante la ejecución de la evaluación, se producen resultados intermedios y finales. Para lograr un máximo de objetividad de las acciones, éstas deben ser revisadas por personal de evaluación que no haya participado consuetudinariamente en las mismas. Todos los resultados de la evaluación deben ser revisados. En la revisión debe participar al menos una persona no involucrada directamente en el proceso. El informe de revisión deberá incluirse en los registros de evaluación. Una vez revisados, los resultados de la evaluación se deberán incluir en el borrador del informe de evaluación.

Nombre del Proceso: *Conclusión de la Evaluación* **Código:** 5

Objetivo: el propósito de este proceso consiste en la revisión del borrador entre las partes (solicitante y evaluador) y en poner a disponibilidad los documentos finales.

Nombre de los Subprocesos:

5.1 Revisión conjunta del informe de evaluación

Descripción de los Subprocesos:

5.1 Revisión conjunta del informe de evaluación: El borrador del informe de evaluación debe ser distribuido al solicitante. Luego se debe organizar una reunión de revisión conjunta. El solicitante debe tener la oportunidad de realizar comentarios sobre el informe. En el caso de realizarlos, se deberá incluir dichos comentarios en un capítulo separado del informe final de evaluación. Finalmente, el documento se distribuirá al solicitante.

Otros Aspectos (de Control, etc.): Una vez que el documento final se distribuyó formalmente al solicitante, el evaluador deberá deshacerse de los datos correspondientes a la evaluación. Esto se deberá hacer, dependiendo del tipo de datos, de alguna de estas formas:

- los documentos y artefactos sometidos a evaluación se deberán devolver al solicitante o se deberán archivar por un período de tiempo acordado, o se deberán destruir en un lugar seguro
- los registros de evaluación, y el informe de evaluación se deberán archivar por un período de tiempo acordado
- otros datos cualesquiera, se deberán archivar por un período de tiempo acordado, o se deberán destruir en un lugar seguro

Cuando el período de archivado expire para algún dato, se deberán archivar otra vez por un período de tiempo acordado, o se deberán destruir en un lugar seguro.

En caso en que el solicitante acuerde, los resultados de los datos intermedios podrán ser usados por el evaluador con el fin de estudiar técnicas de evaluación y métricas de software.

11.3 Consideraciones Finales.

La descripción del estándar ISO/IEC 14598 realizada en la sección anterior, tuvo como objetivo mostrar el espíritu impreso al mismo, no siendo una transcripción (ni traducción) total y literal del estándar, aunque refleja los aspectos más relevantes del proceso para evaluadores.

Como indicamos en la introducción, los modelos de proceso para evaluación de la ISO descriptos, son en su especificación original suficientemente genéricos y abstractos, y por lo tanto, personalizables a las distintas necesidades del proceso de evaluación de calidad de artefactos de software y de artefactos Web (ver la discusión “artefactos Web como software” del capítulo 2).

De los estudios realizados, hemos podido observar que los procesos de evaluación comentados no prescriben ni recomiendan metodologías, métodos ni procedimientos específicos para realizar las actividades sino que representan un marco conceptual (y normativo) genérico, esto es, un modelo en donde distintos métodos, técnicas, procedimientos y herramientas se puedan aplicar.

No obstante, es importante realizar algunas observaciones respecto de los estándares descriptos en las secciones 11.1 y 11.2. Primero, es oportuno remarcar que el estándar ISO/IEC 9126 de características de calidad de productos de software, y que además prescribe un modelo de proceso de evaluación, fue publicado en 1991, en tanto que el documento 14598-5 del estándar ISO/IEC, que trata solamente del proceso para evaluadores de productos de software, fue publicado en julio de 1998. Por lo tanto, si bien no es apropiado realizar comparaciones entre ambos, sí se puede indicar que el estándar ISO/IEC 14598 es una evolución respecto del anterior, en donde la experiencia y madurez ganada a través de los años por la comunidad científica y técnica del área, se ve reflejada en el contenido del producto⁵.

Por ejemplo, en este estándar internacional se observa un espíritu contractual entre las partes involucradas (para los roles de solicitante y evaluador), especificando guías y normativas generales a cumplimentar (en el anexo informativo E, se especifica un contrato de evaluación que consta de dos secciones principales, a saber: el punto E.1,

⁵ El estándar ISO/IEC 14598, consta de seis documentos, los cuales, al momento de esta escritura, solamente se ha publicado el documento 14598-5, estando los restantes en modo borrador (draft). Por otra parte, el nuevo estándar ISO/IEC 9126-1, también en modo draft, no especifica un modelo de proceso de evaluación sino que se ha trasladado el modelo de evaluación de producto software al estándar ISO 14598.

en donde se detalla un modelo de contrato de evaluación, y el punto E.2, en donde se describe un marco legal de términos y condiciones). En el escrito se indican responsabilidades generales asignadas a dichos roles y puntos de control (en el anexo informativo D se detalla, de un modo secuencial, las potenciales interacciones entre el solicitante y el evaluador). Por otra parte, se especifica con mayor profundidad, en un nivel de granularidad mayor, a procesos y subprocesos, y se detalla claramente las entradas y salidas de cada uno de los cinco procesos. Por último, se prescribe una plantilla de documento (en el anexo normativo A del estándar) para el informe final de evaluación.

De manera que, a modo de conclusión, podemos decir que se está desarrollando un marco normativo y conceptual del proceso, robusto, general y flexible, para ser empleado por las organizaciones de evaluación de software, con el objeto de evaluar y certificar a la calidad de artefactos de software (el cual es, en nuestra consideración, aplicable a artefactos Web).

Capítulo 12

Conclusiones y Líneas de Trabajo Futuro.

Como declaramos al inicio de este trabajo, los artefactos y desarrollos Web tienen cada vez más complejidad añadida y, asimismo, están creciendo paulatinamente. Asistimos en nuestros días a una demanda creciente de sitios Web con funcionalidad de aplicación de software. Además, un sitio Web debe satisfacer necesidades tan diversas que pueden ir desde la distribución y presentación de documentos hiperenlazados cuyo contenido puede estar compuesto por diversos medios (texto, sonido, imagen, video), hasta la automatización de procesos de negocios más o menos complejos. A pesar de este crecimiento en la demanda y complejidad, según observamos, modelos de proceso de producto bien definidos que promuevan un ciclo de vida de desarrollo repetible y eficiente, y modelos de proceso y metodologías de evaluación que promuevan la comprensión, el control y la mejora de la calidad de artefactos Web, no están acompañando este rápido proceso. Más bien las prácticas actuales para desarrollar sitios Web son ad hoc, y el aseguramiento y el control de la calidad son, por lo general, procesos poco tenidos en cuenta. Esto motiva y urge a un cambio de paradigma en la prosecución de estos procesos y prácticas para minimizar la crisis en los desarrollos en la Web y, por lo tanto, propender a la calidad y a una efectiva evolución de los artefactos. En esta dirección, el empleo sistemático y disciplinado de procesos, métodos, modelos, y técnicas de Ingeniería de Software para el desarrollo, el mantenimiento, y el aseguramiento de la calidad de los sitios Web, principalmente en los proyectos de mediana o gran escala, promoverá el cambio de paradigma de estrategias ad hoc hacia estrategias más objetivas y justificables.

En este contexto, presentamos en la tesis a la *Metodología de Evaluación de Calidad de Sitios Web* (Web-site Quality Evaluation Method, o Web-site QEM) la cual realiza un aporte ingenieril al sentar las bases y construir las estructuras para el uso de un enfoque sistemático, disciplinado y cuantitativo que se adecue a la evaluación, comparación y análisis de la calidad de aplicaciones centradas en la Web más o menos complejas. Discutimos las fases y actividades principales y los constructores de proceso para realizarlas (capítulos 4 a 8).

Durante la fase de Definición y Especificación de los Requerimientos de Calidad (capítulo 5) se trata con actividades y procedimientos para la elicitación, modelado y especificación de los requerimientos. Particularmente, discutimos distintos modelos de calidad (McCall, Boehm, ISO 9126, IEEE 1061 e ISO 9126-1), y propusimos un enfoque de modelo de calidad mixto, jerárquico y flexible. Este enfoque se emplea para representar y acordar el árbol de requerimientos de calidad de artefactos Web, para una meta, perfil de usuario y dominio dados. En consideración de la naturaleza de los sitios Web (capítulo 2) varios atributos y requerimientos los hacen distintos respecto de los

artefactos de software tradicional. No obstante, las características de calidad de alto nivel de abstracción prescritas en el estándar ISO, son suficientes para modelar y contener las particularidades del dominio de aplicaciones Web. Finalmente, el producto de esta fase es empleado como entrada en la preparación de la evaluación elemental y global (y retroalimenta a actividades de planificación y programación).

Durante la fase de Definición e Implementación de la Evaluación Elemental (capítulo 6), utilizamos criterios de evaluación de calidad elemental para cada atributo cuantificable, escalas, valores y rangos críticos, funciones para determinar la preferencia elemental, entre otros asuntos. Presentamos además, una plantilla de especificación descriptiva para contener a la información esencial de las características y atributos a partir de un árbol de requerimientos. La misma está pensada para ser un documento hiperenlazado, referenciando a otras plantillas, documentos y tablas, favoreciendo así la implementación de un modelo de seguimiento hacia atrás y hacia adelante. Una vez definidos y consensuados los criterios para medir cada atributo, se ejecuta el proceso de recolección de datos, el de cómputo de las variables y las preferencias elementales, y la documentación de los resultados. Los valores elementales obtenidos sirven de entrada al proceso de agregación.

Durante la fase de Definición e Implementación de la Evaluación Global se debe seleccionar un modelo que permita derivar y calcular los valores para las subcaracterísticas y características de más alto nivel de abstracción, a partir de los valores elementales obtenidos para cada atributo del árbol. Discutimos dos enfoques cuantitativos de utilidad en las actividades de estructuración de atributos y cálculo de puntaje para la calidad de cada artefacto. Específicamente, presentamos fortalezas y debilidades de los modelos de puntaje aditivos y lineales, con respecto al modelo de agregación multiatributos no-lineal como lo es LSP (capítulos 3 y 7). Este fue el modelo seleccionado debido a la cantidad de atributos, subcaracterísticas y características intervinientes en los casos de estudio realizados (más de 80 atributos, para el estudio de sitios académicos en la Web), y la necesidad de modelar diferentes tipos de relaciones entre los mismos. En la sección 7.3.1 describimos las actividades a considerar en el proceso de agregación de las preferencias elementales y parciales para producir la preferencia de calidad global (derivada de los requerimientos). Una vez que fueron estructurados y acordados todos los criterios, pesos y operadores de agregación, los evaluadores ejecutan la herramienta que calcula los resultados parciales y finales.

Durante la fase de Análisis de Resultados, Recomendaciones y Documentación los evaluadores realizan actividades de análisis, comparación y justificación de los resultados. En consideración de las metas establecidas y el perfil de usuario seleccionado, el proceso culmina con un documento distribuible conteniendo las conclusiones y recomendaciones del caso. Las salidas de las fases anteriores sirven de entrada y retroalimentación a este proceso. Web-site QEM fue usado hasta el presente, en dos proyectos de evaluación en la fase operativa de sitios Web, y como herramienta

de monitoreo y control de calidad en un proyecto de desarrollo. Actualmente estamos incursionando en el dominio de aplicaciones Web con funcionalidad de comercio electrónico.

Durante la fase de Validación de métricas se debe asegurar que las medidas sean una caracterización numérica apropiada del atributo, mostrando que se satisfaga la condición de representación, es decir, que la correspondencia entre el dominio empírico y el nuevo dominio numérico preserve a la relación funcional. En el capítulo 10, describimos propiedades y un marco conceptual para la validación de métricas y, además, discutimos la validación teórica de algunas métricas empleadas en el proceso.

Por último, en el capítulo 11, describimos detalladamente modelos de proceso de evaluación al que Web-site QEM puede adherir. Específicamente, describimos al modelo de proceso del estándar ISO 9126, y al recientemente editado proceso para evaluadores contenido en el estándar ISO 14598. Estos modelos de proceso de evaluación pueden ser personalizados en actividades de aseguramiento y certificación de calidad. Discutimos que los procesos de evaluación comentados no prescriben ni recomiendan metodologías, métodos ni procedimientos específicos para realizar las actividades sino que representan un marco conceptual (y normativo) genérico en donde distintos métodos, procedimientos y herramientas se puedan aplicar.

Web-site QEM es todavía una metodología de evaluación joven; con todo, el campo de evaluación de artefactos Web es relativamente nuevo y apenas está siendo considerado en la literatura más reciente. El enfoque propuesto, es esencialmente integral, flexible y robusto, y cubre la mayor parte de las actividades en el proceso de evaluación, comparación, y selección de artefactos Web.

Las principales contribuciones de esta tesis para el área de proyectos de evaluación y comparación de calidad de aplicaciones centradas en la Web fueron apuntadas en la sección 1.1. En la siguiente sección, discutimos algunos aspectos que no consideramos en esta investigación y que deberán ser abordados en próximos avances. Algunos representan futuras líneas de investigación.

12. 1 Futuros Avances

12.1.1 Evaluación basada en Preferencias de Calidad y Costo.

En particular, en esta tesis nos hemos concentrado en la característica “calidad” de artefactos Web como característica de estudio. No obstante, otros indicadores como costo, o relación costo-calidad, entre otros, pueden ser de interés para investigaciones en evaluación, comparación y selección de sistemas en general. Por ejemplo, en la evaluación de la compra de una computadora, sea una PC o una computadora portátil,

una persona observa de un modo más o menos intuitivo el valor del costo final que suman todos los componentes requeridos, y, además, la calidad de los mismos (principalmente, podría ser el caso de un profesional del área). Desde un punto de vista más riguroso, y ante la decisión de seleccionar el mejor producto (computadora) con respecto a la relación costo-calidad entre dos o tres sistemas alternativos, se debería emplear un procedimiento que pueda calcular y combinar dichas características en un indicador global.

Muchos artefactos de la realidad se podrían evaluar y seleccionar en función de la relación costo-calidad a partir de un árbol de requerimientos de calidad por un lado, y de requerimientos de costos por otro; además, hay artefactos que sólo pueden requerir una evaluación de costo o de calidad.

En el caso de arribar a un indicador global de calidad IGC_a, y a un indicador global de costos, IGC_o, entonces se podría determinar la relación de calidad y costo, RC&C, para cada artefacto seleccionado, mediante algún modelo matemático. El modelo más simple para obtener un valor, es el siguiente:

$$RC\&C = IGC_a / IGC_o$$

Con el fin de analizar y decidir por “el mejor” artefacto o el “más conveniente”, se podría determinar a partir de ciertas restricciones, un *espacio de aceptabilidad*. Dicho espacio puede estar determinado (en un sistema de ejes cartesianos de dos dimensiones) por el plano formado entre la preferencia de calidad mínima, IGC_{a_min}, a partir del cual el sistema es inaceptable (p.ej. el 60%), y el del costo máximo IGC_{o_max}, a partir del cual, para valores mayores, se deshecha la alternativa. Entre los puntos que caen dentro del plano de aceptabilidad, los evaluadores (o el gerente de sistemas) podrían seleccionar al punto de menor costo, con preferencia de calidad entre el 60 y el 100%. Otros modelos matemáticos, en los cuales intervenga el peso de los indicadores globales de preferencia de calidad y costo, se podrían utilizar.

Esto podría ser de utilidad en la evaluación, comparación y selección de sistemas centrados en la Web o componentes del mismo. Este criterio no era de aplicabilidad para los casos de estudio realizados, en donde la meta era determinar el estado del arte en la calidad de artefactos Web de sistemas, en la fase operativa. No obstante, es nuestra intención la de avanzar en esta línea, analizar modelos e incorporarlo a la metodología, para cuando dicho requerimiento deba ser tomado en cuenta.

12.1.2 Hacia Requerimientos de Calidad Generales para Dominios Web.

Un aspecto de importancia para el reuso (como apuntábamos en la sección 5.5), que si bien hemos avanzado en parte aún es motivo de futura investigación, reside en determinar un árbol de calidad estándar que pueda ser reusado y personalizado para

distintos dominios de aplicaciones, dentro del dominio de la Web. Contar con un árbol de requerimientos de calidad general pero a su vez fácilmente adaptable a diversos dominios, sería de gran beneficio para tareas de aseguramiento y control de calidad posibilitando potencialmente la automatización de las mismas.

Lo anterior, no contradice el empleo del enfoque de modelo de calidad mixto propuesto en la sección 5.4.3, sino que potencialmente lo fortalece. La idea es la de contar con plantillas de modelos de calidad orientado a dominios y perfiles de usuario, y extenderla y consensuarla con los participantes, para cada proyecto de evaluación en particular. Dicha plantilla prescribiría características de alto nivel (las del estándar ISO 9126), subcaracterísticas al segundo nivel, y, en algunos casos, hasta del tercer nivel. Asimismo varios atributos y criterios elementales podrían reusarse. Un trabajo inicial en esta dirección se puede encontrar en [Olsina et al 99d]. No obstante, debemos ganar más experiencia en diferentes dominios de evaluación y perfiles de usuario.

12.1.3 Patrones de Diseño para Dominios Web.

La experiencia ganada hasta el momento, en la determinación de requerimientos de calidad de artefactos Web, ha sido una tarea compensadora, dado que nos ha ayudado a reconocer patrones de diseño, en varios casos, independientes del dominio, como *Esquema de Organización Global del Sitio*, *Visita Guiada*, *Indicador de Última Actualización*, *Retroalimentación Personalizada*, *Estrategias de Búsqueda*, *Carrito de Compras*, por citar algunos. No obstante, este trabajo si bien avanzado, está en proceso de discusión con investigadores del área. Vale destacar que, recientemente, han sido documentados patrones de diseño para la Web como los catalogados en [Rossi et al 99]. Otros patrones, esencialmente para el campo de Hipermedia, se pueden encontrar en [Rossi et al 97].

12.1.4 Validación de Métricas Predictivas para proyectos Web.

Un sistema de medición, y las métricas asociadas, se usan para evaluar un ente (producto, proceso o recurso) existente, al realizar una correspondencia en un dominio numérico de los atributos observables del ente en el mundo empírico, como fue tratado en el capítulo 10. Sin embargo, como indican varios autores. [Fenton et al 97, Zuse 98], un sistema predictivo, y sus métricas asociadas, se utilizan para predecir algún atributo o característica de un ente futuro, involucrando a modelos matemáticos y procedimientos de predicción. Por ejemplo, el bien conocido modelo COCOMO [Boehm 81, Boehm et al 95] dentro de la comunidad de Ingeniería de Software, representa un modelo de predicción de esfuerzo y costos.

Por otra parte, los autores en [Mendez et al 99], presentan estudios iniciales para validar un conjunto de cinco métricas de Hipermedia con fines predictivos, específicamente, con vistas de predecir esfuerzo. Un objetivo esencial para proyectos de Hipermedia y en

la Web sería poder contar en etapas tempranas con un modelo que ayude a predecir el esfuerzo y el costo para producir un artefacto. Esta es una línea abierta de investigación en donde hay mucho por hacer.

12.1.5 Integración del Ambiente WebQEM_Tool.

Tal cual comentamos en el capítulo 6 y 7, estamos desarrollando un ambiente integrado, denominado WebQEM_Tool, para el soporte de automatización de métricas (Web-siteMA), y para la edición, cálculo y documentación en el proceso de agregación de preferencias, sustentado en el modelo LSP. Si bien hemos avanzado en el diseño del ambiente, nos falta implementar algunos módulos e integrarlo con el de automatización de algunas métricas. La idea final es realizar una aplicación centrada en la Web, en la que se integre documentación de requerimientos, evaluación y conclusiones añadido a la funcionalidad de los principales procedimientos de las distintas fases. El ambiente debe permitir seguimiento hacia atrás y hacia adelante.

Finalmente, esperamos que la metodología Web-site QEM evolucione y madure en la medida que evolucione el estado del arte en el dominio de la Web, es decir, la calidad de sus procesos y de sus productos. Es nuestro anhelo y empeño futuro para que la misma pueda ser empleada en procesos de aseguramiento y de certificación de calidad en conformancia con estándares internacionales. En esta dirección, estamos estudiando la evolución del estándar ISO/IEC 9126, como apuntamos en la sección 5.4.2.1, y del estándar ISO/IEC 14598, como lo apuntamos en la sección 11.2 respectivamente.

General Pico - La Pampa, 21 de Julio de 1999

13. Glosario de Términos

13.1 Aclaraciones

Los conceptos del glosario están ordenados alfabéticamente. Un concepto puede ser un único término como **calidad** o una frase como **evaluación de sitios Web**.

Utilizaremos la/s palabra/s resaltada/s en **negrita** para describir un nuevo concepto. Emplearemos un estilo *itálico* para indicar que dicho término o frase se referencia en el glosario. Por último usaremos paréntesis para referirnos a sinónimos o palabras fuertemente relacionadas; por ejemplo, el concepto **artefacto** (producto, distribuible, documento) se halla relacionado o cuenta con dichos sinónimos.

Un término o una frase puede tener más de una acepción en nuestra área de investigación o en un contexto relacionado, por lo que lo indicamos con números.

13.2 Definiciones

abstracción el proceso de identificar a un comportamiento y propiedades comunes de una colección de *objetos* o *entes*.

actividad. 1. (proceso, tarea) representa una unidad abstracta de trabajo a realizar por un *agente*, no considerando gerenciamiento de recursos 2. es un *subproceso* que no requiere más descomposición.

agente. es el *ente* ejecutor de un *proceso*; el agente puede ser tanto un *ente* humano como un ente computarizado (herramienta o dispositivo).

agregación (composición) 1. un tipo de *asociación* que modela la relación todo-parte entre *clases* o *instancias*. 2. (ver agregación de preferencias)

agregación de preferencias. es un *proceso* que consiste en estructurar recursivamente a las *preferencias elementales* a partir del *árbol de requerimientos* en consideración de pesos y operadores lógicos de simultaneidad, reemplazabilidad y neutralidad. Finalmente se computa la *preferencia de calidad global* a partir de la estructura resultante de *preferencias*.

ambiente de soporte a procesos de software. es un entorno de trabajo computarizado que ofrece asistencia a los usuarios participantes en la guía o ejecución en un *proyecto de software*.

anchor, ancla (punto de partida) en un sistema de *Hipermedia* -y por ende la *Web*-, es el origen de los *enlaces*. Representa puntos de partida para la acción de navegación, esto es, cuando se selecciona y activa el anchor, sucede la navegación al *objeto* destino pudiendo permanecer o cambiar de *nodo*, *contexto de navegación* o *sitio*.

árbol de requerimientos, árbol de requerimientos de calidad es una estructura jerárquica que representa a la descomposición en *características*, *subcaracterísticas* y *atributos* que conforma todos los *requerimientos de calidad* a evaluar y/o comparar para un *perfil de usuario* y dominio determinados.

artefacto. (producto, distribuible, documento, versión) 1. es el producto creado, evolucionado, mantenido o destruido durante el *proceso de desarrollo de software* ya como un resultado requerido por un *agente* o para facilitar la prosecución de otro *proceso* 2. es un *objeto* temporario o persistente que representa al producto de realizar una *tarea*.

artefacto Web. (artefacto) es un artefacto de software en el dominio de los sistemas de información centrados en la *Web*. Un artefacto Web puede ser un *sitio Web* o sus componentes como *subsitios*, *páginas*, *applets*, *scripts*, *formularios*, *tablas*, entre otros.

aseguramiento de calidad. es un *proceso* que se aplica en el contexto de una estrategia de calidad con el fin de asegurar que un *ente* cumpla con los *requerimientos de calidad* establecidos.

atributo. 1. lo que se atribuye de lo que es propio de un *objeto* o *ente*. 2. en *métricas* de software, *característica* elemental -que no soporta descomposición- de un *ente*, la cual se la puede medir directamente; esto es, aplicar una *métrica directa*. Sin embargo, hay atributos que sólo se los puede medir por medio de una ecuación (o modelo), compuesta por otros parámetros y variables; en este caso la *métrica es indirecta* 3. *dato* miembro de una *clase*, estructura de dato, variable, estado, contenedor de *información*. 4. (Ver *métrica interna* y *externa*).

calidad¹. 1. El conjunto de las *características* de un *ente* (*artefacto*, *proceso*, o *recurso*) que le confiere la capacidad de satisfacer *requerimientos* o necesidades explícitas e implícitas, en consideración de un *perfil de usuario*. 2. es una *característica* de muy alto nivel, que se la puede medir mediante *métricas indirectas* útil en el proceso de *evaluación* y comparación de ciertas propiedades y relaciones atribuidas a un *ente*.

calidad de un artefacto Web. El conjunto de *características* y *atributos* de *calidad* de un *artefacto Web* que satisfacen los *requerimientos de calidad* establecidos o implicados, considerando un *perfil de usuario* y dominio determinados.

¹ Ver [ISO 8402] para conceptos relacionados a Calidad

característica, características de calidad de un sitio Web. (árbol de requerimientos) un conjunto de *atributos* de un *ente* -o específicamente, de un *artefacto Web*-, por medio de los cuales se describe y evalúa su *calidad*. Una característica permite ser descripta y evaluada por la descomposición recursiva en *subcaracterísticas* y/o *atributos*. Entre las características que describen a la *calidad* desde cierto *perfil de usuario*, se pueden citar a la *usabilidad, funcionalidad, confiabilidad, eficiencia, portabilidad, y mantenibilidad*.

certificación de calidad. la certificación con respecto a una norma o estándar de *calidad* de *producto* o *proceso* es la emisión de un *documento* oficial indicando la satisfacción o *conformidad* de las *características de calidad* del *ente* objeto de juicio con dicha norma o estándar. Para ello es necesario realizar un *proceso* de *evaluación*. Existen certificaciones en conformidad a estándares internacionales como por ejemplo, *ISO*, o certificaciones en conformidad a estándares de facto como *CMM*, -de alta aceptación mundial, aunque principalmente en EEUU.

ciclo de vida. es una serie de *fases* y *procesos* por la que pasa un *artefacto* de software, desde el inicio de un *proyecto de software* -cuando se reconoce la existencia de un *requerimiento* insatisfecho o complementario- pasando posiblemente por *fases* de exploración, desarrollo y *operación*, hasta finalizar con la destrucción del mismo.

clase. (entidad, objeto, instancia) 1. representación abstracta de un conjunto de objetos que exhiben semejante comportamiento y *atributos*. 2. La clase es el molde a partir del cual se crean objetos o instancias.

CMM del inglés Capability Maturity Model y se traduce en Modelo de Madurez de Capacidad de los *procesos* organizacionales. Es una iniciativa del SEI (Software Engineering Institute); es un *modelo* para evaluar y certificar la madurez de los procesos de una organización que desarrolla software. En general, se aplica principalmente en grandes empresas de software.

condición. de un proceso es la declaración del estado de situación que debe ocurrir para el inicio, ejecución y finalización de un *proceso*. Una *actividad* puede comenzar cuando se cumple un conjunto de precondiciones y puede finalizar cuando se alcanzan las postcondiciones establecidas.

conformidad. cumplimiento de los *requerimientos* especificados.

confiabilidad. es una *característica* de *calidad* de alto nivel -que se la puede medir mediante cálculo a partir de *métricas directas e indirectas*-, y que representa el grado en que un *artefacto* está libre de errores y/o fallas u operable bajo ciertas condiciones especificadas.

constructor de proceso (método) es un *enfoque* específico o *método* que se puede usar para realizar una *tarea* o conjunto de tareas en dominios semejantes.

contexto de navegación. 1. *subsitio* 2. es una primitiva de diseño (de *OOHDM*) que está compuesta por un conjunto de *nodos*, *enlaces* y otros contextos -posiblemente anidados. Este *constructor de proceso* permite representar unidades cohesivas de conceptos y establecer relaciones semánticas y estructurales apropiadas favoreciendo la orientación del usuario en la aplicación.

criterio elemental. (criterio elemental absoluto, criterio elemental relativo, nivel de puntaje, lógica de preferencia continua) es una correspondencia o mapeo del valor de la variable X_i -asociado al *atributo* A_i -, en el valor de la *preferencia de calidad elemental* IE_i . Analíticamente se define como una función de *evaluación*.

criterio elemental absoluto. es un *criterio elemental* en donde interviene un *requerimiento* absoluto y el nivel de cumplimiento del mismo depende solamente de un *atributo* de un *artefacto* en el *proceso* de *evaluación*. A su vez, un criterio elemental absoluto puede clasificarse en criterio de variable continua y en criterio de variable discreta.

criterio elemental relativo. es un *criterio elemental* formado a partir de una relación *entre variables* de más de un sistema competitivo -al menos dos-, en el *proceso* de *evaluación*.

dato/s. 1. consiste de los valores recolectados de un *atributo* -sensado, medido, observado, etc.- para un *ente* en un *dominio de aplicación* dado. 2. ítem elemental o primitivo de distinta naturaleza o medios, que sirve para contener y/o comunicar *información* elemental -con muy bajo nivel de elaboración.

descripción de proceso es una manera de representar y especificar la secuencia parcial de *actividades* de un *proceso*. Una descripción completa de proceso debe considerar las actividades y las *operaciones* asociadas, las *condiciones* para cada actividad, y a otros *entes* intervinientes en el *proceso de software* como *artefactos*, *agentes* y *roles*.

distribuable. (producto, artefacto, documento) es un *artefacto* requerido por algún *proceso* o *agente* interno o externo.

documento. (producto, artefacto, distribuable) es un *artefacto* que posee un contenido o *información* y que puede contener, referenciar o enlazar a otros artefactos.

dominio de aplicación. es un sistema real o abstracto del universo que existe independientemente del sistema de *evaluación*. Consiste de un conjunto de *entes* a los

que se le atribuyen propiedades (*atributos, características*) y manifiestan un comportamiento y se relacionan.

dominio del atributo. es el conjunto de posibles valores de un *atributo*.

eficiencia. es una *característica* de *calidad* de alto nivel -que se la puede medir mediante cálculo a partir de *métricas directas e indirectas*- y que representa principalmente a la relación entre el grado de performance del *artefacto* y la cantidad de recursos (tiempo, espacio, etc.) usados bajo ciertas condiciones.

elemento de proceso. (subproceso, paso de proceso)

enfoque. (estrategia)

enlace. 1 un vínculo entre *objetos* de navegación 2. En *OOHDM* los enlaces son la realización navegacional de las relaciones definidas en el modelo conceptual.

ente, entidad (clase, objeto) cualquier cosa, tangible o intangible, que está o exhibe comportamiento en el mundo real. El mundo real puede ser el mundo físico -hombre, computadora, teléfono-, o un mundo intangible o abstracto -organización, *evaluación, proyecto de software*-.

escala de preferencia. es una representación gráfica de un *criterio elemental*.

estrategia. (enfoque) define las características más generales y representativas de un *proceso* -de desarrollo, de evaluación, etc., sus *principios* fundamentales y los *objetivos* y *metas* a alcanzar.

etapa. (fase, paso) *tareas* fundamentales de un *modelo de proceso*.

evaluación. (evaluación cuantitativa, evaluación de sitios Web) es un *proceso* que consiste en determinar el grado en que un *atributo, subcaracterística* o conjunto de *características* de un *ente* (*artefacto, proceso, o recurso*) satisface a un conjunto de *requerimientos* establecidos o implicados, considerando una *meta*, un *dominio de aplicación*, y un *perfil de usuario* determinados.

evaluación cuantitativa. *evaluación* basada principalmente en *métricas objetivas*, cuantificables, precisas -dentro de la tolerancia de error permitida-, útil para determinar y justificar una óptima toma de decisión.

evaluación de sitios Web. (evaluación cuantitativa) es un *proceso* que consiste en determinar el grado en que un *atributo, subcaracterística* o conjunto de *características*

de un *artefacto o sitio Web* satisface a un conjunto de *requerimientos* establecidos o implicados, considerando un *dominio* de aplicación *Web* y a un *perfil de usuario*.

evento. alguna ocurrencia que provoca el cambio de estado de un *objeto*.

experticia. conjunto de habilidades específicas de un *agente*, surgido del estudio, del conocimiento y de la *práctica*.

fase. (etapa) es una agrupación de *procesos* fuertemente relacionados o cohesivos realizados en cierto orden. Las distintas fases pueden exhibir comportamientos diferentes.

fase operativa es la fase de mantenimiento y evolución en el *ciclo de vida* de un *artefacto*. Se pueden enumerar otras fases en un *proyecto de software* como la fase de exploración y la fase de desarrollo.

funcionalidad. es una *característica* de *calidad* de alto nivel -que se la puede medir mediante cálculo a partir de *métricas directas e indirectas*-, y que representa la existencia de un conjunto de funciones y comportamientos -para un *artefacto*- que satisfagan a un conjunto especificado de *requerimientos* en consideración de un *dominio* y *perfil/es de usuario/s*.

GQM del inglés, Goal-Question-Metric approach, que se traduce en enfoque Meta-Pregunta-Métrica. Dado un conjunto seleccionado de *metas* del proyecto en el contexto de una organización -teniendo en cuenta las *características* y *atributos* deseables de los *artefactos*, *productos* o *recursos*, se construye y refina un conjunto de preguntas para cada meta y, en función de cada pregunta se eligen las *métricas* apropiadas.

heurística. *principio*, criterio o regla *práctica*, surgida de la *experticia*.

Hipermedia. 1. es un conjunto organizado de *información* de diferentes medios, vinculados por relaciones estructurales y semánticas. 2. es la ciencia que se ocupa de estructurar, presentar y permitir acceso directo al contenido y relaciones, en un espacio organizado de *información*.

información. interpretación, elaboración y presentación de los *datos*, en cierto contexto.

ingeniería. (Ingeniería de Software) es el empleo sistemático, disciplinado, y cuantificable de *heurísticas*, *principios*, *modelos*, *métodos*, *técnicas* y/o herramientas derivado de la ciencia y la experiencia, en respuesta de alguna necesidad o deseo humano para comprender o resolver algún problema o aspecto de un *ente*. Debe haber una adecuada relación costo-beneficio en la prosecución y solución del problema.

Ingeniería de Software. (ingeniería) es el empleo sistemático, disciplinado, y cuantificable de *heurísticas, principios, modelos, métodos, técnicas* y/o herramientas derivado de la ciencia y la experiencia, principalmente en el dominio del software, en respuesta de alguna necesidad o deseo humano para comprender o resolver algún problema o aspecto de un *ente* (relacionado a sistemas de información o partes del mismo, esencialmente en su componente software).

instancia. (objeto) una ocurrencia específica de una *clase*

ISO del inglés International Standard Organization, que se traduce en Organización de Estándares Internacionales. Normatiza diversas características de *procesos* y *productos* ya sean del dominio del software o de otros dominios. Por ejemplo, normas sobre las *características de calidad* del producto “software” se encuentran especificadas en el estándar 9126; para la proceso de *evaluación de calidad* de productos, en el estándar 14598; para la evaluación de los procesos de ciclo de vida del software, en el estándar 12207; etc.

lógica de preferencia continua. se interpreta a la *preferencia elemental* como una variable de lógica continua en el intervalo unitario -perteneciente al dominio de los números reales. El valor 0 de la *preferencia* IE_i denota que no satisface los *requerimientos*, en tanto que el valor 1 denota total satisfacción de los mismos. Los valores entre 0 y 1 indican parcial cumplimiento de los requerimientos establecidos. Consecuentemente, la preferencia se puede interpretar como el porcentaje de requerimiento/s satisfecho/s considerando el rango de 1 a 100%

LSP del inglés Logic Scoring of Preference approach. *Modelo* y procedimiento basado en la *lógica de preferencia continua* para determinar la *preferencia de calidad global* (y *ranquin*) del o los sistemas o *artefactos* intervinientes, a partir de un conjunto de *preferencias de calidad elementales* relacionados por medio de operadores lógicos.

medición. 1. el *proceso* de medir. Es un *proceso* sistemático y disciplinado que consiste en *recolectar datos* de un *ente* (*artefacto, proceso, o recurso*) con el fin de analizarlos y suministrar *información* para comprender, controlar, mejorar y/o predecir el fenómeno o *ente* bajo estudio. 2. correspondencia o mapeo desde un dominio empírico o real a un mundo formal, matemático, relacional. Por lo tanto, una *medida* es un valor numérico o nominal asignado al *atributo* de un *ente* por medio de dicha correspondencia o mapeo.

medida. *producto* de aplicar una *medición*.

meta. (objetivo) representa a un conjunto de declaraciones de los resultados que se desean alcanzar generalmente de un *proceso*, en el contexto de la estrategia organizacional. Los resultados pueden estar en función de *artefactos, procesos*, etc.

método. 1. (constructor de proceso, procedimiento) modo específico de realizar una *tarea* o resolver un *proceso*. Curso de acción u *operaciones* y conjunto de estándares y procedimientos de modelado a usar para tratar con algún *proceso* de un *proyecto*. 2. implementación de un servicio u *operación* de una *clase*.

metodología. conjunto de *métodos* asociados a un *enfoque* con el fin de cubrir una o más *fases* o una parte significativa de una fase de un *proyecto*.

métrica. 1. (medida) es un valor numérico o nominal asignado a *características* o *atributos* de un *ente* computado a partir de un conjunto de *datos* observables y consistentes con la intuición. Para que la métrica sea válida debe poseer un conjunto de características entre las que podemos enumerar: robustez, consistencia, permitir escalas y rangos, significancia, y facilidad en la *recolección de datos*. Una métrica puede ser directa o indirecta, interna o externa –a veces, dentro de esta última categoría, se la suele llamar también objetiva o subjetiva 2. procedimiento y criterios usados para determinar el valor de *características* o *atributos* de un *artefacto*, *proceso* o *recurso*.

métrica directa. es lo resultante de una correspondencia o mapeo directo entre un *atributo* de un *ente* (del dominio empírico) y el valor (del dominio numérico), y que sirve como referencia para describir y evaluar aspectos o situaciones del mundo empírico. Los atributos se miden generalmente por medio de métricas directas (aunque no siempre) y las *características* y *subcaracterísticas* se miden por medio de *métricas indirectas*.

métrica externa. (métrica objetiva, subjetiva) es el valor resultante del atributo al aplicar una *métrica indirecta* y que siempre involucra al *ente* y su comportamiento con el entorno.

métrica indirecta. es lo resultante de una correspondencia o mapeo entre relaciones de dos o más *atributos*, o de un atributo compuesto (del dominio empírico) y el nuevo valor (del dominio numérico), y que sirve como referencia para describir y evaluar *características* y/o *subcaracterísticas* y/o *atributos compuestos* de un *ente* (*artefacto*, *proceso*, o *recurso*) del mundo empírico. Los atributos se miden generalmente por medio de *métricas directas* y las *características* y *subcaracterísticas* se miden por medio de métricas indirectas (cálculos).

métrica interna. (métrica objetiva) es un valor numérico o nominal del atributo que siempre involucra al *ente* en sí, ya sea obtenido por una *métrica directa* o *indirecta*.

métrica objetiva. es el valor resultante del atributo de un *ente* (*artefacto*, *proceso*, o *recurso*), comprobable, independiente del juicio o subjetividad humana. No obstante, es importante destacar que existen grados de objetividad.

métrica subjetiva. es un valor numérico que siempre involucra al juicio humano por medio de *heurísticas* o criterios de preferencia directa.

modelo es una representación abstracta (abstracción) de *entes* o fenómenos del mundo real en la que se consideran los aspectos relevantes de los mismos y se desechan los menos relevantes sin que por ello deje de representar significativamente a esa realidad. Es una estructura en un dominio usado para representar a *entes* de otro dominio, con el propósito de comprenderlo y/o controlarlo.

modelo de calidad. conjunto de *características*, *subcaracterísticas* y las relaciones entre ellas que proveen la base para especificar *requerimientos de calidad* con el fin de realizar un *proceso de evaluación*.

modelo de ciclo de vida. (modelo de proceso de software, modelo de proceso de producto).

modelo de proceso de desarrollo de software. (modelo de ciclo de vida) 1. es una *estrategia* apropiada para abstraer, organizar, ejecutar y/o controlar a las distintas *fases*, *tareas*, *recursos* y *artefactos* de un *proyecto de software* con el objeto de alcanzar las *metas* establecidas. 2. una descripción más o menos formal del *proceso de desarrollo de software*. Por lo que un modelo de proceso expresa: a) un cierto nivel de *abstracción*; b) una *perspectiva* particular del *proceso de desarrollo*.

modelo de proceso de evaluación de software. 1. es una *estrategia* apropiada para abstraer, organizar, ejecutar y/o controlar a las distintas *fases*, *tareas*, *recursos* y *artefactos* de un *proyecto de evaluación* con el objeto de alcanzar las *metas* establecidas. 2. Es una descripción más o menos formal del *proceso de evaluación de software*.

modelo de proceso de producto. (modelo de proceso de software, modelo de ciclo de vida).

modelo de proceso Web. 1. es un *modelo de proceso* para abstraer, organizar, ejecutar y/o controlar a las distintas *fases*, *tareas*, *recursos* y *artefactos* de un *proyecto de software Web* con el objeto de alcanzar las *metas* establecidas. 2. una descripción más o menos formal *del proceso de desarrollo de software Web*.

nivel de puntaje (rating). obtenida la *preferencia de calidad*, el rango de valores que permiten clasificar al *atributo* o *característica* del *artefacto* en niveles de cumplimiento de los *requerimientos* establecidos. Tres niveles principales de aceptabilidad son, a saber: insatisfactorio, marginal y satisfactorio definidos por consenso en tres rangos de valores.

nodo. 1. Es una ventana lógica o *página* en un *sitio Web* 2. En *OOHDM* es una *clase* navegacional que contiene elementos de *información*, esencialmente *atributos* y *puntos de partida* o *anchors*.

objetivo. (meta) representa a un conjunto de declaraciones de los resultados que se desean alcanzar en el tiempo. Los resultados pueden estar en función de *artefactos*, *procesos*, o recursos. El objetivo, con respecto al tiempo, es de menor plazo que la meta y debe ser cuantificable.

objeto. (entidad, clase) 1. es la representación de un comportamiento que ocurre en el mundo real, tangible o intangible. Un objeto exhibe un comportamiento y se le atribuyen propiedades o *atributos* 2. una instancia u ocurrencia de una clase.

OOHDM que en inglés se traduce en Object-Oriented Hypermedia Design Method, es una *metodología* basada en *modelos* y *principios* de Orientación de Objetos útil para la especificación y construcción de *artefactos* de *Hipermedia*.

operación. es una acción específica que se puede invocar para realizar una tarea. -En plural- conjunto de servicios que dispone una *clase* para responder a solicitudes externas.

página, página Web. (artefacto, sitio Web) es un *artefacto Web*. Es la unidad básica o *nodo* de la cual se compone un *sitio Web*. A su vez, una página puede estar compuesta de *artefactos* HTML, scripts, applets, e *información* o *datos* como sonidos, imágenes, texto, etc.

paso, paso de un proceso. 1. (subproceso, elemento de proceso) 2. (etapa)

perfil de usuario. (clase, perspectiva) es el punto de vista de usuario a partir del cual se construyen, especifican, evalúan, o comparan los *requerimientos*. En el *ciclo de vida* de un *artefacto Web*, la *evaluación de calidad* puede ser considerada generalmente desde el perfil del visitante, del desarrollador, o del administrador. Dada una clase de usuario, se puede aplicar un mecanismo de especialización.

perspectiva. (vista, submodelo) es un enfoque particular para especificar y comunicar *información del modelo*.

preferencia. (lógica de preferencia continua) el grado de verdad alcanzado, en términos de lógica continua, de un valor medido de un *atributo* o *característica* especificado en el *proceso de evaluación*. El grado de *conformidad* o cumplimiento de un atributo o característica del *requerimiento* establecido.

preferencia de calidad. es la *preferencia* en consideración de la *calidad* de un *artefacto, proceso o recurso*.

preferencia elemental, preferencia de calidad elemental. es la *preferencia de calidad* de un *atributo* como resultante de aplicar un *criterio elemental* en el *proceso de evaluación*.

preferencia de calidad global. es la *preferencia de calidad* de un *ente* complejo o de partes del mismo abarcando una o varias *características de calidad* en el *proceso de evaluación*.

principio. una regla general o verdad fundamental que puede servir de guía en la toma de decisiones.

procedimiento. (método)

proceso. (proceso de software, actividad, tarea) 1. *ente* usado para producir *artefactos*. 2. *ente* que mediante la función correspondiente, transforma entradas en salidas.

proceso de evaluación. (proceso) un conjunto parcialmente ordenado de *subprocesos* a los que se le asocia una colección de *recursos, agentes, condiciones, artefactos y constructores de proceso*, con el fin de producir los *distribuibles* conforme a las *metas* establecidas, en un *proyecto de evaluación*.

proceso de desarrollo de software, proceso de desarrollo es la realización de un *proyecto de software*.

proceso de Hipermedia, proceso Web que en inglés se traduce en Hypermedia Process o Web Process es un *proceso de software* general, abstracto, personalizado al dominio de aplicaciones de *Hipermedia* o *Web*.

proceso de software. (proceso) un conjunto parcialmente ordenado de *subprocesos* a los que se le asocia una colección de *recursos, agentes, condiciones, artefactos y constructores de proceso*, con el fin de producir los *distribuibles* conforme a las *metas* establecidas en un *proyecto de software*.

producto, producto de software (artefacto)

proyecto de evaluación es un *ente* que comprende a un conjunto de *tareas*, tanto técnicas como de gerenciamiento, a un conjunto de *recursos*, a un conjunto de *estrategias, métodos y heurísticas*, con el propósito de lograr los *objetivos* y las *metas* de evaluación acordadas.

proyecto de software es un *ente* que comprende a un conjunto de *tareas*, tanto técnicas como de gerenciamiento, a un conjunto de *recursos*, a un conjunto de *estrategias*, *métodos* y *heurísticas*, con el propósito de lograr los *objetivos* y las *metas* de exploración, desarrollo o mantenimiento acordadas.

proyecto de software Web. es un *proyecto de software* con *recursos*, *actividades*, *constructores de proceso*, *roles*, *atributos* y *requerimientos* específicos del área de sistemas de información centrados en la *Web*.

punto de partida. (anchor, ancla)

ranquin (ranking). posición o puntaje final obtenido en el *proceso* de *evaluación* y comparación de los sistemas o *artefactos* seleccionados, en consideración de las *preferencias (de calidad) global* resultantes.

recolección de datos. es la *actividad* manual, semiautomática o automática de recolección de valores de *atributos* de un *ente*, con el fin de servir de entrada al *proceso* de *medición*.

recurso. 1. es un *ente* requerido como entrada por un *proceso* para producir alguna salida especificada 2. es un *ente* necesario para que las *tareas* de un *proyecto de software* se puedan efectuar. Recursos de un proyecto son: humanos, monetarios, materiales, tecnológicos, temporales.

requerimiento. (requerimientos de calidad) una necesidad o deseo del usuario, explícito o implícito, que se traduce en *atributos*, *características* o funciones necesarias de un *artefacto* de software (o *ente*), y que se puede considerar desde el contexto del sistema en un *proceso* dado.

requerimientos de calidad. (árbol de requerimientos) *características* y *atributos* de *calidad*, por medio de las cuales se especifican *metas*, necesidades, y prioridades a satisfacer por un *artefacto*, en consideración de un *perfil de usuario* en el *proceso* de *aseguramiento de calidad* y/o de *evaluación* o comparación.

RMM. que en inglés se traduce en Relationship Management Methodology, es una *metodología* basada en un *enfoque* estructurado y en el *modelo* de Entidad-Relación útil para la especificación y construcción de *artefactos* de *Hipermedia*.

rol. es un conjunto de permisos y obligaciones que se debe asociar a un *agente* durante la realización de un tipo de *tarea*. El agente debe tener un conjunto de permisos para realizar las actividades de la tarea conforme a la submeta establecida y obligado a satisfacer un conjunto de *condiciones*.

sitio, sitio Web. (artefacto, subsitio) es un *artefacto Web*. Generalmente, es una colección estructurada de *páginas Web* que representa a un *ente* real -organización, grupo, individuo-, o partes de él -sistemas de información, *datos*, procedimientos, etc-

subcaracterística. (característica, atributo) parte o componente de una *característica* de un *ente* (*artefacto, proceso, o recurso*), usualmente como resultado de aplicar un mecanismo de descomposición.

submodelo. parte o componente de un *modelo*

subproceso. (paso de proceso, elemento de proceso), cualquier *ente* resultante de la descomposición recursiva de un *proceso*. El nodo raíz del árbol es el *proceso* más general y abstracto, el nodo hoja es el *proceso* más específico y lo denominamos *actividad*.

subsitio. (contexto de navegación) parte de un *sitio* que representa una unidad cohesiva de contenido y/o funciones.

tarea. (actividad, proceso) representa una unidad concreta de trabajo a realizar por un *agente* considerando gerenciamiento de recursos. Es un *proceso* al que se le asocian componentes de gestión, es decir, se le pueden asignar *agentes, recursos*, se la puede planificar, programar, ejecutar y controlar. Las tareas manuales sólo involucran a *agentes* humanos; las tareas automáticas sólo dependen de *agentes* computarizados, en tanto que las tareas interactivas necesitan de ambos tipos de *agentes*.

técnica. *procedimientos y heurísticas* específicas usadas por un *método*.

usabilidad. es una *característica* de *calidad* de alto nivel -que se la puede medir mediante cálculo a partir de *métricas directas e indirectas*-, y representa la capacidad o potencialidad del producto para ser utilizado, comprendido y operado por los usuarios, además de ser atractivo. Incluye principalmente a subcaracterísticas como comprensibilidad, operabilidad, y comunicatividad, entre otras características como estéticas y de estilo que hacen del artefacto que sea agradable de usar.

validación. *proceso* para asegurar que las *medidas* sean una caracterización numérica apropiada del *atributo* de un *ente*, mostrando que se satisfaga la condición de representación. Existen dos *estrategias* de validación: teórica y empírica.

versión. es un *artefacto* revisado y aprobado, que representa cierta variante del mismo. Generalmente, una versión posterior se corresponde a una adaptación, extensión o mejora respecto de la versión anterior.

Web. (World Wide Web) sistema de información en Internet que vincula mediante *hipermedia* a *documentos* o *artefactos* ubicados en servidores distribuidos en todo el mundo y que, desde el punto de vista del usuario, permite en principio acceder e interactuar con los mismos independientemente de la ubicación física.

Web-site QEM. que en inglés se traduce en Web-site Quality Evaluation Method, es una *metodología* basada en *métodos, modelos, principios* y herramientas de *Ingeniería de Software* útil para la *evaluación y comparación cuantitativa* de la *calidad* de *artefactos Web*, principalmente en la *fase operativa* del ciclo de vida. No obstante, se puede utilizar *en fases de exploración y desarrollo*, en este caso se debe sincronizar con el *modelo de proceso de desarrollo*.

14. Referencias

- [Albrecht et al 83] **Albrecht, A.; Gaffney, J.**, 1983, “*Software, source lines of code and development effort prediction: a software science validation*”, IEEE Transaction on SE, 9 (Nov83) pp. 639-647
- [Arnold et al 77] **Arnold, S.; Ma, S; Tigert, D.**, 1977, “*A comparative analysis of determinant attributes in retail store selection*”, Advances in Computer Research. Assoc. for Consumer Research (5), pp. 663-667
- [Basili et al. 84] **Basili, V.R., Weiss, D.**, 1984, “*A methodology for Collecting Valid Software Engineering Data*”, IEEE Transactions on Software Engineering, Vol. 10 (6), pp. 728-738.
- [Basili et al. 88] **Basili, V.R., Rombach, H.D.**, 1988, “*The TAME Project: Towards Improvement-oriented Software Environments*”, IEEE Transactions on Software Engineering, Vol. 14 (6), pp. 758-773.
- [Basili et al. 94] **Basili, V.R., Caldiera, C., Rombach, H.D.**, 1994, “*Goal Question Metric Paradigm*”, Encyclopedia of Software Engineering, Vol. 1, John Wiley & Sons, pp. 528-532 .
- [Berard 93] **Berard, E.**, 1993, “*Essays on Object-Oriented Engineering*”, Vol 1. Prentice Hall.
- [Bevan 95] **Bevan, N.**, 1995, “*Measuring Usability as Quality of Use*”, Software Quality Journal, 4 (2), pp. 115-130.
- [Bevan 97] **Bevan, N.**, 1997, “*Usability Issues in Web Site Design*”, Proceed. of HCI Int’l Conference. San Francisco. Elsevier Press.
- [Boehm et al 76] **Boehm, B.; Brown, J.R.; Lipow, M.**, 1976, “*Quantitative Evaluation of Software Quality*”, Proceed. 2nd Int’l Conference on Software Engineering, pp. 592-605.
- [Boehm et al 78] **Boehm, B.; Brown, J.R.; Kaspar, J.R.**, et al. 1978, “*Characteristics of Software Quality*”, TRW Series of Software Technology.
- [Boehm 81] **Boehm, B.**, 1981, “*Software Engineering Economics*”, Prentice Hall
- [Boehm 88] **Boehm, B.**, 1988, “*A Spiral model of Software Development and Enhancement*”, IEEE Comp. 21, 5 (May88).
- [Boehm et al 95] **Boehm, B; Clark, E.; Horowitz, R;** et al., 1995, “*Cost models for future life cycle processes: COCOMO 2.0*”, Annals of Software Engineering 1(1), pp. 1-24.
- [Booch 94] **Booch, G.**, 1994, “*Object-Oriented Analysis and Design with Application*”, Benjamin/Cummings.

- [Botafogo 92] **Botafogo, R. Rivlin, E., Shneiderman, B.**, 1992, "*Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics*". ACM Transactions on Office Information Systems, 10(2), pp. 142-180.
- [Briand et al 96] **Briand, L., Morasca, S., Basili, V.**; 1996, "*Property-Based Software Engineering Measurement*", IEEE Transaction on Software Engineering, 22(1), pp. 68-85.
- [Briand et al 97] **Briand, L., Devandu P., Melo M.**; 1997, "*An Investigation into Coupling Measures for C++*". Proceedings of ICSE'97, Boston, MA, USA, pp. 412-421.
- [Brooks 87] **Brooks, F.**, 1987, "*No Silver Bullet: Essence and Accidents of Software Engineering*", IEEE Comp. 20,4 (Apr 87); pp.10-19
- [Chidamber et al 94] **Chidamber, R.; Kemerer, C.**, 1994, "*A Metrics Suite for Object Oriented Design*", IEEE Transaction on Software Engineering, 20: pp. 476-493
- [Curtis et al 92] **Curtis, B.; Kellner, M.; Over, J.**, 1992, "*Process Modelling*", Comm. ACM 35, 9; pp. 75-90.
- [Davis 93] **Davis, A.**, 1993, "*Software Requirements, Objects, Functions and States*", Prentice Hall
- [Deshpande et al 99] **Deshpande, Y., Hansen, S; Murugesan, S.**; 1999, "*Web Engineering: beyond CS, IS, and SE*", Proceed. of ICSE '99 (Int'l Conference on Software Engineering) Web Engineering Workshop, Los Angeles, US, pp 10-16.
- [Dowson et al 91] **Dowson, M. ; Nejme, B.; Riddle, W.** 1991, "*Fundamental Software Process Concepts*", Proceed. First European Workshop on Software Process Modelling, AICA Press.
- [Dujmovic 74] **Dujmovic, J.J.**, 1974, "*Weighted Conjunctive and Disjunctive Means and their Application in System Evaluation*", Journal of the University of Belgrade, EE Dept., Series Mathematics and Physics, N° 483, pp.147-158.
- [Dujmovic et al 82] **Dujmovic, J.J.; Elnicki, R.**, 1982, "*A DMS Cost/Benefit Decision Model: Mathematical Models for Data Management System Evaluation, Comparison, and Selection* ", National Bureau of Standards, Washington D.C. N° GCR 82-374. NTIS N° PB 82-170150.
- [Dujmovic 91] **Dujmovic, J.J.**, 1991, "*Preferential Neural Network. Chapter 7 in Neural Network – Concepts, Applications, and Implementations*", Edited by P. Antognetti and V. Milutinovic. Prentice Hall Advanced Reference Series, Vol. 2, pp.155-206.
- [Dujmovic 96] **Dujmovic, J.J.**, 1996, "*A Method for Evaluation and Selection of Complex Hardware and Software Systems*", The 22nd Int'l Conference for the Resource Management and Performance

- Evaluation of Enterprise CS. CMG 96 Proceedings, Vol. 1, pp.368-378.
- [Feiler et al 93] **Feiler, P. H; Humphrey, W.** 1993, “*Software Process Development and Enactment: Concepts and Definitions*“, Proceed. Second Int. Conf. On the Software Process, US (Feb. 93).
- [Fenton 91] **Fenton, N.E.**, 1991, “*Software Metrics a Rigorous Approach*”, Chapman & Hall, London.
- [Fenton 96] **Fenton, N.E.**, 1996, “*Do standards improve product quality?*”, IEEE Software, 13(1), pp. 22-24.
- [Fenton et al 97] **Fenton, N.E., Pfleeger, S.L.**, 1997, “*Software Metrics: a Rigorous and Practical Approach*”, 2nd Ed., PWS Publishing Company.
- [Furano et al 97] **Furano, F.; Orsini, R.; Celentano, A.**, 1997, “*Museum-on-demand: dynamic management of resources in World Wide Web museums*”, Hypertext and Hypermedia: Products, Tools, Methods (H2PTM'97) , Hermes Editorial, Paris, Fr, V. 1, N° 2-3-4/97 pp. 115,124.
- [Garzotto et al 91] **Garzotto, F.; Paolini, P., Schwabe, D.**; 1991, “*HDM, a model for a design of Hypertext Application*”, Proceed. of Hypertext'91, ACM Press.
- [Garzotto et al 93] **Garzotto, F.; Schwabe, D.; Paolini, P.**, 1993, “*HDM, a model based approach to Hypermedia Application Design* ”, ACM Transaction on Information System, Vol. 11, 1, Jan 93, pp. 1-26.
- [Garzotto et al 95] **Garzotto, F.; Mainetti, L.; Paolini, P.**, 1995, “*Hypermedia Design, Analysis, and Evaluation Issues*”, CACM 38,8 (Ago-95); pp 74-86.
- [Garzotto et al 97a] **Garzotto, F.; Mainetti, L.; Paolini, P.**, 1997, “*Designing Modal Hypermedia Applications*”, The Eighth ACM Conference on Hypertext, Southampton, England, pp 38-47.
- [Garzotto et al 97b] **Garzotto, F.; Matera, M.**, 1997, A Systematic Method for Hypermedia Usability Inspection, *The New Review of Hypermedia and Multimedia*, 3, 1997, pp. 39-65.
- [Gibbs 94] **Gibbs, W.**, 1994, “*Software's chronic crisis*”, Scientific American.
- [Gilb 69] **Gilb, T.**, 1969, “*Weighted Ranking by Levels*”, IAG Journal, Vol 2 (2), 1969, pp. 7-22 .
- [Gilb 76] **Gilb, T.**, 1976, “*Software Metrics*”, Chartwell-Bratt, Cambridge, MA.
- [Gilb 88] **Gilb, T.**, 1988, “*Principles of Software Engineering Management*”, Adisson-Wesley.

- [Goldberg et al 95] **Goldberg, A.; Rubin, K.**, 1995, “*Succeeding with Objects: decision frameworks for project management*”, Addison-Wesley.
- [Grady 94] **Grady, R.B.**, 1994, “*Successfully Applying Software Metrics*”, IEEE Computer, 27 (9), pp. 18-25.
- [GVU 99] **GVU's WWW User Survey**, The tenth edition, http://www.gvu.gatech.edu/user_surveys/survey-1998-10/tenthreport.html
- [Henderson et al 90] **Henderson-Sellers, B; Edwards, J.**, 1990, “*The Object-Oriented systems lifecycle*”, Comm. ACM 33, 9.
- [Henry et al 81] **Henry, S.M.; Kafura, D.G.**, 1981, “*Software Structure Metrics Based On Information Flow*”, IEEE Transactions on Software Engineering, Vol 7, N°5, pp. 510-518.
- [Hetzel 93] **Hetzel, W.C.**, 1993, “*Making Software Measurement Work: Building an Effective Software Measurement Program*”, QED Publishing Group.
- [Humphrey 88] **Humphrey, W.S.**, 1988, “*Characterizing the Software Development Process: A Maturity Framework*”, IEEE Software, 5,2 (March88).
- [Humphrey et al 89] **Humphrey, W.S., Kellner, M.I.**, 1989, “*Software Process Modelling: Principles of Entity Process Models*”, Proceed. of the 11th Int. Conference on Software Engineering, IEEE Comp. Society.
- [Humphrey 89] **Humphrey, W.S.**, 1989, “*Managing the Software Process*”, Addison-Wesley.
- [IEEE 1061] **IEEE Std 1061-1992**, “*IEEE Standard for a Software Quality Metrics Methodology*”, IEEE Computer Society Press
- [IEEE 830] **IEEE Std 830-1993**, “*IEEE Recommended Practice for Software Requierements Specifications*”
- [IEEE WPG] **IEEE Web Publishing Guide**, <http://www.ieee.org/web/developers/style/>
- [Isakowitz et al 95] **Isakowitz,T.; Stohr, E.; Balasubramanian, P.**, 1995, “*RMM: a methodology for structured hypermedia design*”, Comm. ACM 38, 8 (Aug 95) pp. 34-48
- [ISO 8402] **ISO 8402:1994** *Quality - Vocabulary*
- [ISO/IEC 12207] **ISO/IEC 12207:1995** *Information technology – Software life cycle processes.*
- [ISO/IEC 14598-1] **ISO/IEC DIS 14598-1** *Information technology -- Software product evaluation -- Part 1: General overview*
- [ISO/IEC 14598-5] **ISO/IEC 14598-5:1998** *Information technology -- Software product evaluation -- Part 5: Process for evaluators*
- [ISO 9001] **ISO 9001: 1994**, “*Model for Quality Assurance in design, development, production, installation and servicing*”

- [ISO 9000-3] **ISO/IS 9000-3**, 1990, “*Quality Management and Quality Assurance Standards – Part 3: Guidelines for the Application of ISO 9001 to the Development, Supply and Maintenance of Software*”
- [ISO/IEC 9126] **ISO/IEC 9126-1991 International Standard**, “*Information technology – Software product evaluation – Quality characteristics and guidelines for their use*”.
- [ISO/IEC 9126-1] **ISO/IEC 9126-1 International Standard** (Draft), 1998 “*Information technology – Software product evaluation – Part 1: Quality Model*”.
- [ISO/IEC 9126-2] **ISO/IEC 9126-2 International Standard**, 1999 (Draft) “*Information technology – Software product evaluation – Part 1: External Metrics*”.
- [ISO/IEC 9126-3] **ISO/IEC 9126-3 International Standard**, 1999 (Draft) “*Information technology – Software product evaluation – Part 1: Internal Metrics*”.
- [ISO/IEC 9126-4] **ISO/IEC 9126-4 International Standard**, 1999 (Draft) “*Information technology – Software product evaluation – Part 1: Quality in use Metrics*”.
- [ISO_SPICE] **ISO-International Standard Organization**, 1993, “*SPICE Baseline Practice Guide, Product Description*”
- [Kendall et al 98] **Kendall, K.; Kendall, J.** , 1998, “*Systems Analysis and Design*” Prentice Hall College Div., 4th edition.
- [Kitchenham et al 89] **Kitchenham, B.A., Walker, J.G.** 1989, “*A Quantitative Approach to Monitoring Software Development*” *Software Engineering Journal*, 4(1), pp. 2-13.
- [Kitchenham et al 95] **Kitchenham, Pickard, L., Pfleeger, S. L.**, 1995, *Case Studies for Method and Tool Evaluation* , *IEEE Software*, July 95, pp. 52-62.
- [Kitchenham et al 96] **Kitchenham, B., Pfleeger, S. L., And Fenton, N.**, 1996, *Towards a Framework for Software Measurement Validation*. *IEEE Transactions on Software Engineering*, 21(12), pp. 929-944.
- [Kitchenham et al 97] **Kitchenham, B., Stell, J.**, 1997, “*The danger of using Axioms in Software Metics*”, *IEEE Proceedings on Software Engineering*, Vol. 144, N° 5-6, pp. 279-285.
- [Kirakowski et al 94] **Kirakowski, J., Corbett, M.**, 1994, “*SUMI: the Software Usability Measurement Inventory*”, *Brit J Ed Technology*, 23.
- [Kirakowski et al 98] **Kirakowski, J., Cierlik, B.**, 1998, “*Measuring the Usability of Web Sites*”, *Human Factors and Ergonomics Society Annual Conference*, Chicago, US.
- [Krantz et al 71] **Krantz, D., Luce, R. D., Suppes, P.; Tversky, A.** 1971, “*Foundation of Measurements – Additive and Polynomial*”

- Representation*”, Academic Press, Vol. 1.
- [Lafuente et al 99] **Lafuente, G.H., Lafuente, G.J., Nicolai, P.; Olsina, L.**, 1999, “*Un Survey sobre Atributos de Calidad de Sitios Web*”, CD-ROM del V Congreso Argentino de Ciencias de la Computación. (CACIC), Tandil, Argentina.
- [Lonchamps 93] **Lonchamps, J.**, 1993, “*A Structured Conceptual and Terminological Framework for Software Process Engineering*”, ICSP 2, Berlin, IEEE Press.
- [Lohse et al 98] **Lohse, G.; Spiller, P.**, 1998, “*Electronic Shopping*”, CACM 41,7 (Jul-98); pp 81-86.
- [Lowe et al 98] **Lowe, D., Webby, R.**; 1998, “*The Impact Process Modelling Project*”, 1st International Workshop on Hypermedia Development, at ACM Hypertext 98, Pittsburgh, US, <http://ise.ee.uts.edu.au/hypdev/>
- [Lowe et al 99] **Lowe, D., Hall, W.**; 1999, “*Hypermedia and the Web: an Engineering Approach*”, John Wiley & Sons
- [McCall et al 77] **McCall, J.A.; Richards, P.K.; Walters, G.F.**; 1977, “*Factors in Software Quality*”, RADC TR-77-369.
- [Mendez et al 99] **Mendes, M. E. X., Harrison, R., And Hall, W.**; 1999. *Applying Measurement Principles to Improve Hypermedia Authoring*, To appear in New Review of Hypermedia and Multimedia Journal.
- [Miller 77] **Miller, J.R.**; 1970, “*Professional Decision-Making*”, Praeger Publisher.
- [Madhavji 91] **Madhavji, N.H.**, 1991, “*The Process Cycle*”, IEEE Software Engineering. Journal, (Aug 91).
- [Mu_Gal_Art 98] **National Gallery of Art Museum**, 1998 (versión a Oct. 98), <http://www.nga.gov>
- [Mu_Louvre 98] **Louvre Museum**, 1998 (versión a Oct. 98), <http://www.louvre.fr>
- [Mu_Met 98] **Metropolitan Museum**, 1998 (versión a Oct. 98), <http://www.metmuseum.org>
- [Mu_Prado 98] **Museo del Prado**, 1998 (versión a Oct. 98), <http://museoprado.mcu.es>
- [Nanard et al 95] **Nanard, J.; Nanard, M.**, 1995, “*Hypertext Design Environment and the Hypertext Design Process*”, Comm. ACM 38, 8 (Aug 95) pp. 49-56
- [Nielsen 93] **Nielsen, J.**; 1993, “*Usability Engineering*”, Academic Press, 1993.
- [Nielsen_Alert] **Nielsen, J.**; The Alertbox, <http://www.useit.com/alertbox/>
- [Olsina et al 95] **Olsina, L.; Nicolau, S.; Irastorza, J.; Bertone, E.**, 1995, “*Estrategias y Criterios de Diseño e Implementación Hipermediales en el proyecto Facultad de Ingeniería*”, Informe

- de proyecto de I&D, Dep. de Informática, Fac. Ingeniería, UNLPam.
- [Olsina 96] **Olsina, L.**, “*View of a Process Model to Develop Hypermedia*” (in Spanish), Proceed. of the IV Congress of the SCCC (Computer Science Chilean Society), Valdivia, Chile, 1996, pp. 219-228.
- [Olsina 97a] **Olsina, L.**, 1997, “*Systematic use of Flexible Process Model to build Hypermedia Artifacts*”. Poster Sesion, Hypertext 97, Southampton, England
- [Olsina 97b] **Olsina, L.**, 1997, “*Object-Oriented Flexible Prototyping to support Hypermedia Flexible Process Model*”. III Workshop em Sistemas Multimídia e Hiperemídia (WoMH 97), Sao Carlos, Brasil, pp. 3-14.
- [Olsina 97c] **Olsina, L.**, 1997, “*Applying the Flexible Process Model to build Hypermedia Products*”. Hypertext and Hypermedia: Tools, Products, Methods (HHTPM 97), Paris, France, Ed. Hermes, Vol I N. 1-2-3 pp. 211-221.
- [Olsina 98a] **Olsina, L.**, 1998, “*Hypermedia Engineering: Flexible Process Model to support Hypermedia Application Development*”, Master Thesis on Software Engineering (in Spanish), UNLP, La Plata, Ar, Feb 98.
- [Olsina 98b] **Olsina, L.**, 1998, “*Functional View of the Hypermedia Process Model*”. The Fifth International Workshop on Engineering Hypertext Functionality at ICSE'98 (Int'l Conference on Software Engineering) , Kyoto, Japan, (Disponibile electrónicamente en <http://www.ics.uci.edu/~kanderso/htf5/papers/>).
- [Olsina 98c] **Olsina, L.**, 1998, “*Building a Web-based Information System applying the Hypermedia Flexible Process Modeling Strategy*”; 1st International Workshop on Hypermedia Development, at ACM Hypertext 98, Pittsburgh, US, <http://ise.ee.uts.edu.au/hypdev/>
- [Olsina 98d] **Olsina, L.**, 1998, “*Toward Improvements in Hypermedia Process Modeling*”, Proceed. at the Symposium of Software Technology, 27 JAIIO, Bs As., pp. 219-226
- [Olsina 98e] **Olsina, L.**, 1998, “*Overview of a framework to Hypermedia Process Modeling*, Proceed. del IV Congreso Argentino de Ciencias de la Computación. (CACIC), Neuquén, Argentina, Oct 1998, Vol I, pp. 207-217
- [Olsina 99] **Olsina, L.**, 1999, “*Web-site Quantitative Evaluation and Comparison: a Case Study on Museums*”, (Int'l Conference on Software Engineering) ICSE '99 Workshop on Software Engineering for the Internet, Los Angeles, US (Disponibile

- electrónicamente en <http://sern.cpsc.ucalgary.ca/~maurer/ICSE99WS/ICSE99WS.html>).
- [Olsina et al 99a] **Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.;** 1999, "*Quality Characteristics and Attributes for Academic Web Sites*", WWW8 Web Engineering'99 Workshop, Toronto, Canada (Disponible electrónicamente en http://budhi.uow.edu.au/webengineering99/web_engineering.html).
- [Olsina et al 99b] **Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.;** 1999, "*Specifying Characteristics and Attributes for Web Sites*", Proceed. of ICSE '99 (Int'l Conference on Software Engineering) Web Engineering Workshop, Los Angeles, US, pp 84-93.
- [Olsina et al 99c] **Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.;** 1999, "*Assessing the Quality of Academic Websites: a Case Study* ", A publicar en New Review of Hypermedia and Multimedia (NRHM) Journal, Taylor Graham Publishers, UK.
- [Olsina et al 99d] **Olsina, L.; Rossi, G.;** 1999, "*Towards Web-site Quantitative Evaluation: defining Quality Characteristics and Attributes*", Proceedings of IV Int'l WebNet Conference, World Conference on the WWW and Internet, Hawaii, USA, Vol 1-2, pp.834-839.
- [Osterweil 87] **Osterweil, L** 1987, "*Software Processes are Software Too*". Proceed. of the Ninth International Conference of Software Engineering, pp2-13, Monterey CA.
- [Paulk et al 93] **Paulk, M.C.; Curtis, B.; Chrissis, M.B.; Weber, C.V,** 1993, "*Capability Maturity Model for Software, Version 1.1*", SEI Technical Report SEI-CMU-93-TR-24, Pittsburgh, US.
- [Paulk et al 94] **Paulk, M.C.; Weber, C.V; Curtis, B.,** 1994, "*Capability Maturity Model for Software: Guidelines for Improving the Software Process*", Addison-Wesley, MA.
- [Powell et al 98] **Powell, T.; Jones, D; Cutts, D.,** 1998, "*Web Site Engineering: Beyond Web Page Design*", Prentice Hall PTR.
- [Pressman 93] **Pressman, R.,** 1993, "*Ingeniería de Software. Un enfoque Práctico*", Mac Graw Hill.
- [Pressman 98] **Pressman, R.,** 1998, "*Can Internet-based Applications be Engineered*", IEEE Software (Sep/Oct 98).
- [Roberts 79] **Roberts, F.,** 1979, "*Measurements Theory with Applications to Decisionmaking, Utility, and the Social Sciences*", Encyclopedia of Mathematics and its Applications Addison-Wesley Publishing Co.
- [Rosenfeld et al 98] **Rosenfeld, L., Morville, P.,** 1998, "*Information Architecture for the WWW*", O'Reilly.

- [Rossi 96] **Rossi, G.**, 1996, “*Uma metodologia Orientada a Objetos para o projeto de aplicativos Hipermídia*”, Tesis Doctoral, PUC-RIO, Río de Janeiro, Br.
- [Rossi et al 97] **Rossi, G., Schwabe, D.; Garrido, A.**, 1997, “*Design reuse in Hypermedia Application Development*”, Proceedings of ACM Hypertext 97, Southampton, England pp 57-66
- [Rossi et al 99] **Rossi, G., Schwabe, D.; Lyardet, F.** ; 1997, “*Improving Web Information Systems with navigational Patterns*”, Proceedings of WWW8 Congress, Toronto, Canada, pp. 589-600
- [Royce 70] **Royce W.**, 1970, “*Managing the Development of Large Software System*”, IEEE WESCON pp 1-9. Reprinted in Ninth IEEE International Conference on Software Engineering, Washington DC: Computer Society Press of IEEE, 1987 pp.328-338.
- [Rumbaugh et al 91] **Rumbaugh, J; Blaha, M; Premerlani, W; Eddy, F; Lorensen, W.** ,1991, “*Object-Oriented Modelling and Design*”, Prentice Hall.
- [Schneiderman 98] **Schneiderman, B.**, 1998, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd Edition, Reading, MA: Addison Wesley.
- [Schwabe et al 94] **Schwabe, D.; Barbosa, S.**, 1994, “*Navigational Modelling in Hypermedia Application*”, Technical Report MCC 42/94, Dep. de Informática PUC-Rio, Brasil
- [Schwabe et al 95a] **Schwabe, D.; Rossi, G.**, 1995, “*Building Hypermedia Application as Navigational Views of Information Model*”, Proceed. of the 28th Hawaii International Conference on System Science, Jan 95, Vol 3, pp. 231-240
- [Schwabe et al 95b] **Schwabe, D.; Rossi, G.**, 1995, “*The Object-Oriented Hypermedia Design Model*”, Comm. ACM 38,8 (Aug 95) pp. 45-46
- [Schwabe et al 96] **Schwabe, D.; Rossi, G. Barbosa, S.**, 1996, “*Systematic Hypermedia Application Design with OOHDM*”, Hypertext 96, US
- [Sommerville 92] **Sommerville, I.**, 1992, “*Software Engineering*”, Addison-Wesley, 4th Ed.
- [Thüring et al 95] **Thüring, M.; Hannemann, J.; Haake, J.**, 1995, “*Hypermedia and Cognition: Designing for Comprehension*”, Comm. ACM 38, 8 (Aug 95) pp. 57-66
- [Tilson et al 98] **Tilson, R., Dong, J., Martin, S., Kieke, E.**, 1998, “*Factors and Principles Affecting the Usability of Four E-commerce Sites*”, 4th Conference on Human Factors & the Web, Baking Ridge, NJ, U S.

- [UML 97] UML Notation Guide, 1997, Ver 1.0, y superiores, <http://www.rational.com>
- [Un_Chile 99] **Universidad de Chile**, 1999 (versión a Feb. 99), <http://www.uchile.cl>
- [Un_España 99] **Universidad Politécnica de Cataluña**, 1999 (versión a Feb. 99), <http://www.upc.es>
- [Un_Singapur 99] **University of Singapore**, 1999 (versión a Feb. 99), <http://www.nus.sg>
- [Un_Canadá 99] **University of Quebec**, 1999 (versión a Feb. 99), <http://www.uqam.ca>
- [Un_Estados_Un 99] **University of Stanford**, 1999 (versión a Feb. 99), <http://www.stanford.edu>
- [Un_Australia 99] **University Technological of Sidney**, 1999 (versión a Feb. 99), <http://www.uts.edu.au>
- [Verlag et al 98] **Verlage, M; Webby, R; et. al.**, 1998, “*The Spearmint Approach to Process Definition and Process Guidance*”, Workshop on S.Eng. over the Internet at ICSE’98, Kyoto, Japan. (Disponible electrónicamente en <http://sern.cpsc.ucalgary.ca/~maurer/ICSE98WS/ICSE98WS.html>)
- [White et al 63] **White, D.R.J.; Scott, D.L; Schultz**, 1963, “*POED: a Method of Evaluating System Performance*”, IEEE TEM, (Dec. 63).
- [W3C 99] **WWW Consortium**, 1999, W3C Working Draft, “WAI Accessibility Guidelines: Page Authoring”, W3C, 1999, W3C Working Draft, “WAI Accessibility Guidelines: Page Authoring”, <http://www.w3c.org/TR/WD-WAI-PAGEAUTH/>
- [Yu et al 98] **Yu, J.J., Prabhu, P.V; Neale, W.C.**; 1998, “*A user-centred approach to designing a new top-level structure for a large and diverse corporate Web site*”, 4th Conference on Human Factors and the Web, Basking Ridge, NJ, US. Available at URL: <http://www.research.att.com/conf/hfweb/proceedings/kirakowski/index.html>
- [Zeff et al 97] **Zeff, R., Aronson, B;** 1997, “*Advertising on the Internet*”, John Wiley & Sons, Inc.
- [Zuse 98] **Zuse, H.**, 1998, “*A Framework of Software Measurement*”, Walter de Gruyter, Berlín-NY.

APENDICE A

Información Adicional de los Estudios de Campo Empleados.

A.1 Información Adicional para el Caso de Estudio de Sitios Académicos

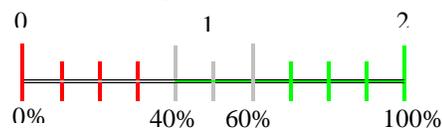
A.1.1 Plantillas de Características y Atributos de Calidad (Ver Capítulos 6 y 9)

Título: *Nivel de Personalización en la Recuperación*; **Código:** 2.1.2.1; **Tipo:** Atributo
Característica de más Alto Nivel: Funcionalidad
Super-característica: Mecanismo de Recuperación

Definición / Comentarios: Este atributo representa el nivel de configuración permitido a los usuarios para ver la información en la pantalla, luego que la búsqueda se ha ejecutado. Por ejemplo, se puede personalizar la cantidad de información a ser recuperada para cada ocurrencia hallada, como así también la cantidad de ocurrencias por pantalla. Además, las ocurrencias se pueden mostrar al usuario ordenadas de diferente manera.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde se describe el nivel de personalización: 0 = no disponible; 0.5 = bajo; 1=alto.

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional

Ejemplo/s: Un ejemplo, es el nivel de seteo en la recuperación de información permitido al usuario cuando por ejemplo está en la búsqueda de personas, en el sitio Stanford (<http://sin.stanford.edu:2000/frame?person>) como se ilustra en la figura 6.10. La preferencia elemental computada es 100%.

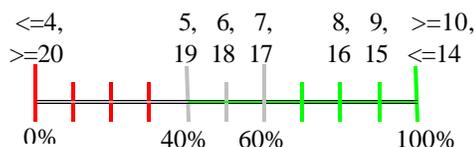
Título: *Promedio de Enlaces por Página*; **Código:** 2.2.1.2.; **Tipo:** Atributo
Característica de más Alto Nivel: Funcionalidad
Super-característica: Navegabilidad

Definición / Comentarios: Cada página o nodo tiene generalmente dos tipos principales de enlaces (o puntos de partida): enlaces semánticos (distribuidos en la información de las páginas del sitio o contenido), y enlaces estructurales (a cerca de controles principales y navegacionales). Considerando la heurística o regla práctica de G. Miller (publicada en el artículo: “*The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*”), una persona puede tener siete más-menos-dos threads de procesamiento de información, en un momento dado. Podemos considerar que una página en promedio tiene esta cantidad de enlaces

semánticos. También, a partir de estudios realizados, una página en promedio tiene cinco controles principales y/o navegacionales.

Este atributo nos da una idea sobre la cantidad de puntos de partida en una página promedio que indica cómo una página promedio está interconectada hacia nodos destinos. Un nivel alto de interconectividad puede ser contraproducente para la navegabilidad al permitir potencialmente seleccionar destinos diversos que pueden causar desorientación del usuario. En un nodo altamente poblado de anchors el usuario no tiene un camino estructurado para seguir un concepto distribuido en varios nodos. El caso opuesto (un nivel bajo de anchors en los nodos) generalmente puede ser una indicación de un diseño pobre [Botafogo 92].

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde entre 10 y 14 anchors en promedio, se corresponde a una preferencia de 100%; 9 o 15 produce el 90%; 8 o 16 produce el 80%; 7 o 17 produce el 60%; 6 o 18 produce el 50%; 5 o 19 produce el 30%; y ≤ 4 o ≥ 20 anchors en promedio, produce el 0%.



Escala de Preferencia:

Tipo de Recolección de Datos: Automatizado.

Ejemplo/s: Por ejemplo, Stanford arrojó un promedio de todas sus páginas de 11 anchors, UQAM arrojó 12, produciendo ambos sitios una preferencia elemental de 100%. El nivel más bajo fue observado en el sitio de Chile (5), y en el de la UPC (6). El nivel más alto fue para UTS (17 anchors, produciendo ambos sitios una preferencia elemental de 60 %)

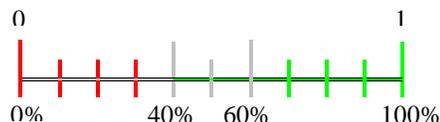
Título: *Indice de Carreras*; **Código:** 2.3.1.3.1; **Tipo:** *Atributo*

Característica de más Alto Nivel: *Funcionalidad*

Super-característica: *Información de Carreras*

Definición / Comentarios: Este mecanismo generalmente implementado en el subsitio (no en la página principal), permite estructurar por medio de enlaces a través de un ordenamiento de carreras, la navegación al contenido de las mismas.

Tipo de Criterio Elemental: es un criterio binario, discreto y absoluto: sólo se pregunta si está disponible (1) o si no está disponible, (0).



Escala de Preferencia:

Tipo de Recolección de Datos: Manual, Observacional.

Ejemplo/s: Todos los sitios evaluados lo poseían

Título: *Programa Sintético/Analítico*; **Código:** 2.3.1.3.4.2 ; **Tipo:** *Atributo*

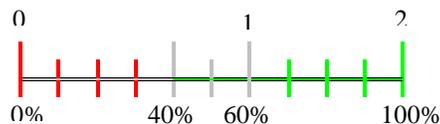
Característica de más Alto Nivel: *Funcionalidad*

Super-característica: Descripción de Cursos

Definición / Comentarios: Este atributo representa la disponibilidad de descripciones sintéticas o analíticas de los cursos ofrecidos en una o más carreras.

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto; en donde: 0 = no disponible; 1 = disponible parcialmente (no en todos los cursos); 2 = disponible totalmente

Escala de Preferencia:



Tipo de Recolección de Datos: Manual, Observacional.

Ejemplo/s: Solamente los sitios de la UChile y UPC lo tenían .

Título: *Número de Vistas considerando Marcos (frames); Código:* 4.2.2.1;

Tipo: Atributo

Característica de más Alto Nivel: Eficiencia

Super-característica: Accesibilidad de Ventanas

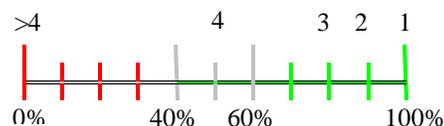
Definición / Comentarios: Los frames or marcos organizan a una ventana en diferentes áreas o subvistas tanto de control como de contenido. Cuanto mayor es la cantidad de frames, menor es la accesibilidad de la ventana, principalmente para personas con discapacidades.

La guía provista por W3C en WAI Accessibility Guidelines [W3C 99], dice: *‘For visually enabled users, frames may organize a page into different zones. For non-visual users, relationships between the content in frames (e.g., one frame has a table of contents, another the contents themselves). must be conveyed through other means’*

Este atributo está relacionado con el 4.2.2.2 (*Versión sin Marcos*)

Tipo de Criterio Elemental: es un criterio multi-nivel, discreto y absoluto: X=Número de vista o subventanas (considerando frames). De manera que, X = 1-> 100 %; X = 2 -> 90 %; X = 3 -> 80%; X = 4 -> 50 %; y, X > 4 -> 0 %

Escala de Preferencia:



Tipo de Recolección de Datos: Automatizado con Web-siteMA.

Ejemplo/s:

- 1) Sitios que usan frames son NUS, UPC, Chile, y UQAM. El lector puede apreciar, en la figura 4.4 los frames en la página principal del sitio de la UQAM. Analizando con la herramienta el código HTML de este sitio arrojó que tiene un máximo de 4, lo cual produce una preferencia de 50%.
- 2) Un fragmento de código HTML mostrando la etiqueta (tag) FRAME, entre otros, los cuales permiten ver cómo automatizar la medida por medio del lenguaje WebL:

```

<FRAMESET cols="10%, 90%"
  title="Our library of electronic documents">
  <FRAME src="nav.html" title="Navigation bar">
  <FRAME src="doc.html" title="Documents">
  <NOFRAMES>
    <A href="lib.html" title="Library link">
      Select to go to the electronic library</A>
  </NOFRAMES>
</FRAMESET>

```

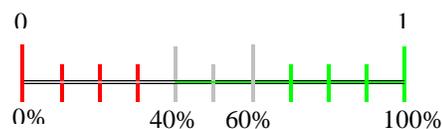
Título: *Versión sin Marcos (frames)*; **Código:** 4.2.2.2; **Tipo:** Atributo

Característica de más Alto Nivel: Eficiencia

Super-característica: Accesibilidad de Ventanas

Definición / Comentarios: Cuando un sitio utiliza la estrategia de frames o marcos es deseable que cuente con una opción de versión del sitio sin frames, en consideración de la accesibilidad, de acuerdo a lo comentado para el atributo 4.2.2.1. Este atributo está relacionado con el 4.2.2.1 (*Número de Vistas considerando Marcos*)

Tipo de Criterio Elemental: es un criterio binario, discreto y absoluto: sólo se pregunta si está disponible la versión sin frames (1) o si no está disponible, o no es aplicable (en caso de no estar diseñado con frames) (0).



Escala de Preferencia:

Tipo de Recolección de Datos: Manual, Observacional.

Ejemplo/s: De los sitios que poseían frames sólo UChile disponía versión sin frames

A.1.2 Resultados de las Preferencias Elementales para los Sitios Universitarios

Tabla A.1.1 Resultados de las Preferencias Elementales para los Sitios Académicos

	UPC	UChile	UTS	NUS	Stanford	UQAM
Usabilidad						
1.1.1.1	100	0	0	0	0	0
1.1.1.2	100	0	100	100	100	0
1.1.1.3	0	0	100	0	100	0
1.1.2	90	90	90	80	90	80
1.1.3	0	0	100	0	100	0
1.1.4	100	100	100	100	50	100
1.2.1.1	50	40	90	50	70	60
1.2.1.2	100	60	100	60	100	100
1.2.2.1	60	0	100	100	100	0
1.2.2.2	0	0	100	100	100	100
1.2.3.1	100	100	100	100	100	100
1.2.3.2	100	100	100	100	100	100
1.2.3.3	100	100	100	100	100	100
1.2.4	100	0	0	0	80	0

1.2.5.1	0	0	100	100	0	0
1.2.5.2	0	0	0	0	0	0
1.2.5.3	100	100	0	100	0	100
1.3.1	100	100	100	70	100	100
1.3.2.1	80	80	100	30	100	50
1.3.2.2	80	50	80	30	80	100
1.3.2.3	100	80	100	30	80	80
1.3.3.1	60	60	60	20	60	100
1.3.3.2	60	60	60	20	60	60
1.3.3.3	100	0	0	0	0	0
1.3.4	80	60	90	40	70	50
1.4.1	90	0	0	0	0	90
1.4.2	80	80	100	100	100	100
1.4.3	0	100	0	0	0	0
Funcionalidad						
2.1.1.1.1	60	100	60	100	100	100
2.1.1.1.2	0	0	0	0	100	0
2.1.1.1.3	0	0	100	0	100	100
2.1.1.2	60	60	60	0	100	100
2.1.2.1	50	50	0	0	100	50
2.1.2.2	50	0	0	0	100	100
2.2.1.1.1	0	0	0	0	70	0
2.2.1.1.2	70	70	70	70	70	100
2.2.1.2	50	40	60	60	100	100
2.2.2.1.1	80	50	80	0	80	30
2.2.2.1.2	100	70	100	0	70	70
2.2.2.2.1	80	80	80	80	80	80
2.2.2.2.2	50	50	100	50	0	50
2.2.3.1	100	50	0	70	100	0
2.2.3.2	90	90	90	90	90	80
2.3.1.1.1	100	100	100	100	100	100
2.3.1.1.2	100	100	100	100	100	57,14
2.3.1.2.1	100	100	0	100	100	100
2.3.1.2.2	0	0	0	100	0	0
2.3.1.3.1	100	100	100	100	100	100
2.3.1.3.2	100	100	100	0	100	100
2.3.1.3.3	100	100	100	0	0	100
2.3.1.3.4.1	0	100	100	0	100	100
2.3.1.3.4.2	100	100	0	0	0	0
2.3.1.3.4.3	100	0	100	0	0	0
2.3.1.4.1	100	100	100	0	100	100
2.3.1.4.2	100	0	100	100	100	0
2.3.1.4.3	0	100	100	100	100	100
2.3.1.4.4	100	0	100	100	100	100
2.3.1.4.5	100	100	100	100	100	100
2.3.1.5.1	100	100	100	100	100	100
2.3.1.5.2	100	100	100	100	100	100
2.3.1.5.3	0	100	100	100	100	100
2.3.2.1	0	0	100	100	100	100
2.3.2.2	0	0	0	100	100	100
2.3.2.3	100	0	0	0	0	100
2.3.2.4	100	0	0	0	0	100

Confiabilidad						
3.1.1.1	0	75,02	74,1	68,06	58,32	0
3.1.1.2	100	100	100	100	100	100
3.1.1.3	100	75	100	100	100	100
3.1.2.1	100	100	100	100	100	100
3.1.2.2	100	100	100	100	100	100
3.1.2.3	50	100	100	100	100	100
Eficiencia						
4.1.1	75,3	50,46	82	51,46	100	83,44
4.2.1.1	0	0	0	0	0	0
4.2.1.2.1	34,38	45,36	81,65	36,22	47,29	53,15
4.2.1.2.2	80	50	50	60	50	50
4.2.2.1	80	90	100	100	100	50
4.2.2.2	0	100	0	0	0	0

A.2 Información Adicional del Caso de Estudio de Museos en la Web

Tabla A.2.1 Resultados de las Preferencias Elementales para los Sitios de Museos

1. Usabilidad

A_i	Louvre IE_i	Prado IE_i	Metropolitan IE_i	Gallery of Art IE_i
1.1.1.1	0	0	0	0
1.1.1.2	0	0	0	100
1.1.1.3	0	0	0	100
1.1.2.1	80	80	80	100
1.1.2.2	0	80	50	0
1.1.3	50	100	50	100
1.1.4	100	80	80	0
1.2.1.1	60	60	60	100
1.2.1.2	0	100	0	100
1.2.2.1	100	100	0	0
1.2.2.2	0	0	0	0
1.2.3.1	100	100	100	100
1.2.3.2	100	100	100	100
1.2.3.3	100	100	100	100
1.2.4	0	0	100	80
1.2.5	100	0	0	0
1.3.1	30	70	70	70
1.3.2.1	0	80	80	100
1.3.2.2	100	0	0	0
1.3.2.3	100	80	100	100
1.3.3	40	80	80	90
1.3.4	100	60	60	100
1.4.1	90	60	0	24
1.4.2	0	0	0	0

2. Funcionalidad

2.1.1.1	0	100	0	100
2.1.1.2	0	0	0	70
2.2.1.1	80	80	80	80
2.2.1.2.1	0	70	70	70
2.2.1.2.2	70	70	70	70

2.2.2.1	80	80	80	80
2.2.3.1.1	0	30	80	30
2.2.3.1.2	0	70	100	30
2.2.3.2.1	70	70	70	70
2.2.3.2.2	100	100	100	100
2.2.4.1	0	0	70	70
2.2.4.2	80	80	80	90
2.3.1	80	80	80	80
2.3.2	80	80	80	80
2.3.3.1.1	0	100	100	100
2.3.3.1.2	0	80	80	80
2.3.3.2	0	0	100	100
2.3.4.1	0	100	50	0
2.3.4.2	100	100	100	100

3. Confiabilidad

3.1.1.1	100	100	0	100
3.1.1.2	100	75	100	100
3.1.1.3	100	50	0	100
3.1.2.1	100	100	100	100
3.1.2.2	100	50	100	100
3.1.2.3	0	100	0	0
3.1.2.4	100	100	100	100

4. Eficiencia

4.1.1	0	0	0	0
4.1.2.1	0	0	50	50
4.1.2.2	80	80	70	90

A.3 Un Survey sobre 24 Museos

A.3.1 Preparación del Survey

Para este survey hemos tomado una muestra de 24 museos publicados en servidores distribuidos alrededor del mundo. La muestra fue seleccionada de uno de los índices en la Web más relevantes para el dominio, y el mismo estaba ordenado por continentes y países. Para la selección de la muestra se utilizaron criterios específicos como tamaño del sitio (se empleó la herramienta SiteSweeper para obtener el tamaño en páginas), que sea típico en el dominio de museos y, además, por estratos geográficos y culturales. En la tabla A.3.2 (al final) se especifica el nombre de cada museo y su URL respectivo.

Se evaluaron 29 atributos correspondientes a tres características de alto nivel, prescritas en el estándar ISO/IEC 9126. La figura A.3.1 muestra un árbol mínimo de requerimientos de calidad en consideración del visitante general, extraído del caso de estudio de museos (ver sección 9.2). En la tabla A.3.3 del apéndice se realiza una somera descripción para cada atributo. En cuanto a la muestra seleccionada, como se mencionó previamente, se tuvo en cuenta que los museos debieran cumplimentar algunos criterios tales como:

- ✓ Que sean típicos en su dominio y de renombre nacional o internacional.
- ✓ Que el sitio posea un tamaño no menor que el establecido en cuanto a cantidad de nodos (no menor de 100 páginas)
- ✓ Que pertenecieran a diferentes continentes y regiones (no pudiendo seleccionar aquellos que tuvieran un único lenguaje (y nativo) no comprensible para los evaluadores, como mandarín, hebreo, árabe, etc.).

1. Usabilidad

1.1 Comprensibilidad Global del Sitio

- 1.1.1 Esquema de Organización Global
 - 1.1.1.1 *Mapa del Sitio*
 - 1.1.1.2 *Índice (Alfabético, cronológico, etc.)*
 - 1.1.1.3 *Tabla de Contenidos*
- 1.1.2 Mecanismos de visita Guiada
 - 1.1.2.1 *Visita Guiada Convencional*
 - 1.1.2.2 *Visita Guiada Virtual*
- 1.1.3 *Mapa de Imagen (Edificio y Salas)*

1.2 Mecanismos de Ayuda y Retroalimentación

- 1.2.1 Calidad de la Ayuda
 - 1.2.1.1 *Ayuda Explicatoria Orientada al Visitante*
 - 1.2.1.2 *Ayuda de la Búsqueda*
- 1.2.2 Indicador de Última Actualización
 - 1.2.2.1 *Global (de todo el sitio Web)*
 - 1.2.2.2 *Restringido (por subsitio o página)*
- 1.2.3 Directorio de Direcciones
 - 1.2.3.1 *Directorio E-mail*
 - 1.2.3.2 *Directorio TE-Fax*
- 1.2.4 *Comentarios/Sugerencias*
- 1.2.5 *Facilidad FAQ*

1.3 Aspectos de Interfaces y Estéticos

- 1.3.1 Permanencia y Estabilidad en la Presentación de los Controles Principales
 - 1.3.1.1 *Permanencia de Controles Directos*
 - 1.3.1.2 *Permanencia de Controles Indirectos*

1.3.1.3 *Estabilidad*

1.4 Misceláneas

- 1.4.1 *Soporte a Lenguaje Extranjero*
- 1.4.2 *Indicador de Tamaño de la Imagen*
- 1.4.3 *Zooming de la Imagen*

2. Funcionalidad

2.1 Aspectos de Búsqueda y Recuperación

- 2.1.1 Mecanismo de Búsqueda en el Sitio
 - 2.1.1.1 *Búsqueda Restringida*
 - 2.1.1.2 *Búsqueda Global*

2.2 Aspectos de Navegación y Exploración

- 2.2.1 Navegabilidad
 - 2.2.1.1 Orientación
 - 2.2.1.1.1 *Indicador del Camino*
 - 2.2.1.1.2 *Etiqueta (label) de la Posición Actual*
- 2.2.2 Objetos de Control Navegacional
 - 2.2.2.1 Permanencia y Estabilidad en la Presentación de los Controles Contextuales (Subsitio)
 - 2.2.2.1.1 *Permanencia de los Controles Contextuales*
 - 2.2.2.1.2 *Estabilidad*
- 2.2.3 Predicción Navegacional
 - 2.2.3.1 *Enlace con Título (enlace con texto explicatorio)*

3. Eficiencia

3.1 Accesibilidad

- 3.1.1 Accesibilidad de Información
 - 3.1.1.1 *Soporte a Versión sólo Texto*
 - 3.1.1.2 *Imagen con Título*

3.2 Performance

- 3.2.1 *Tamaño de la Página (*)*

Figura A.3.1. *Arbol de requerimientos de características, subcaracterísticas y atributos en consideración del visitante general para museos en la Web.*

Si bien los atributos a evaluar pueden poseer uno mayor interés que otros desde el punto de vista del visitante, no se consideraron pesos relativos para cada atributo (a diferencia de los casos de estudio), pero en cuanto a la determinación de los criterios elementales se tuvieron en cuenta dos tipos, dentro de los absolutos y discretos, a saber: binarios y multinivel (particularmente, de tres niveles). Los criterios son: *disponible* (1); *no disponible* (0); y *parcialmente disponible* (0.5). Estos valores medidos para los atributos (1, 0, y 0.5) se deben interpretar respectivamente como que satisface

completamente el requerimiento de calidad (el 100%); que no lo satisface (0%), o que lo cumplimenta parcialmente (50%), en consideración de la audiencia especificada.

El valor “*parcialmente disponible*” fue aplicado a los siguientes atributos en el proceso de evaluación, conforme al criterio antes mencionado:

- 1.2.1.1 *Ayuda Explicatoria Orientada al visitante*
- 1.2.2.2 *Indicador de Última Actualización (Restringido a subsitios o páginas)*
- 1.2.3.1 *Directorio E-mail*
- 1.2.3.2 *Directorio TE-Fax*
- 1.3.1.2 *Permanencia de los Controles Indirectos*
- 1.4.1 *Soporte a Lenguaje extranjero*
- 1.4.2 *Indicador de Tamaño de la Imagen*
- 1.4.3 *Zooming de la Imagen*
- 2.2.1.1.1 *Indicador de Camino*
- 2.2.1.1.2 *Etiqueta de la Posición Actual*
- 2.2.2.1.1 *Permanencia de los Controles Contextuales*
- 2.2.2.1.2 *Estabilidad (de los Controles Contextuales)*
- 2.2.3.1 *Enlace con título (texto explicatorio)*
- 3.1.1.1 *Soporte a Versión sólo Texto*
- 3.1.1.2 *Imagen con Título*

Finalmente, los mecanismos válidos que se pueden utilizar para recolectar datos pueden ser: *Manual* y *observacional*, *Semiautomático*, y *Automático*. En este survey el mecanismo de recolección que se empleó fue el manual y observacional (excepto la herramienta utilizada para determinar el tamaño del sitio). Por último, los datos fueron capturados (con la participación de tres evaluadores) en el período que va desde el 15 de noviembre, al 20 de diciembre de 1998 (fue realizado con posterioridad al caso de estudio sobre museos). En tal período no se observaron cambios (con respecto a cada sitio) que pudieran haber afectado el proceso de evaluación.

A.3.2. Análisis de los Datos y Tendencias

En la tabla A.3.1, se encuentran registrados los valores individuales de cada atributo para cada museo y los porcentajes de disponibilidad del atributo.

A.3.2.1 Análisis para Usabilidad. Observando los resultados obtenidos de la encuesta, podemos apreciar que en cuanto a los museos que presentan un *Esquema de Organización Global* (subcaracterística que favorece la rápida comprensión tanto de la estructura global como del contenido del sitio) a través de los atributos *Mapa del Sitio* (código 1.1.1.1), *Índice* (1.1.1.2) y *Tabla de Contenidos* (1.1.1.3), el 41,67% de la muestra presenta al menos uno de estos mecanismos de organización (ver figura A.3.2). Los porcentajes individuales por atributos son los siguientes:

- ✓ *Mapa del Sitio arrojó 0%*
- ✓ *Índice Global (alfabético, cronológico, etc.), 29,17%*
- ✓ *Tabla de Contenidos, 41,67%*

De estos resultados podemos decir que de los museos que poseen *Indice* también poseen *Tabla de Contenidos*. Este aspecto de mostrar la *Organización Global del Sitio* es importante para poder acceder a la información más rápidamente, o para conocer en que área del sitio se encuentra cierta información a explorar (browsing). Los resultados muestran que ningún museo posee *Mapa del Sitio*, el cual sería un atributo deseable para conocer la estructura organizativa o arquitectura del sitio.

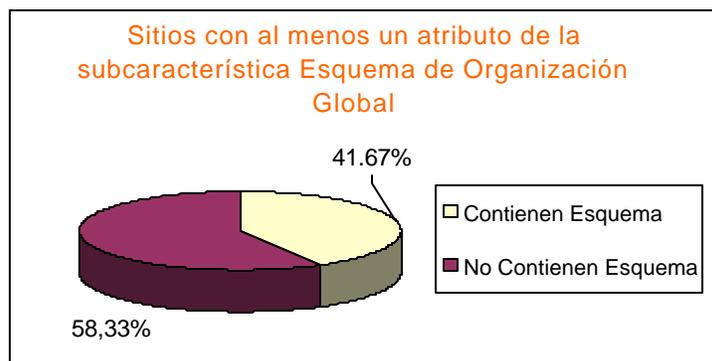


Figura A.3.2. Sitios que poseen al menos un atributo de la subcaracterística Esquema Global del Sitio

En consideración del dominio, lo que es de interés para los visitantes generales son las obras que los museos poseen, y una característica deseable es alguna forma de recorrer rápidamente este contexto navegacional del sitio. Los mecanismos más comunes utilizados por los museos son la *Visita Guiada Convencional* (1.1.2.1) y la *Visita Guiada Virtual* (1.1.2.2). Este último atributo no sólo permite mostrar las obras, sino que también se puede pasear virtualmente por la sala en donde se exponen las obras, permitiendo no sólo apreciar a las mismas sino que también se puede apreciar la ubicación dentro de la sala o galería, entre otros detalles.

Los porcentajes totales que la encuesta arrojó para estos atributos fueron del 54,17% para el atributo *Visita Guiada Convencional*, y el 37,5% para el otro atributo. Además podemos decir que:

- ✓ El 25% de la encuesta posee los dos mecanismos de visita (1.1.2.1 y 1.1.2.2).
- ✓ El 66,67% de la encuesta posee al menos un mecanismo de visita guiada (atributos 1.1.2.1 o 1.1.2.2).
- ✓ El 29,17% posee solamente el atributo *Visita Guiada Convencional*.
- ✓ El 12,5% posee solamente el atributo *Visita Guiada Virtual*.

Si bien la encuesta indica que el 66,67% posee al menos un mecanismo de visita, debemos tener en cuenta que para poder acceder a las visitas virtuales los navegadores deben soportar el formato “mov”, lo cual impediría a algunos visitantes no poder visualizarlas debido a que esta tecnología es más recientemente soportada en los

navegadores. Si no consideráramos esta tecnología, podemos deducir que el 54,17% de la totalidad de la muestra poseen mecanismos de visitas guiadas que están adaptados para ser utilizados por cualquier navegador (que soporte gráficos).

Además, considerando una audiencia general, es de utilidad en un museo conocer la disposición de los componentes del edificio en pisos y salas de exhibición, es decir, es deseable contar con el plano navegable del museo (atributo 1.1.3 - *Mapa de Imagen*). Esto permite a la audiencia conocer dónde están las obras que jerarquizan al museo, las cuales son de mayor o menor atracción turística y educativa del mismo. Para este atributo se tuvo en cuenta el criterio de algún mecanismo de navegación a partir del mapa o imagen del edificio. El porcentaje que arrojó la muestra es del 29,17 %.

Ahora bien, si ponemos el foco de atención en las imágenes estáticas (importantes, por ejemplo, para las *Visitas Guiadas Convencionales*), uno de los aspectos a tener en cuenta es el *Zomming de la Imagen* de las obras (atributo 1.4.3) para una mejor apreciación, en donde la encuesta arrojó el 83,33% de disponibilidad. En cuanto al atributo *Indicador de Tamaño de la Imagen* (1.4.2 – y no nos referimos al tamaño real de la obra en el museo, sino al tamaño de la imagen que es mostrada a los visitantes lo cual permite conocer el tamaño antes de descargarla), el porcentaje de cumplimiento del requerimiento fue del 16,67%.

En cuanto a las ayudas proporcionadas a los usuarios, ya sean relacionadas al sitio (y su contenido) o ayudas de las búsquedas, son muy bajas, teniendo en cuenta la complejidad y amplitud de los temas que el dominio de museos en la Web puede proporcionar a los visitantes en general (por ejemplo, áreas temáticas como colecciones, exhibiciones, visitas guiadas, información de interés educativo, historia, etc.). El porcentaje de ayuda del sitio global (1.2.1.1) fue del 18,75%. En cuanto a la ayuda de las búsquedas (1.2.1.2), si bien el muestrario nos da un porcentaje de 25%, hay que tener en cuenta que sólo el 41,67% poseen mecanismos de búsquedas, lo cual nos indica que el 60% del muestrario posee ayuda en dicho atributo.

Por otra parte, sólo el 29,17% de la población de los museos encuestados poseen la fecha de la última actualización (global) en la página principal (1.2.2.1), mientras que el 14,58% poseen fecha de actualización en los subsitios y/o páginas (1.2.2.2) lo cual son cifras muy bajas, teniendo en cuenta que los usuarios que visitan al sitio más de una vez, querrían saber si hubo alguna modificación en el nodo antes de navegar.

Unos de los atributos deseables a tener en cuenta en el diseño de los sitios Web es el de la retroalimentación por parte de los usuarios; mecanismo que podría permitir modificar el sitio según las preferencias de los visitantes. Un atributo para evaluar este concepto son los comentarios y sugerencias (1.2.4) que pueden ser completados por los visitantes a su voluntad. En cuanto a la implementación de este atributo, el 16,17% de los sitios encuestados cumplen con dicha propiedad.

En cuanto a la subcaracterística *Directorio de Direcciones* los porcentajes para cada atributo que la compone son:

- ✓ 1.2.3.1 - Directorio E-mail 60,42 %
- ✓ 1.2.3.2 - Directorio TE-Fax 81,25 %

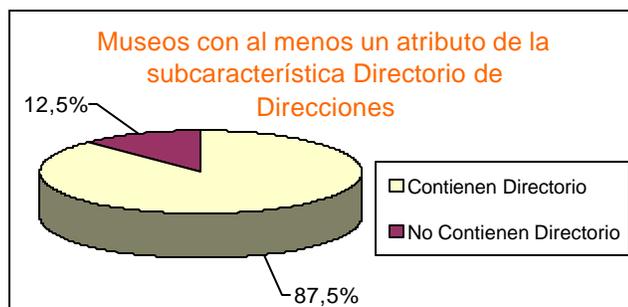


Figura A.3.3. Este gráfico demuestra que el 87,5% de los museos cumplen con la subcaracterística *Directorio de Direcciones*

Además, otro atributo evaluado es el 1.2.5 (*Facilidad FAQ*). Este atributo permite a los visitantes disponer de un listado de las preguntas más frecuentemente formuladas (y que a su vez pueden estar enlazadas con otras páginas de información), y así conocer y aprender rápidamente el contenido, sin tener que recurrir al envío de un e-mail u otros mecanismos de comunicación. El porcentaje que la encuesta proporciona para este atributo es del 16,67 %.

Por otra parte, una característica esencial de las aplicaciones Web dado su carácter de orientación a usuarios, es su interface. Cuando una persona está navegando, muchas veces desea volver a la página inicial del sitio o a la página inicial del subsitio que está visualizando. Para esto, la interface del sitio debe proporcionar controles principales y controles contextuales (por subsitio). Aquí no sólo se debe tener en cuenta la *Permanencia de los Controles Principales y Contextuales*, sino que también su *Estabilidad*, de modo que los visitantes tengan acceso permanente a los principales nodos del sitio, y realizando el menor esfuerzo cognitivo y de ajuste de interfaces [Thüring et al 95]. Los resultados para los atributos de la subcaracterística (1.3.1) fueron los siguientes:

- ✓ 1.3.1.1 - *Permanencia de Controles Directos*: 62,5 %
- ✓ 1.3.1.2 - *Permanencia de Controles Indirectos*: 39,58 %
- ✓ 1.3.1.3 - *Estabilidad*: 87,50 %

Finalmente, otro aspecto de importancia para el dominio de museos es el soporte a multi-lenguajes. El atributo *Soporte a Lenguajes Extranjeros*, modela este concepto en donde sólo el 58,33% de los sitios encuestados cumplen con el mismo.

A.3.2.2 Análisis para Funcionalidad. Una de las características fundamentales de funcionalidad de los sitios Web, principalmente para encontrar documentos en espacios grandes y complejos, es la búsqueda por palabras, operadores booleanos y/o filtros. Hemos considerado a la búsqueda desde dos perspectivas básicas: *Búsqueda Global en sitio Web* (2.1.1.2); y *Búsqueda Restringida* (2.1.1.1). En cuanto a la *Búsqueda Global* en el sitio podemos decir que de la totalidad de los museos evaluados, solamente el 37,5% cumplen con este atributo. Por otro lado, en cuanto a la *Búsqueda Restringida* solamente el 25% de los sitios encuestados tienen este atributo. Además, sobre la base del análisis realizado a partir de los datos de la tabla A.3.1, podemos decir que:

- ✓ El 41,67% del total de los museos de la muestra poseen al menos uno de los dos mecanismos de búsquedas (ver figura A.3.4)
- ✓ El 20,83% posee los dos mecanismos de búsqueda
- ✓ El 20,83% posee solamente uno de los dos mecanismos de búsqueda
- ✓ De los sitios que poseen búsqueda, solamente posee búsqueda global el 16,67%
- ✓ De los sitios que poseen búsqueda, solamente posee búsqueda restringida el 4,17%

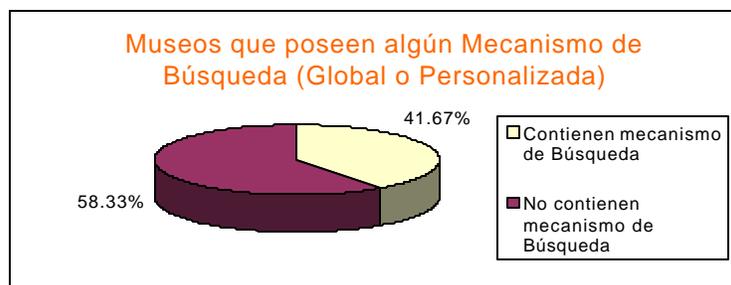


Figura A.3.4. El gráfico demuestra el porcentaje de los museos que cumplimentan con la subcaracterística Mecanismos Búsqueda en el Sitio (2.1.1)

Uno de los atributos que permiten conocer a los visitantes la ubicación a partir de donde proviene (considerando una estructura jerárquica de árbol), es el *Indicador de Camino* (2.2.1.1.1). Este atributo favorece a la *Orientación* del visitante. Otro atributo que favorece dicha subcaracterística es la *Etiqueta de la Posición Actual*. Ahora bien, si el 93,75% posee el atributo 2.2.1.1.2, solamente el 16,67% de la muestra contiene el *Indicador de Camino*. Esto nos puede indicar que no se ha tenido muy en cuenta en etapa de desarrollo, aspectos de orientación del usuario (lo cual afecta a la navegabilidad).

Finalmente, en cuanto a la subcaracterística *Permanencia y Estabilidad de los Controles Contextuales* 2.2.2.1, los resultados obtenidos fueron los siguientes:

- ✓ 2.2.2.1.1 *Permanencia de los Controles Contextuales*: 83,33%
- ✓ 2.2.2.1.2 *Estabilidad*: 81,25%

A.3.2.3 Análisis para Eficiencia. Una de las características importantes para evaluar a la Eficiencia, es la *Performance* (3.2) del sitio. Sin embargo, al igual que en el caso de estudio de museos [Olsina 99], nos encontramos en la disyuntiva de si evaluarla o no. Esto es debido a que el criterio usado depende del tamaño de la página (con todos sus componentes textuales, tabulares y gráficos). Pero debido a que las imágenes en los museos es un contenido esencial y debido a la importancia del atributo de ampliación de imagen (zooming), y de recorridos virtuales (1.1.2.2), esto produciría una baja en la velocidad potencial para cargar a las páginas. (No obstante, en el caso de estudio de sitios académicos [Olsina et al 99c] fue incluido y recolectado automáticamente. El criterio de evaluación fue uno absoluto de variable continua y múltiple).

Por otra parte, en cuanto a la *Accesibilidad* de información, la encuesta arrojó como preocupante el soporte a la versión de sólo texto para el sitio Web, con la cifra del 8.33%. Es oportuno decir que no todos los posibles usuarios poseen navegadores que soporten características gráficas sofisticadas; o bien, algunos usuarios pueden adolecer de discapacidades visuales (ver dentro de las recomendaciones de la W3C, las características de accesibilidad [W3C 99]).

Por último, muchos usuarios de la Web tienen la particularidad de desactivar la descarga de imágenes desde la opción del navegador para mejorar la velocidad. Una vez que la página ha sido descargada, los visitantes recorrerán las páginas y navegarán por medio del texto. Por lo tanto, es posible que las imágenes pudieran haber sido de relevancia para la información que estaban buscando. Por esta razón se debe considerar que las imágenes tengan un texto alternativo. El atributo que modela este concepto se llama *Imagen con Título* (3.1.1.2) en donde la encuesta arrojó como resultado de disponibilidad sólo el 35,42% de los sitios.

A.3.3. Observaciones Finales

Este survey, como una investigación adicional al caso de estudio sobre museos en la Web, discutido en el capítulo 9, analiza y muestra resultados de indicadores de disponibilidad de atributos pertenecientes a algunas características de calidad. Este estudio se realizó sobre artefactos Web en la fase operativa, lo que permite comprender el estado del arte en atributos de calidad (para un dominio) y analizar tendencias, con el fin de ser tenidos en cuenta como factores que favorecen a la calidad en el diseño y construcción de sitios Web actuales o en la evolución de los existentes.

Como conclusión del estudio podemos afirmar que el nivel de disponibilidad de atributos y subcaracterísticas de calidad básicas, es más bien bajo y la lista de atributos ausentes no está vacía. En general factores esenciales de los sitios Web como Esquema de Organización Global del sitio, Mecanismos de Búsqueda, Mecanismos de Orientación y Navegación, entre otros, no superan en conjunto una disponibilidad del

45% (lo que confirmaría para este caso la declaración que el diseño en la Web es pobre, -en consideración de un perfil de usuario y para el dominio específico).

Por ejemplo, en cuanto al Esquema de Organización Global (subcaracterística que favorece la rápida comprensión tanto de la estructura global como del contenido del sitio), la encuesta arrojó que un 41,67% al menos dispone uno de los tres atributos (considerando una relación de reemplazabilidad entre atributos, es decir, uno sería suficiente para tener una visión de la estructura y del contenido). En cuanto a Mecanismos de Búsqueda (una funcionalidad esencial para grandes espacios de información como museos), sólo el 41,67% del total de los museos muestreados poseen al menos uno de los dos mecanismos de búsquedas. Del estudio se puede concluir que, en general, los sitios de los museos, están mejor diseñados para exploración (browsing) que para búsqueda (en tanto que ambas características son igualmente importantes [Rosenfeld et al 98])

Algunas tendencias observadas son que los museos están incorporando la tecnología de visitas virtuales (principalmente como un complemento de los tours convencionales). Además, es satisfactorio la disponibilidad de atributos como ampliación de imágenes, algunos mecanismos de retroalimentación y aspectos de permanencia y estabilidad de los controles principales y contextuales en la interface, en tanto los usuarios navegan.

Tabla A.3.1 Muestra los valores para los atributos de cada Museo y los porcentajes de disponibilidad

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	%
1.1.1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00
1.1.1.2	0	0	0	0	0	1	1	0	0	1	0	1	1	0	0	0	1	0	0	0	0	0	1	0	29,17
1.1.1.3	0	0	0	1	0	1	1	0	0	1	0	1	1	0	0	0	1	0	1	0	0	1	1	0	41,67
1.1.2.1	0	1	0	1	1	0	1	0	1	1	1	0	0	0	1	0	1	1	1	0	1	0	0	1	54,17
1.1.2.2	0	0	0	0	1	1	1	0	1	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0	37,50
1.1.3	1	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	29,17
1.2.1.1	0	0	0	0	0	1	0	0	0,5	1	0	0	0	0	0	0	1	0	0,5	0	0	0	0	0,5	18,75
1.2.1.2	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	25,00
1.2.2.1	0	1	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	29,17
1.2.2.2	0	1	0	0	0	0	1	0,5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	14,58
1.2.3.1	0,5	1	0	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	60,42
1.2.3.2	0,5	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	81,25
1.2.4	0	0	0	0	1	1	0	0	0	0	1	0	1	0	1	0	1	1	0	1	1	0	1	0	41,67
1.2.5	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	16,67
1.3.1.1	0	0	1	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	1	1	62,50
1.3.1.2	1	1	0	1	0	1	0,5	0	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0	39,58
1.3.1.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	87,50
1.4.1	1	1	0	1	1	0	1	0	1	0,5	0,5	1	1	0	1	0	1	0	0	1	1	1	0	0	58,33
1.4.2	0,5	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0,5	16,67
1.4.3	0,5	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0,5	1	1	0	1	1	1	1	1	83,33
2.1.1.1	0	1	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	25,00
2.1.1.2	0	1	0	1	0	1	0	0	0	1	0	1	1	0	0	1	1	0	0	0	0	0	0	1	37,50
2.2.1.1.1	0	0	1	0,5	0	1	0	0	0,5	0,5	0	0	0	0	0	0	0	1	0	0	0	0	0	0,5	16,67

1.1.2.2	Visita Guiada Virtual	Permite recorrer un espacio virtual (salas, edificios) a través de un conjunto de imágenes unidas las que tienen la propiedad de ser vistas con los movimientos del mouse o del teclado. Es como mirar alrededor del lugar en que se encuentra el visitante en un espacio físico.
1.1.3	Mapa de Imagen	Muestra el plano o componentes del edificio o campus (salas, aulas, etc.), permitiendo navegación directa o indirectamente entre partes del mismo.
1.2.1.1	Ayuda Explicativa al Visitante	La ayuda explicativa debe estar basada en la riqueza, concisión, y oportunidad del texto (a veces podría ser acompañado por iconos, imágenes o animaciones).
1.2.1.2	Ayuda de la Búsqueda	Es igual al atributo 1.2.1.1 pero con la diferencia que debe estar relacionada al mecanismo de búsqueda.
1.2.2.1	Indicador de Última Actualización Global	Permite conocer al visitante si se ha realizado alguna modificación o agregado en el sitio Web. Se debe encontrar en la página principal.
1.2.2.2	Indicador de Última Actualización Restringido	Permite conocer al visitante si se ha realizado alguna modificación o agregado en alguna página del subsitio (la que está visitando).
1.2.3.1	Directorio E-mail	Es el lugar en el sitio donde se agrupan las direcciones electrónicas para enviar E-mail (como mecanismo de retroalimentación)
1.2.3.2	Directorio de TE-Fax	Es el lugar en el sitio donde se agrupan las direcciones de números para contactarse vía telefónica o por medio de Fax
1.2.4	Comentarios/ Sugerencias	Este atributo permite hacer conocer las inquietudes u otros intereses que el usuario formule mediante el llenado de un formulario con mayor o menor nivel de estructuración.
1.2.5	Facilidad FAQ	La disponibilidad FAQ es un conjunto de preguntas que se realizan con mayor frecuencia, y que están ya publicadas en el sitio con sus respuestas. Esto le permite aprender y/o ayudar a los visitantes (evitando la demora cognitiva).
1.3.1.1	Permanencia de Controles Principales Directos	Este atributo representa la permanencia directa de los controles del menú principal del sitio que permiten navegación.
1.3.1.2	Permanencia de los Controles Indirectos	Es un control o referencia indirecta a la página principal (en donde se encuentran los controles a los subsitios)
1.3.1.3	Estabilidad	Se refiere a la ubicación de los controles principales directos o indirectos en los nodos del sitio. Un control es estable si se encuentra siempre en la misma ubicación al navegar por los nodos.
1.4.1	Soporte a Lenguaje Extranjero	Este atributo modela la disponibilidad parcial o total de lenguajes extranjeros soportados por el sitio Web. No se computa el lenguaje nativo como lenguaje extranjero.
1.4.2	Indicador de Tamaño	Este atributo representa la disponibilidad del indicador del tamaño de la imagen (en bytes, Kb) que es mostrada a los visitantes antes de accederla, visible por medio del objeto enlazado o texto alternativo.
1.4.3	Zooming de la Imagen	Atributo que permite ampliar la imagen para una mejor apreciación
2.1.1.1	Búsqueda Restringida	Algunas veces tiene sentido brindar búsqueda restringida a un subsitio o parte de un sitio, debido a que el mismo es altamente cohesivo o distintivo del resto de la información del sitio Web (por ejemplo, colecciones, libros, autores, etc.).
2.1.1.2	Búsqueda Global	Este atributo permite realizar búsquedas sobre cualquier palabra/s en el sitio (proveyendo algunas veces operadores y/o filtros)
2.2.1.1.1	Indicador de Camino	Los usuarios al navegar por el sitio, deben tener pistas visuales (con elementos de diseño consistentes), que les indique con precisión en dónde se encuentran posicionados dentro de la estructura del espacio de información del sitio. Este atributo trata con la orientación del usuario en tanto navegan el hiperespacio.

2.2.1.1.2	Etiqueta de la Posición Actual	Este atributo permite hacer conocer a los usuarios en qué nodo están posicionados, por medio de una etiqueta o título.
2.2.2.1.1	Permanencia de los Controles Contextuales	Es la permanencia de los controles de navegación de un subsitio
2.2.2.1.2	Estabilidad	Se refiere a la ubicación de los controles contextuales en los nodos del subsitio. Un control es estable si se encuentra siempre en la misma ubicación durante la navegación de los nodos del subsitios.
2.2.3.1	Enlace con Título	Este atributo trata de predecir los temas o contenidos que están asociados al enlace.
3.1.1.1	Soporte a Versión solo Texto	Este atributo representa la accesibilidad a la información que está en las páginas, principalmente para las personas con invalidez o cuando la velocidad es un problema. Es de relevancia que el sitio entero sea editado en una versión sólo texto; sin embargo, a veces, una disponibilidad parcial también puede ser deseable
3.1.1.2	Imagen con Título	Se debe proveer texto alternativo para cada imagen u objeto gráfico para que comuniquen información visual. Este atributo mide el nivel de disponibilidad del texto alternativo a la imagen, útil principalmente cuando se desactiva la propiedad de ver imágenes del navegador.

APENDICE B

Una Estructura Conceptual para un Modelo de Proceso Genérico.

El objetivo de este apéndice¹ es definir una estructura o modelo conceptual útil para comprender y soportar a la modelización de procesos. Para una mejor comprensión de esta parte, la estructuraremos del siguiente modo:

- ✓ Primero realizamos una definición, bajo un enfoque lexicográficamente descriptivo, de los principales conceptos que intervienen en procesos de software y modelado de los mismos, remitiendo al lector a distintas fuentes de la literatura investigada. El abordaje no pretende ser amplio (en cuanto a la cantidad de conceptos tratados) sino el necesario para facilitar la comprensión del marco conceptual del punto B.2.
- ✓ Luego, presentamos un modelo conceptual para el modelado de procesos de software. Mediante un diagrama de clases, representamos a los principales entes y relaciones intervinientes.

B.1 Conceptos sobre Procesos de Software y Modelado de Procesos y Productos

El conocimiento público de carácter científico, se ve favorecido si se identifica y comunica, de un modo no ambiguo y representativo, a un conjunto de conceptos primitivos de un dominio.

En ingeniería de procesos de software se manejan conceptos como el de proceso, tarea, actividad, agente, artefacto, recurso, rol, constructores de proceso, descripción de proceso, modelo de proceso, perspectiva o vista de proceso, arquitectura de proceso, ambientes de ingeniería de software orientados a procesos, entre otros conceptos.

Esfuerzos importantes se han venido realizando con el fin de establecer una base conceptual sólida [Curtis et al 92, Dowson et al 91, Feiler et al 93, Lonchamps 93, Madhavji 91, más recientemente Verlag et al 98, entre otros]. Trabajos seminales en el área son los de Humphrey, Feiler y Dowson et al. Pensamos que nuestro modelo conceptual (fig. B.6), por medio de un diagrama de clases y relaciones, contribuye a

¹ Nota: El texto de este apéndice fue extraído y adaptado ligeramente, principalmente en cuanto a tamaño, del Capítulo 3 de la Tesis de Magister [Olsina 98a].

representar y sintetizar de un modo no ambiguo el dominio del problema. A seguir introduciremos los conceptos² usando un enfoque lexicográficamente descriptivo.

Una **fase** es una agrupación de *procesos de software* fuertemente relacionados o cohesivos realizados en cierto orden. Por ejemplo, en todo *modelo de proceso de producto* que represente a todo el ciclo de vida de *artefactos* se puede apreciar generalmente tres fases: la de exploración, la de desarrollo y la operativa (en la sección 4.2.1.1. se describe sucintamente a cada fase).

Humphrey ha dicho que un **proceso de software** (o proceso, para abreviar) es un conjunto parcialmente ordenado de *subprocesos* que tienen el fin de alcanzar alguna meta establecida. De un modo más amplio, podemos definir a un proceso de software como a un conjunto parcialmente ordenado de subprocesos a los que se le asocian una colección de *recursos, agentes, condiciones, artefactos y constructores de proceso*, con el fin de producir los *distribuibles* conforme a las metas establecidas.

Por **distribuible** entendemos a un *artefacto* (producto) requerido por algún proceso o agente interno o externo.

Ejemplos de nombres de procesos a diferente nivel de granularidad son: administración de un proyecto Web, especificación de requerimientos de calidad, análisis conceptual, diseño navegacional, diseño de interfaces, determinación y resolución de riesgos, evaluación elemental de calidad, aseguramiento de calidad, prototipación, validación, etc.

Un proceso se puede descomponer en **subprocesos** (paso o elemento de proceso), por lo tanto, le cabe una definición recursiva. El nodo raíz del árbol es el más general y abstracto, el nodo hoja es el más específico y lo denominamos *actividad*. Por lo que una **actividad** es un subproceso que no requiere más descomposición. Se dice que una actividad es un elemento de proceso descrito a un bajo nivel de granularidad (granularidad fina)

Es oportuno destacar que en la literatura se ha establecido una distinción, a veces sutil, entre el concepto de actividad y el de **tarea**. En el área de administración de proyectos una tarea es una unidad de trabajo a realizar por un agente. Una tarea es un subproceso a la que se le asocian componentes de gestión, es decir, se le pueden asignar agentes, recursos, se la puede planificar, programar, ejecutar y controlar.

En el área de ingeniería de procesos de software también se ha utilizado a la tarea como a la entidad atómica de abstracción, sobre todo en enfoques prescriptivos y orientados al

² Utilizaremos la/s palabra/s resaltada/s en **negrita** para describir un nuevo concepto, con *itálica* los conceptos a describir en párrafos subsiguientes y, en esta sección, usaremos paréntesis para referirnos a (sinónimos). Ver además el glosario de términos en la sección 14.

análisis, como apunta Lonchamps [Lonchamps 93]. Sin embargo, existen enfoques en que a las tareas se las dividen en actividades, siendo las actividades entes que no necesitan ser gerenciados (sobre todo en entornos o herramientas de automatización). Por lo tanto, en el caso en que se establezca la diferencia entre tarea y actividad, ésta corresponderá a que un proceso pueda ser o no gerenciado por un agente. En la figura B.1 observamos la relación entre entes que representan a las definiciones previas de proceso, actividad y tarea. (En general usamos la notación de Rumbaugh [Rumbaugh et al 91], y en la figura B.6. una adaptación de UML [UML 97])

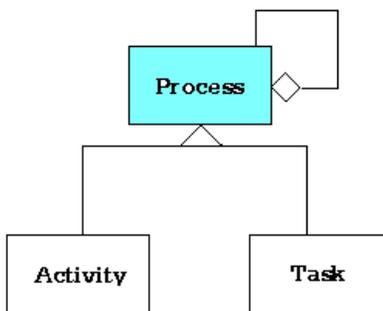


Fig. B.1 Diagrama de clases que relaciona a los entes Proceso, Tarea y Actividad.

En cuanto al concepto de **agente** lo definimos como al ente ejecutor de un proceso. El agente puede ser tanto un ente humano como una herramienta o dispositivo computarizado. Podemos realizar una taxonomía de tareas en función de los agentes teniendo en cuenta que una tarea puede llevarse a cabo por uno o más agentes. Clasificamos a las tareas en tres tipos:

- ✓ manuales
- ✓ automáticas
- ✓ interactivas

Las tareas manuales solo involucran a agentes humanos; las tareas automáticas solo dependen de agentes computarizados, en tanto que las últimas necesitan de ambos tipos de agentes con el fin de alcanzar la submeta establecida. Para llevar a cabo un proceso, el agente humano o computarizado debe contar con un conjunto de habilidades (experticia). Para cumplimentar un *rol* se deben requerir habilidades específicas por parte de los agentes. Por ejemplo, un agente humano debe tener experticia sobre criterios estéticos y cognitivos de interfaces de usuario y aspectos de diseño del look & feel para cumplimentar el rol de “Diseñador de Interfaces”.

Un **artefacto** es el producto de realizar una tarea. Un artefacto es el producto creado, evolucionado, mantenido o destruido durante el proceso de desarrollo ya como un resultado requerido por un agente o para facilitar la prosecución de otro proceso. Con lo dicho podemos razonar que un artefacto puede servir de entrada a un proceso y, mediante la transformación correspondiente, ser la salida del mismo. Además, un

artefacto puede ser un objeto compuesto, es decir, se da una relación de agregación entre componentes.

En nuestra visión los procesos de desarrollo de software fundamentalmente existen para crear, modificar y reusar artefactos. Entre los objetos considerados artefactos se encuentran los documentos -por ejemplo: documento del plan del proyecto Web, especificaciones de requerimientos de calidad, prototipos y versiones finales de productos -código fuente y/o ejecutable-. Los artefactos pueden ser administrados bajo una estrategia de configuración de cambios.

En cuanto a un **recurso** es un ente del mundo real necesario para que las tareas de un *proyecto de software* se puedan efectuar. Una tarea usa recursos. Ejemplos de recursos son: recursos humanos, monetarios, materiales, tecnológicos. La figura B.2 muestra una jerarquía general de recursos.

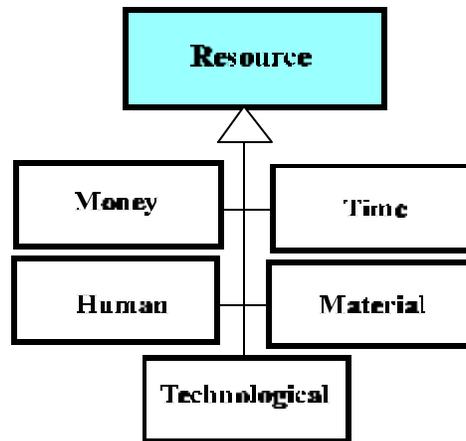


Fig. B.2 Diagrama general de jerarquía de recursos.

Un **rol** es un conjunto de permisos y obligaciones que se debe asociar a un agente durante la realización de un tipo de tarea. El agente debe tener un conjunto de permisos para realizar las actividades de la tarea conforme a la submeta establecida y obligado a satisfacer un conjunto de *condiciones*.

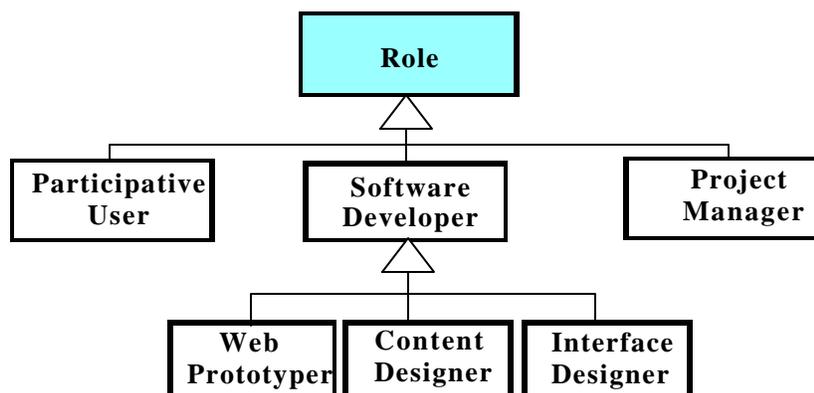


Fig. B.3 Ejemplo reducido de jerarquía de roles de un proyecto Web.

Se puede establecer una jerarquía de roles: roles más generales y roles más específicos. En la figura B.3 observamos que el rol de diseñador de interfaces es más específico que el rol de desarrollador de software.

En la página 493 de [Goldberg et al 95], los autores presentan un diagrama -de Venn- en donde se muestra los roles ocupados por agentes, agrupados estos en equipos de trabajo. Cuando un proceso se está llevando a cabo puede existir una asignación dinámica de roles a agentes y viceversa. Los agentes pueden ocupar –jugar- diferentes roles a diferentes momentos; un rol dado puede ser ocupado por distintos agentes a diferentes momentos [Dowson et al 91].

Una **condición** de un proceso es la declaración del estado de situación que debe acontecer para el inicio, ejecución y finalización de un proceso. Una actividad puede comenzar cuando se cumple un conjunto de precondiciones y puede finalizar cuando se alcanzan las postcondiciones establecidas.

Un **constructor de proceso** es un enfoque específico (método) que se puede usar para realizar una tarea. Ampliando el concepto podemos decir que para llevar a cabo una tarea especificada en la *descripción de proceso* se pueden usar diferentes métodos que responden a distintos enfoques (ver fig. B.4).

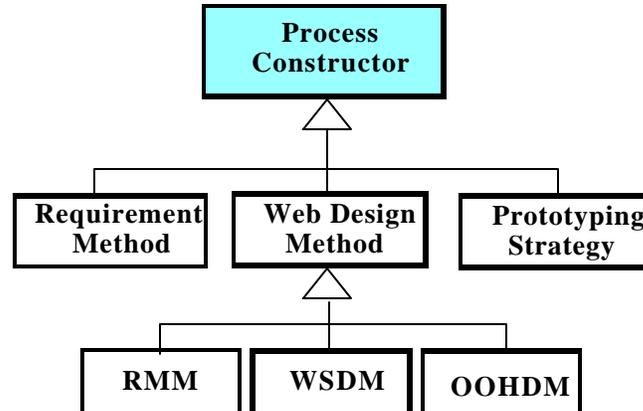


Fig. B.4 Ejemplo reducido de jerarquía de constructores de proceso de un proyecto Web.

Por ejemplo, para una tarea de “modelado del dominio de la aplicación”, podemos usar el método especificado en RMM [Isakowitz et al 95], denominado “diseño de E-R”, o podemos usar para esa tarea el método especificado en la metodología OOHDM [Rossi 96, Schwabe et al 95b], denominado “modelado conceptual”. Son enfoques diferentes para una descripción de proceso dada.

En el trabajo seminal sobre el área de ingeniería de procesos, Osterweil dijo [Osterweil

87] que los procesos de software son también software. Si bien el enfoque era formal, centrado en descripciones procedurales y declarativas, hay diferentes paradigmas para describir a los procesos. Primero veamos el concepto.

Una **descripción de proceso** es una manera de representar y especificar la secuencia parcial de actividades de un proceso. Una descripción completa de proceso debe considerar las actividades y las operaciones asociadas, las precondiciones y postcondiciones para cada actividad, y otros entes (objetos) intervinientes en el proceso de software como artefactos, agentes y roles. Pueden coexistir descripciones alternativas para un proceso y se puede considerar a una descripción de proceso como a un artefacto especial.

Una descripción de proceso está para que pueda ser interpretada y ejecutada por un agente humano o un agente computarizado. Un área de investigación atractiva es la de soporte a procesos de software. Se han realizado algunas contribuciones y avances, en el área de *ambientes de ingeniería de software centrado en procesos*. En estos sistemas muchas de las tareas son interpretadas y ejecutadas automáticamente por un *motor de procesos*. Puede coexistir para una cierta actividad descripciones alternativas bajo un mismo enfoque, como indicamos anteriormente, o descripciones multi-paradigmáticas.

Los aspectos de mayor o menor formalidad en la descripción de procesos es una cuestión de relevancia. Una representación precisa, formal, es ejecutable por una máquina, en tanto que una representación informal, esto es en lenguaje natural, o una representación semiformal estructurada en un script, puede ser ejecutada por un agente humano y no por un agente computarizado -aunque en algunos casos una representación semiformal puede ser interpretada.

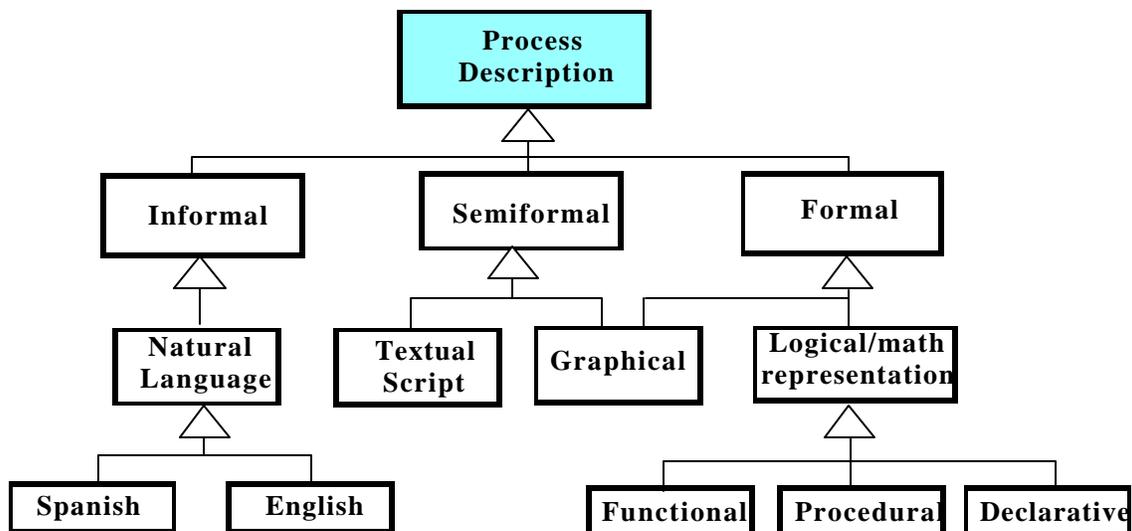


Fig. B.5 Diagrama reducido de tipos de descripciones de proceso en función de la formalidad de su representación.

En la figura B.5 el lector puede observar una taxonomía de descripciones de procesos en función de la formalidad de su representación. Otra taxonomía se podría realizar o complementar teniendo en cuenta el nivel de granularidad de las actividades descriptas. Los autores en [Curtis et al 92], han expresado que la experiencia sugiere que los procesos automatizados requieren una representación de granularidad más fina, en tanto que una representación semiformal como los script, de granularidad más alta, son adecuadas para ser interpretadas y ejecutadas por agentes humanos.

Relacionado a la representación de las descripciones de procesos y a las distintas *perspectivas* de un *modelo de proceso* está el concepto de *lenguajes de modelado de procesos*, que describiremos en párrafos subsiguientes.

Los procesos de software se pueden abstraer en un **modelo de proceso** (modelo de ciclo de vida). Definimos a un **modelo** como a una representación abstracta de entes o fenómenos de la realidad en la que se consideran los aspectos relevantes de los mismos y se desechan los menos relevantes sin que por ello deje de representar significativamente a esa realidad.

Definimos a un modelo de proceso de software como a una estrategia apropiada para abstraer, organizar, ejecutar y/o controlar a las distintas fases, tareas, recursos y artefactos de un proyecto con el objeto de alcanzar las metas establecidas. Un modelo de proceso es una descripción más o menos formal del proceso de desarrollo de software. Por lo que un modelo de proceso expresa: 1) un cierto nivel de abstracción; 2) una *perspectiva* o vista particular del proceso de desarrollo.

Modelos de ciclo de vida como los modelos de cascada, espiral, recursivo/paralelo [Berard 93], etc. son modelos muy abstractos y generalmente semiformales. No se especifica con suficiente detalle y rigor formal, las entradas y salidas de las actividades, condiciones de comienzo y finalización, flujos de artefactos, sincronización de actividades, entre otros asuntos.

Un modelo de proceso integra diferentes tipos de información de los procesos de software. Una **perspectiva** (vista o submodelo) es un enfoque particular para especificar y comunicar información del modelo. Un modelo de proceso debe ayudar a responder preguntas tales como: qué hacer, cómo y cuándo hacerlo, dónde y quiénes lo harán, qué dependencias existirán entre tareas y cómo se sincronizarán, etc.

Curtis et. al. agrupan la información en cuatro perspectivas con el fin de abstraer la complejidad subyacente en los procesos de software, y son, a saber: *perspectiva funcional*, *perspectiva informacional*, *perspectiva de comportamiento* y *perspectiva organizacional*. Pensamos que una *perspectiva metodológica* [Olsina 97c] es fundamental para representar a los diferentes métodos (constructores de proceso) y, además, puede ser integrado en *ambientes de ingeniería de software centrado en*

procesos.

La **perspectiva funcional** representa qué tareas se deben realizar en cada fase, qué estructura jerárquica de actividades existe para cada tarea, y qué dependencias funcionales -flujos de información y control, documentos, artefactos-, son relevantes para las entradas y las salidas.

La **perspectiva de información** se centra en los artefactos producidos o requeridos por los procesos de software; en la estructura de los artefactos y sus interrelaciones; en las estrategias de administración de cambios de artefactos y seguimiento (traceability) de los mismos.

La **perspectiva de constructores de proceso** representa los métodos a aplicar para efectuar las distintas actividades especificada en la descripción de procesos. Para un método en particular puede haber una o más herramientas que den soporte a la tarea.

La **perspectiva de comportamiento** representa aspectos dinámicos del modelo de proceso: el secuenciamiento y la sincronización de las tareas; información de cómo se realizan las actividades, esto es, iteraciones, ciclos de retroalimentación, paralelismo, condiciones de inicio y terminación, entre otros asuntos. Asimismo puede especificar el ciclo de vida de un ente como un artefacto, proceso, agente con formalismos como: diagrama de transición y estados, statecharts, redes de Petri, etc.

En la **perspectiva organizacional** se consideran aspectos de qué agentes participan/planifican/ejecutan y controlan a qué tareas; qué roles se asignan a los agentes participantes; qué estrategias de comunicación y dinámica de grupos se aplican, etc.

Por lo tanto, un modelo de proceso integral y completo, necesita de una **arquitectura de proceso** (meta-modelo de proceso) para conectar a las diferentes perspectivas. Una arquitectura de proceso es una estructura conceptual útil para describir las perspectivas relevantes y sus relaciones.

Un **lenguaje de modelado de procesos** es un paradigma capaz de representar a un esquema de meta-modelo de proceso. Como todo lenguaje debe proveer una sintaxis precisa y una semántica no ambigua y amplia. Desafortunadamente, es difícil contar con un lenguaje que represente a las distintas perspectivas de un modelo de proceso y provea constructores de soporte a las distintas actividades de modo que pueda ser interpretado y ejecutado con eficacia en diferentes *proyectos de software*. En la práctica, la mayoría de las descripciones de proceso, cuando existen, se basan en mecanismos informales o semiformales difíciles de ser ejecutados en *ambientes centrados en procesos*.

No obstante algunos aspectos de un modelo de proceso pueden ser automatizados. Curtis et al. presentan un conjunto de tipos de lenguajes que responden a distintos paradigmas, que pueden ser útiles para representar a las diferentes perspectivas de un modelo de proceso (en esta área se ha avanzado últimamente).

Un **ambiente de ingeniería de software centrado en procesos** (ambiente (o sistema) de soporte a procesos de software) es un entorno de trabajo que ofrece asistencia a los usuarios en la ejecución de un *proyecto de software*. Un componente esencial del ambiente es el **motor de procesos** el cual es el intérprete automático de las descripciones de proceso provistas de un modo estático o dinámico. Con dinámico queremos significar que durante la ejecución del proceso la descripción puede cambiar (evolucionar).

La arquitectura de un ambiente de software centrado en procesos debe ofrecer servicios de administración de objetos e interrelaciones, servicios de herramientas, servicios de trabajo colaborativo, y otros como administración de cambios, versionamiento, seguimiento de artefactos, seguridad.

La realización de un **proyecto de software** comprende a un conjunto de tareas, tanto técnicas como de gerenciamiento, a un conjunto de recursos, a un conjunto de estrategias, métodos y heurísticas, con el propósito de lograr los objetivos y las metas acordadas. Un proyecto de software puede ser autocontenido o parte de otro proyecto mayor y está inserto en el marco de alguna estrategia organizacional.

B.2 Definición de un Modelo Conceptual

En esta sección presentamos un modelo conceptual para el modelado de procesos de software. Pensamos que al definir un marco base que contenga a un conjunto canónico de conceptos que intervienen en modelado de procesos, potencia a las investigaciones en el área, al permitir:

- ✓ establecer una base conceptual reusable y que favorece a la comunicación.
- ✓ la posibilidad de combinar diferentes clases y relaciones para dar soporte a las distintas perspectivas.
- ✓ la posibilidad de crear capacidades en ambientes de ingeniería de software centrado en procesos para que soporten multi-paradigmas en la ejecución de actividades.

El objetivo de esta parte del trabajo consiste en presentar un modelo conceptual especificado mediante un diagrama de paquetes, clases y relaciones fundamentales (figura B.6), soportado por los principios de Orientación a Objetos.

Las clases (o conceptos clave para nuestro dominio del problema que es el modelado de

proceso) representan comportamiento y estado. En definitiva disponen de métodos para la ejecución de operaciones y encapsulan atributos del ente. Existen métodos denominados constructores y destructores entre otros tipos de métodos. Por ejemplo una clase artefacto puede tener como atributos su identidad, su fecha de creación, y terminación (para una versión dada), el identificador del documento asociado, etc. y puede responder a solicitudes de un agente por medio de operaciones de creación, destrucción, modificación del contenido o de sus atributos.

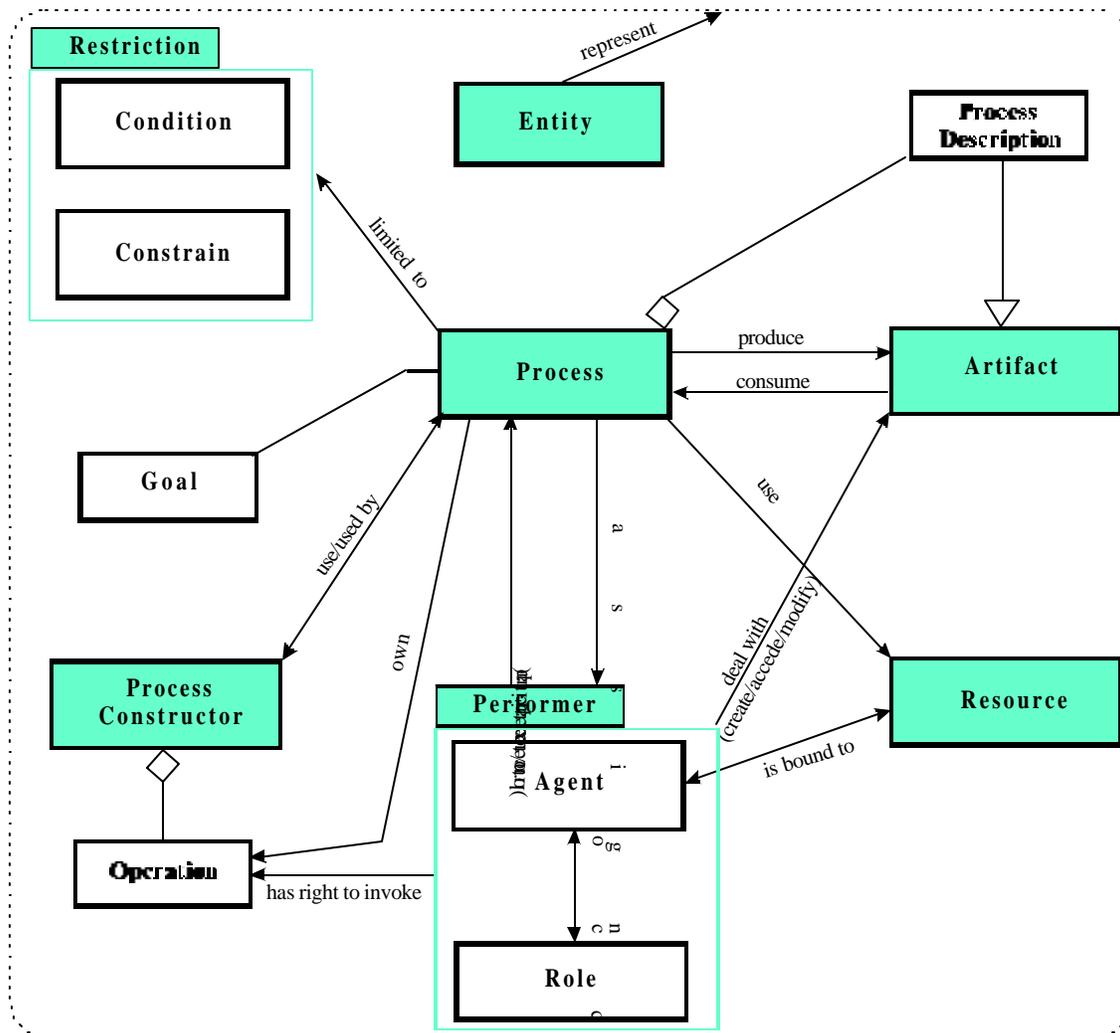


Fig. B.6 Diagrama de paquetes, clases y relaciones fundamentales para un Modelo de Proceso Genérico.

Por otra parte, las clases soportan diferentes mecanismos de relaciones [Booch 94, Rumbaugh et al 91, UML 97]. Las relaciones pueden darse entre entes de una misma clase, o entre clases distintas. Podemos citar mecanismos de herencia, de agregación, de asociación. Por ejemplo, en la fig. B.3 mostramos un diagrama simplificado de la jerarquía de herencia para la clase "Rol" (p.ej. un "Interface Designer" *is-a* "Designer"); un artefacto compuesto está relacionado por un mecanismo de agregación; las clases fundamentales se relacionan entre sí por medio de mecanismos de asociación (por

ejemplo, ver la relación produce y consume entre las clases "Artifact" y "Process").

De este modo, en la figura B.6 una clase de tipo “Agente”³ asociada o jugando un tipo de “Rol” realiza un tipo de “Tarea”. (Así, “un agente, Gustavo” controla “la tarea, diseño de contextos de navegación” cuando ocupa el “rol, coordinador del proyecto”; “un agente, Luis” ejecuta “la tarea, diseño de contextos de navegación” cuando ocupa el “rol, desarrollador del proyecto”; “un agente, Marina” participa “en la tarea, diseño de contextos de navegación” cuando ocupa el “rol, usuario participante”. El lector puede observar que los agentes tienen distintos derechos, y que se puede establecer un mecanismo de delegación).

En la figura B.6, consideramos explícitamente las relaciones fundamentales entre las clases. Sin embargo, con n clases bases podríamos tener un número de relaciones potenciales de n factorial ($n!$). Por ello una división de preocupaciones por perspectivas puede disminuir la complejidad en el modelado de procesos.

B.2.1 Descripción de las Clases y sus Relaciones Fundamentales

Si bien en la sección B.1 introducimos un conjunto de conceptos, a seguir profundizaremos sobre los entes y las relaciones mostradas en la fig. B.6 [Olsina 98b].

Una “tarea” representa una unidad de trabajo que se le asigna a un “agente” (o “Performer”) para su realización. Una “tarea” contiene “descripciones de proceso”, las que pueden comprender una colección de alternativas -instanciables-, para representar a la misma unidad de trabajo. Asimismo, pueden estar especificadas en diferentes formalismos y/o notaciones (figura B.5), las que serán comprensibles para los distintos “agentes” involucrados.

Si observamos (leemos) algunas clases y relaciones del diagrama podemos decir que un “agente” realiza una “tarea” contenida en la “descripción de proceso”, usando algún “constructor de proceso” en particular. Para la efectivización de la tarea el “Performer” tiene el derecho a invocar las “operaciones” que son parte del “constructor”. Esta clase al ser usada alimenta al “proceso” cuya finalidad es producir un “artefacto” que esté dentro de las “metas” establecidas. El “proceso” se lleva a cabo bajo ciertas “condiciones”.

Además de las “operaciones” que posee una “tarea” asociada al constructor de proceso en particular, la tarea también posee métodos para crear una actividad (constructor), y para la terminación de la misma (destructor). Por otra parte, hay métodos para la delegación de actividades.

³ A seguir subrayaremos a las relaciones y pondremos entre comillas a las clases y objetos (en algunos casos acompañada de otra información).

Dado el carácter jerárquico de las tareas, en la que una actividad puede depender de otras actividades, delegar una actividad al agente correcto implica que éste tendrá el derecho de invocación de ciertas operaciones. La idea de delegación surge de la relación jerárquica de actividades y en los permisos de los agentes a controlar, ejecutar y/o participar en el desarrollo de las tareas. De este modo podemos distinguir entre el concepto de que una tarea posee operaciones y el concepto de que el agente tiene el derecho a invocarlas. (Observar el ejemplo de Gustavo, Luis y Marina, dado más arriba, para ejercitar estas ideas).

Un artefacto es un objeto temporario o persistente que representa el producto de realizar una tarea. Esto es, un “proceso” produce “artefactos” por medio de la colaboración de uno o más “agentes”, y a su vez un “proceso” reusa “artefactos”.

Un artefacto puede ser un objeto simple o compuesto y puede estar sujeto a un sistema de configuración de cambios. El mismo consiste en una estrategia -asociado a algún método- para identificar, mantener, y administrar los cambios de los artefactos bajo cierta configuración. Se debe considerar aspectos de procedimientos de aceptación y congelamiento de los cambios, ítems configurables, versionamiento, entre otros asuntos.

Desde el punto de vista de un sistema de configuración y de un modelo de seguimiento, un artefacto es un objeto con propiedades y comportamiento visibles. Un artefacto es un objeto persistente que se caracteriza por su identidad, tipo de artefacto (por ej. un documento es un tipo de artefacto- ver la definición en el glosario), fecha de creación, de última actualización, versión, estado (por ej. en estado de procesamiento o en estado de aprobación).

Además, los agentes humanos y computarizados disponen de un conjunto de servicios para realizar sus acciones. Una taxonomía de métodos podría ser la siguiente:

- ✓ métodos de contenido, para operar sobre el contenido de un objeto (en general se dispone de herramientas específicas asociadas a constructores de proceso)
- ✓ métodos estructurales, para crear nuevos objetos, para agregar a la jerarquía existente, para eliminar artefactos, para consultar las estructuras de relaciones
- ✓ métodos para operar sobre el estado
- ✓ métodos para dar soporte a un modelo de seguimiento

Considerando una perspectiva de comportamiento, y centrándonos solamente en un modelo de ciclo de vida de un artefacto, podemos utilizar formalismos de modelado tradicional como, máquinas de estado finita, cartas de estado (statecharts) y sus derivados, o redes de Petri. Un posible esquema de transiciones y estados para el ciclo de vida de un ente [Humphrey et al 89], en nuestro caso un artefacto, es el que mostramos en la figura B.7.

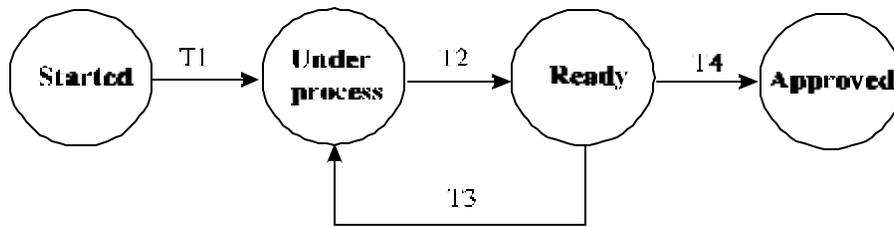


Fig. B.7 Un diagrama de transiciones y estados que representa el ciclo de vida de un artefacto.

Una máquina de estados asociada a un artefacto puede facilitar la definición de la semántica de los estados y transiciones de un producto. Por ejemplo, para un artefacto compuesto un objeto de la jerarquía puede pasar a estado “aprobado” si todos sus componentes están aprobados y pasaron a conformar la línea base del sistema de configuración y administración de cambios.

En cuanto a las transiciones son disparadas por eventos. Para que una transición provoque un cambio de estado, se deben cumplir las precondiciones establecidas; el efecto del cambio se puede representar en las postcondiciones. Una forma de especificar lo anterior es mediante la siguiente notación:

Transición T_n : $\langle \text{Evento}, \text{Estado previo}, \text{Estado siguiente}, \text{Precondición}, \text{Postcondición} \rangle$

Un ejemplo de condición para pasar del estado “listo” al estado “aprobado” de un objeto compuesto es el indicado más arriba; otra condición es que el agente que dispara la transición, tenga el permiso apropiado (en el cumplimiento de su rol) para provocar los cambios de estado de “listo” a “aprobado”.

Finalmente, en cuanto a la clase “agente” abstrae al ente que controla y/o ejecuta y/o participa en la realización de las “actividades” y, además, se relaciona con la clase “artefacto”.

La clase “agente” tiene una relación muchos a muchos con la clase “recurso”. En la figura B.2 mostramos un diagrama general de tipos de recursos. Un recurso se puede asociar a varios agentes y, un agente puede estar asociado o usar varios recursos. Por ejemplo una persona, que es un recurso, puede asociarse a varios agentes, y un agente puede usar varios recursos como, una persona, hardware y sistema operativo de una computadora, espacio físico, etc. Un agente puede usar una composición de recursos del mismo tipo (equipo de personas).

B.2.2 Comentarios finales

Esta división de responsabilidades entre la clase “agente” y la clase “recurso” es

relevante para los objetivos de la modelización. Se puede pensar en la ejecución de una actividad sin que estén disponibles los recursos. Estos representan a objetos del mundo real, en tanto que los agentes representan a objetos del mundo lógico -que en definitiva implica mayor nivel de abstracción.

La división de preocupaciones entre las clases “agente”, “recurso” y “rol” es también de importancia. Principalmente a la separación entre un agente abstracto asociado a un rol para realizar una actividad, de un agente específico, con habilidades particulares. Por ejemplo, la sentencia “El agente X -uno no instanciado-, está asociado al rol de desarrollador del proyecto, cuando realiza el subproceso de diseño navegacional”, es de un nivel más abstracto e independiente de la asignación de recursos y ocupación de roles que la siguiente sentencia: “El agente Gustavo, ocupa el rol de desarrollador del proyecto, dado que posee experticia reconocida en la intervención de los proyectos GIDIS y Pampa I, para la realización de la tarea de diseño navegacional, en el proyecto Pampa II”. Aquí el ocupante del rol, que es un recurso específico provisto por la organización, posee habilidades específicas, para que integre el nuevo proyecto de Hipermedia.

La definición de un modelo conceptual y la distinción entre distintos niveles de abstracción y de preocupaciones son, como indicamos anteriormente, de importancia para la modelización de procesos por varios motivos:

Primero, porque representa a las abstracciones claves del dominio del problema, agrupando comportamientos comunes y haciendo explícitas un conjunto de relaciones. Esto favorece la distribución de responsabilidades e identificación de colaboraciones.

Segundo, favorece el esquema de perspectivas, es decir, la división de preocupaciones en submodelos o vistas del modelo de proceso. Como describimos en la sección B.1, para disminuir la complejidad, es conveniente separar diferentes tipos de información de los procesos para especificar, comunicar y controlar porciones del modelo. Definimos las perspectivas funcional, informacional, de comportamiento, organizacional y metodológica. Por ejemplo, una perspectiva funcional se focaliza principalmente en la clase proceso y sus relaciones; una vista informacional se focaliza en la clase artefacto, sus componentes y relaciones; una vista organizacional se concentra en las clases recurso, agente y rol y la asignación a las distintas tareas.

Por último, y desde otro punto de vista, es importante la separación de preocupaciones de estos dos enfoques para la modelización: aquél enfoque que no depende de los recursos específicos ni de las diferentes instancias de constructores a usar para una descripción de procesos, de aquél otro enfoque que se concentra esencialmente en los recursos específicos, habilidades particulares de los agentes y ocupación de roles.

APENDICE C

Empleo de un Modelo Orientado a Metas (Enfoque Goal-Question-Metric en el Proyecto *Facultad de Ingeniería*)

C.1 Características deseables de Artefactos, Procesos y Recursos en Proyectos de Hipermedia ¹

Una de las metas principales en el desarrollo de aplicaciones de Hipermedia (como en cualquier otro tipo de desarrollo de software) es producir artefactos de calidad, los que deben estar regidos por un conjunto de atributos deseados y observables, utilizando para tal fin los procesos más óptimos y los recursos más apropiados para esas características. Debemos asegurar los mecanismos por medio de los cuales podamos construir artefactos de hipermedia (o en la Web) que cumplimenten tales características a partir del planteo de un conjunto de requerimientos no funcionales [Davis 93].

Podemos ver al desarrollo de hipermedia, como al proceso de producir de un modo efectivo artefactos de software, los cuales contienen una equilibrada mezcla de características y atributos deseados. Para que el desarrollo sea efectivo, en una relación de compromiso entre la calidad de los artefactos y los costos de desarrollo, los procesos deben ser administrados de una manera óptima y eficiente.

Para los agentes intervinientes en el desarrollo, los atributos de calidad representan una serie de restricciones a cumplir por los productos que se están construyendo, los procesos que se están empleando, y los recursos que se están asignando. Para que la descripción del modelo de proceso de desarrollo se complete es preciso definir las características y atributos que representen unidades de medida concreta a observar, analizar, evaluar, validar, y retroalimentar.

Para establecer un marco conceptual para la evaluación, podemos ver en los conceptos mencionados a tres entes intervinientes, a saber: *artefactos*, *procesos* y *recursos*. Por una parte, para comprender al proceso de desarrollo necesitamos comprender primeramente a los productos a ser desarrollados. Los atributos y la funcionalidad de los productos determinarán en parte los atributos de los procesos. Además, para desarrollar procesos óptimos deberíamos considerar el tipo de proyecto, la madurez de la organización y la relación entre actividades, productos y recursos disponibles. Por

¹ Nota: El texto de este apéndice fue extraído y adaptado ligeramente, principalmente en cuanto a tamaño y aspectos de forma, del capítulo 5 de la tesis de Magister [Olsina 98a].

otra parte, necesitamos considerar específicamente características de los procesos y de los recursos. Por ejemplo, podemos tener procesos que produzcan un artefacto ideal (que cumplimente atributos de relevancia de contenido y de enlaces, y navegabilidad, etc.) pero que sea inviable desde el punto de vista de la administración de tiempos y costos.

Lo anterior nos conduce a pensar que el objetivo básico (al menos en proyectos de mediana y gran escala) es tener procesos de calidad que nos aseguren una solución óptima y efectiva en cuanto a costos. De este modo, atributos de los productos, procesos, recursos (y hasta atributos de los constructores de proceso), afectarán al proceso de desarrollo en diferentes formas. Para ilustrarlo con un ejemplo obvio pero conclusivo podemos decir que: en la construcción de páginas de un sitio Web de una organización, el emplear los mejores expertos del mundo para tareas de diseño y autoría (ente "recurso humano" con atributos de gran habilidad y experticia), podrá resultar en un sitio Web (artefacto) de muy alta calidad pero prohibitivo en cuanto a costos. De modo que, para cualquier proceso de desarrollo, se requiere una administrada relación de compromiso entre calidad y costos.

Como indicamos previamente, una de las metas principales en el desarrollo de aplicaciones de hipermedia es producir artefactos de calidad. Por lo tanto, ¿qué implica calidad en Hipermedia?. (Primeramente debiéramos remitirnos al concepto de calidad - ver glosario). Dado que cualquier desarrollo tiene un costo asociado, podemos considerar a un artefacto de calidad como a una aplicación, componente o documento que tiene una combinación óptima del conjunto de características y atributos deseados y observables, balanceados con el costo de cumplimentar con esos atributos. La calidad no es una medida absoluta, sino relativa a un tipo de proyecto, en el contexto de una organización.

Por otra parte, la interpretación de calidad como estar en conformidad con procesos estándares, ha llevado a prescribir normas, guías de estilo, plantillas de descripción de procesos, etc. [ISO 9000-3, IEEE WPG, Olsina 98b, entre otros]. Es importante tener en cuenta que los estándares pueden proveer una estructura y una guía para producir artefactos de calidad pero por sí solos no garantizan los resultados [Fenton 96]. Procesos de calidad combinados con recursos y constructores de proceso de calidad conducen con mayor probabilidad a resultados de calidad.

De lo anterior se concluye que, para poder obtener artefactos de calidad a ésta se la debe planificar, programar y controlar. Consecuentemente, la calidad no podrá ser agregada a los artefactos al final del proceso de desarrollo sino que por el contrario se necesitará considerarla durante todo el ciclo de vida de desarrollo [ISO/IEC 12207]. La calidad es un resultado del proceso, y debe ser un factor directriz del mismo .

Tabla C.1 *Algunas Características y Atributos internos y externos de los entes Artefacto, Proceso y Recurso*

<i>Cracterist Ente</i>	Objetivos (Internos)	Subjetivos (Externos)
Artefacto	<ul style="list-style-type: none"> • Relevancia (de Contenido, de Enlaces) • Nivel de Cohesión de Controles, Permanencia, Estabilidad • Completitud • Reusabilidad • Defectos (Nodos destino ausentes, enlaces a nodos inválidos, etc.) • Tamaño (tamaño de página, cantidad de nodos, cantidad de enlaces por nodo, etc.) • Complejidad (Nivel de Interconexión, Centralidad, etc.) • Nivel de Documentación 	<ul style="list-style-type: none"> • Usabilidad (Esquemas de Organización de Estructura y Contenido, Interfaces, Retroalimentación, Ayudas, etc.) • Navegabilidad (Nivel de Interconexión, Orientación, Camino, etc.) • Confiabilidad (Alcanzabilidad de los nodos, Validez de los enlaces, etc.) • Mantenibilidad • Eficiencia (Performancia) • Calidad (por ej. como agregación de características como Usabilidad, Funcionalidad, Eficiencia, etc.) • Funcionalidad (mecanismos de Búsqueda y Recuperación, Navegabilidad, etc.)
Proceso	<ul style="list-style-type: none"> • Flexibilidad (para incorporar descripciones de proceso alternativas), Escalabilidad • Tiempo • Completitud • Reusabilidad • Performance • Defectos, Frecuencia 	<ul style="list-style-type: none"> • Costo • Calidad • Estabilidad • Mantenibilidad • Comprensibilidad • Esfuerzo •
Recurso	<ul style="list-style-type: none"> • Habilidad (agente humano) • Años de Experiencia • Tamaño (de un equipo de trabajo) • Nivel de Comunicación • Nivel de Estructuración • Velocidad (en hardware, etc.) • Capacidad de Memoria • Luz (en una oficina) 	<ul style="list-style-type: none"> • Costo (por ej. precio de un agente humano, en función de la habilidad y años de experiencia) • Productividad • Calidad • Efectividad • Nivel de Confort

En la tabla C.1 presentamos a un conjunto parcial de características y atributos observables de los entes artefactos, procesos y recursos [Fenton 91, Goldberg et al 95, Olsina 98d] que derivan en atributos mensurables directa o indirectamente.

Desde el punto de vista de la evaluación, control, y mejoramiento de procesos, artefactos y recursos, es preciso realizar mediciones sobre una o varias características. Las observaciones realizadas y los datos recolectados se pueden usar con propósitos de retroalimentación, valoración y recomendación, predicción y control de las

características de los entes. Por una parte, es oportuno recordar estos pares de máximas o principios:

1. *“No se puede medir lo que no se puede comprender”*
2. *“Si no se sabe dónde se está parado, un mapa no ayuda”*
3. *“Lo que no se puede medir no se puede controlar sistemáticamente”*

y, por otra parte, es importante recordar que una medida es un valor nominal o numérico computado a partir de un conjunto de datos observables y consistentes con la intuición, y para que tenga valor debe poseer las siguientes propiedades:

- ✓ La medida debe ser robusta: el cálculo es repetible e insensible a pequeños cambios en el entorno, herramientas, y observadores.
- ✓ La recolección de los datos debe soportar los principios de objetividad científica.
- ✓ La medida debe establecer escalas y rangos (ver sección 6.2.1)
- ✓ La medida debe ser relevante respecto al proceso, producto, o recurso
- ✓ Los datos deben ser fáciles de recolectar y validar

Desafortunadamente, en el área de hipermedia hasta el presente -escrito en 1997-, hay muy pocas métricas investigadas (por ejemplo, las más formalizadas son las contenidas en [Botafogo et al. 92] que han realizado estudios sobre métricas estructurales, respecto a complejidad, de grafos de nodos y enlaces), y, de las métricas existentes en hipermedia, falta la experiencia suficiente para interpretarlas de manera que sean relevantes, efectivas, al mismo tiempo que fáciles de recolectar y validar.

En la tabla anterior hemos seguido algunos criterios existentes en cuanto a la categorización de las métricas, y hemos dividido a las características o atributos de los entes, en internos (más objetivas) y externos (menos objetivas, es decir, características de los entes relacionados a un sujeto o agente del entorno).

En la siguiente sección, veremos mecanismos para seleccionar métricas en función de metas y objetivos, particularmente el enfoque GQM (Goal-Question-Metric) [Basili et al. 84, 94]. Dado un conjunto de metas del proyecto debemos realizar un conjunto de preguntas relevantes y, en función de esto elegir las métricas (seleccionadas de la tabla), que mejor interpreten a las preguntas. Los resultados deben servir para analizar procesos, artefactos, o recursos, con el fin de mejorar, predecir, controlar.

Con objetivos similares a los arriba mencionados, podemos utilizar a las Plantillas de Calidad (Quality Templates) [Gilb 88] como una herramienta necesaria para especificar atributos medibles (y los prerequisites previos a la medición) para un proyecto específico (ver tabla 5.4).

C.2 Goal-Question-Metric como Enfoque Orientado a Metas para Mediciones en Proyectos de Software

Como indicamos previamente, los datos recolectados de las mediciones, una vez interpretados son útiles para valorar, comprender y controlar alguna de las características de los artefactos, procesos o recursos. A partir de la interpretación y análisis de los datos, las métricas proveen de un modo sistemático, el aprendizaje de la experiencia pasada o reciente para aplicarlo a los procesos, artefactos y recursos de un proyecto actual o a situaciones futuras.

Ahora la pregunta es, ¿a partir de qué enfoque o estrategia seleccionamos las características observables (relevancia de enlaces y contenido, navegabilidad, etc.) y en qué contexto las analizamos e interpretamos? De acuerdo con investigaciones realizadas, las métricas para que sean efectivas deben estar focalizadas hacia metas específicas, aplicadas a todo o parte del ciclo de vida de los entes e interpretadas en función de la comprensión del contexto organizacional [Basili et al. 94]. Esto implica que las mediciones se deben definir bajo una estrategia de arriba hacia abajo (top-down).

El enfoque GQM responde a esa estrategia, esto es, deriva mediciones a partir de metas (goals). Dado un conjunto seleccionado de metas (goal) del proyecto en el contexto de una organización (teniendo en cuenta las características y atributos deseables de los artefactos, productos y recursos), se construye y refina un conjunto de preguntas (question) para cada meta y, en función de cada pregunta, se eligen las métricas (metric) apropiadas.

En la figura C.1 se pueden apreciar los componentes del modelo, conformando una estructura jerárquica. A seguir, describimos de un modo informativo, los tres niveles del modelo GQM [Basili et al. 94]:

- ✓ *Nivel Conceptual* (Meta): una meta se define para un ente u objeto (artefacto, etc.), para uno o varios propósitos u objetivos, con respecto a modelos de calidad (características), para algún agente humano en cumplimiento de algún rol (punto de vista), en algún contexto particular.
- ✓ *Nivel Operativo* (Pregunta): se refina un conjunto de preguntas a partir de una meta, identificando al objeto de medición con respecto a características de calidad seleccionadas para un punto de vista.
- ✓ *Nivel Cuantitativo* (Métrica): se refina a un conjunto de métricas para cada pregunta, de modo de responder a cada una de ellas cuantitativamente (a partir de un conjunto de datos recolectados de atributos observables -atributos internos o externos).

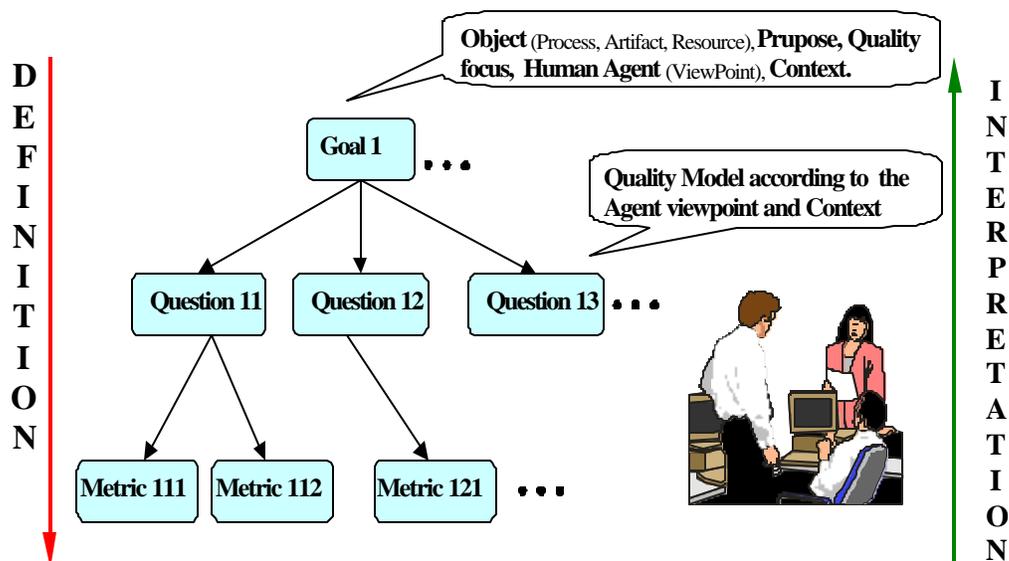


Figura C.1 Estructura Jerárquica del Enfoque GQM. Las flechas representan dos estadios relevantes del proceso: el de definición y el de interpretación.

Para ilustrar con un ejemplo el empleo de este enfoque y teniendo en cuenta una meta propia del campo de Hipermedia, supongamos que la misma consiste en *'Mejorar la navegabilidad de un hiperdocumento desde el punto de vista del usuario final'*. Se pueden formular preguntas y a partir de éstas refinar métricas. En la tabla C.2 presentamos una plantilla para capturar esta información conforme al esquema antes discutido.

Otro aspecto preponderante en el enfoque GQM es la interpretación de los datos recolectados en función de las preguntas a partir de las cuales se derivaron esas medidas. Por ejemplo, para la métrica M111 la podemos interpretar como una valoración de la complejidad de enlaces (ponderada) entre nodos de un grafo. Desde el punto de vista del usuario, una muy alta interconexión indicará que cada nodo tiene muchos puntos de partida y enlaces a nodos intermedios o destinos. Esto puede atentar contra la calidad del producto (en el atributo de navegabilidad) al permitir elegir potencialmente al mismo tiempo diversos destinos, pudiendo en definitiva desorientar al usuario. En un hiperdocumento totalmente interconectado el usuario no tiene un camino bien estructurado para seleccionar un artículo o completar un concepto distribuido en varios nodos. El caso opuesto también es generalmente un indicio de mal diseño.

En la sección C.1, afirmamos que todo proceso de desarrollo de software debe velar continuamente por tres objetivos esenciales: construir artefactos de calidad, emplear los procesos óptimos, y utilizar los recursos apropiados. Esto nos permite definir cuáles son las características y atributos observables que deben contribuir a la calidad de artefactos, procesos y recursos.

Tabla C.2 Plantilla para registrar Metas, Preguntas, Métricas y Comentarios

Meta 1		
Propósito u Objetivo		<i>Mejorar (1)</i>
Característica o Atributo		<i>Navegabilidad (2)</i>
Objeto (tipo)		<i>Hiperdocumento (artefacto) (3)</i>
Agente asignado a un rol		<i>Usuario final</i>
Pregunta	P11	<i>Cuál es el nivel de interconectividad apropiado entre nodos pertenecientes a un grafo de hipermedia?</i>
Métrica	Me111(1)	<p><i>Nivel de Interconexión (NI) = (Max - Suma) / (Max-Min)</i></p> <p><i>Max = (n²-n) * K</i></p> <p><i>n= cantidad de nodos del grafo</i></p> <p><i>K=constante superior a la cantidad de nodos</i></p> <p><i>Min= (n²-n)</i></p> <p><i>Suma representa a la suma total de distancias tomadas a partir de la matriz de distancias convertidas (con el factor K)</i></p> <p><i>Suma=$\sum_i \sum_k D_{ik}$ en donde D_{ik} representa la distancia entre los nodos ik</i></p>
	Me112	<i>Evaluación subjetiva del prototipo o artefacto en revisión conjunta con el usuario, para validar interconexión entre nodos</i>
Pregunta	P12	<i>Cuál es el nivel de alcanzabilidad óptima entre dos nodos?</i>
Métrica	Me121	<i>Distancia mínima (dado que pueden existir caminos alternativos para alcanzar dos nodos, se requiere verificar todos los caminos que permitan navegación entre los mismos)</i>
	Me122	<i>Distancia mínima promedio</i>
	Me123	<i>Desviación estándar (dada la distancia mínima promedio con respecto a cierto umbral)</i>
Comentarios	M1 (1)	<i>El objetivo de Mejorar está inserto en el contexto de un modelo de calidad a seguir para el proyecto y organización.</i>
	M1 (2)	<i>Respecto a la característica de Navegabilidad se puede ver afectada por aspectos tales como el grado de conexión entre nodos de un contexto navegacional; la existencia de caminos apropiados entre dos nodos intervinientes (por ej. la distancia más corta a partir de una deseada) y por la facilidad de alcanzarlos (orientación, ajuste de interface).</i>
	M1 (3)	<i>Hiperdocumento: documento con propiedades de hipertexto (o hipermedia)</i>
	Me111 (1)	<i>NI se refiere al grado de interconexión de nodos de un grafo (o estructura) de hipermedia. Es un valor entre 0 y 1 en donde cero implica que no hay conexiones entre nodos y uno implica conexión total. La experiencia y el análisis recomienda como apropiado (según Botafogo et al. 92) un rango entre 0.3 y 0.8.</i>

Mediante el enfoque GQM (y otras herramientas más específicas como las Plantillas de Calidad), podemos definir y planificar situaciones deseadas, que por medio del análisis de los datos relevantes, nos permitan evaluar y, en definitiva, mejorar procesos, artefactos y recursos.

Finalmente, identificaremos a un conjunto de metas que pueden ser consideradas en proyectos de hipermedia (formulando sólo el nivel conceptual del modelo GQM). La siguiente lista no pretende ser de ningún modo extensiva:

- ✓ *Evaluar la relevancia y completitud del contenido de un hiperdocumento -*
Propósito: *Evaluar* **Característica o Atributo:** *Relevancia y Completitud de contenidos* **Objeto (tipo):** *Hiperdocumento (artefacto)*
- ✓ *Reducir costos durante las tareas de requerimientos, diseño y autoría* **Propósito:** *Reducir* **Característica o Atributo:** *Costos* **Objeto (tipo):** *Tareas de requerimientos, diseño y autoría (proceso)* **Comentario:** *Esta meta se puede refinar en principio en tres submetas, relacionadas a los tres procesos*
- ✓ *Mejorar el personal en el diseño gráfico de páginas Web* **Propósito:** *Mejorar* **Característica o Atributo:** *Habilidad (diseño gráfico en la Web)* **Objeto (tipo):** *Personal (recurso)*
- ✓ *Maximizar el reuso de componentes en la fase de desarrollo* **Propósito:** *Maximizar* **Característica o Atributo:** *Reusabilidad* **Objeto (tipo):** *Componentes (artefacto)*
- ✓ *Mejorar la confiabilidad en la navegación del hiperdocumento* **Propósito:** *Mejorar* **Característica o Atributo:** *Confiabilidad* **Objeto (tipo):** *Hiperdocumento (artefacto)*
- ✓ *Mejorar la performance en actividades de autoría para generar enlaces apropiados* **Propósito:** *Mejorar* **Característica o Atributo:** *Performance* **Objeto (tipo):** *Actividades de autoría (proceso)*

C.3 Midiendo el Nivel de Interconexión en el Contexto “Ubicación” del CD-ROM “Facultad de Ingeniería”

En una sección previa, presentamos el trabajo de autoría utilizado en esta tesis –de **Magister**– para ejemplificar distintos aspectos, el cual se construyó con el modelo de proceso de hipermedia propuesto. La aplicación representa una vista del modelo conceptual de un Sistema de Información Académico [Olsina et al 95, Olsina 98a], y el perfil del usuario considerado es el del estudiante.

Primero, determinamos el alcance del trabajo de autoría a producir y capturamos preliminarmente (en una semana y media) un conjunto de requerimientos iniciales. Se analizó la factibilidad de introducir la aplicación final en un medio de almacenamiento masivo como CD-ROM (año 1995). Las decisiones de planificación consistieron en

fijar estrategias y objetivos de desarrollo, en elegir los constructores de proceso y herramientas, el tipo de usuarios que participarían en las demostraciones (en las primeras iteraciones del prototipo rápido-funcional participaron dos estudiantes y una autoridad académica; y en las demostraciones del prototipo evolutivo participaron grupos de cinco personas), y, entre otros aspectos, el modelo de calidad a seguir.

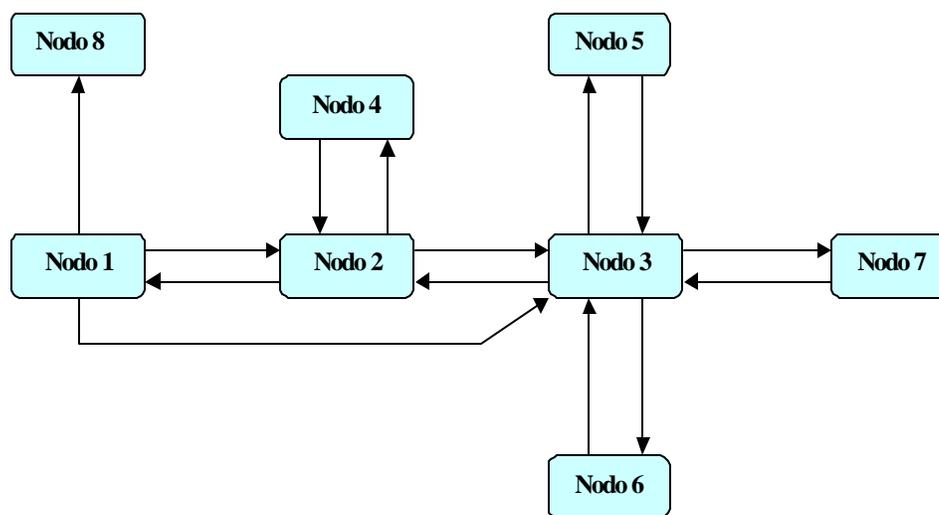


Figura C.2 Estructura o grafo de nodos del contexto “Ubicación”

A la segunda semana comenzamos a prototipar la presentación animada y la clase “*EnteFacultad*”; paralelamente comenzamos a estudiar dos contextos navegacionales “*Ubicación*” y “*Carreras*” distribuyendo el trabajo en dos subgrupos. El total de personas integrantes del equipo fueron cuatro (no considerando en este número a los usuarios participantes).

En cuanto a los atributos de calidad planteados dentro de los requerimientos no funcionales, se encontraban la usabilidad (mecanismos de índices, visitas guiadas, recorrido histórico, ajuste de interface) y aspectos de navegabilidad, entre otras características.

Para ilustrar con un ejemplo el empleo del enfoque GQM tratado en la sección anterior y, teniendo en cuenta que nos propusimos como meta mejorar la navegabilidad de la aplicación desde la perspectiva del estudiante, nos formulamos preguntas para luego observar y analizar el nivel de navegabilidad de los contextos de navegación. Por ejemplo, previo al desarrollo del prototipo del contexto “*Ubicación*” especificamos una estructura de navegación tentativa y mecanismos de recorrido de los nodos. El grafo resultante es el mostrado en la figura C.2. Si observamos los nodos y enlaces del grafo de la figura C.2, encontraremos varias de las correspondencias con objetos de las pantallas mostradas en las figuras C.3 y C.4. Los botones “*La Pampa*” y “*Gral. Pico*” de la fig. C.3.a son los anchors que conducen a los nodos 2 (fig. C.2.b) y 3 (fig. C.4.a) y

el texto en rojo (o anchor) de la misma figura conduce al nodo 8 (fig C.5.b). Asimismo podemos apreciar los controles de navegación direccionales “anterior” y “siguiente”, que se corresponden al doble enlace entre nodos del grafo.



Figura C.3 a) Implementación del nodo 1 -ver fig. C.2- del contexto “Ubicación” ; b) Implementación del nodo 2



Figura C.4 a) Implementación del nodo 3 -ver fig. C.2; b) Implementación del nodo 8, correspondiente al primer nodo del contexto “Carreras”

En la tabla C.3, presentamos una plantilla para volcar preguntas y métricas conforme a la meta planteada: "Mejorar la navegabilidad de la aplicación Facultad de Ingeniería desde la perspectiva del estudiante".

Para la formulación de la pregunta P11 encontramos dos métricas de cuyo análisis (de los valores recolectados, principalmente para la métrica M111), nos permite realizar una evaluación de los objetivos fijados con mayor rigor y objetividad. La métrica M112 consistió en una validación del prototipo en las sesiones de demostración previstas para evaluar el nivel de interconexión entre nodos (para cada contexto

navegacional). Además, las sesiones se utilizaron para evaluar la relevancia de contenidos y enlaces, y la usabilidad del prototipo.

Tabla C.3 *Plantilla para registrar Metas, Preguntas y Métricas específicas al proyecto Facultad de Ingeniería*

Meta 1		
Propósito u Objetivo		<i>Mejorar</i>
Característica o Atributo		<i>Navegabilidad</i>
Objeto (tipo)		<i>Aplicación Facultad de Ingeniería (artefacto)</i>
Agente asignado a un rol		<i>Estudiante</i>
Pregunta	P11	<i>Cuál es el nivel de interconectividad entre nodos pertenecientes a cada contexto de navegación interviniente?</i>
Métrica	Me111	<i>Nivel de Interconexión (NI) = (Max - Suma) / (Max-Min)</i>
	Me112	<i>Validación subjetiva del prototipo en revisión conjunta con el usuario final (estudiante) para evaluar el nivel de interconexión entre nodos (para cada contexto navegacional).</i>
Pregunta	P12	<i>Cuál es el nivel de alcanzabilidad óptima entre dos nodos no superior a un umbral?</i>
Métrica	Me121	<i>Distancia entre dos nodos no superior a cuatro saltos</i>
	Me122	<i>Distancia mínima promedio</i>
Comentarios	M111.1	<ul style="list-style-type: none"> ✓ <i>El valor determinado para el grafo correspondiente al contexto “Ubicación” es de 0,75 (ver fig. C.2).</i> ✓ <i>El valor determinado para el grafo correspondiente al contexto “Carrera” es de ...</i>

Por otra parte, presentamos en la figura C.5 el valor determinado de nivel de interconexión de nodos del contexto “Ubicación” conforme al grafo visto (esto se puede hacer extensivo a los contextos de navegación restantes).

De la interpretación de estos datos y de las demostraciones conjuntas efectuadas con los usuarios finales, se determinó que el nivel de interconectividad era el adecuado para este tipo aplicación. (Recordar que NI es un valor que va en un rango entre 0 y 1 en donde cero implica que no hay conexión entre nodos y uno implica conexión total. La experiencia recomienda como apropiado un rango entre 0.3 y 0.8, dependiendo del tipo de aplicación de hipertexto [Botafogo 92]).

A	Nodo	Suma	Centralidad								
Desde	1	2	3	4	5	6	7	8	Fila		
Nodo 1	0	1	1	2	2	2	2	1	11	0,067	
Nodo 2	1	0	1	1	2	2	2	2	11	0,067	
Nodo 3	2	1	0	2	1	1	1	3	11	0,067	
Nodo 4	2	1	2	0	3	3	3	3	17	0,103	
Nodo 5	3	2	1	3	0	2	2	4	17	0,103	
Nodo 6	3	2	1	3	2	0	2	4	17	0,103	
Nodo 7	3	2	1	3	2	2	0	4	17	0,103	
Nodo 8	9	9	9	9	9	9	9	0	63	0,384	
Suma Total									164		

Para $K = 9$; $n = 8$, entonces $\text{Max} = (n^2 - n) * K = 504$, y $\text{Min} = (n^2 - n) = 56$

$$\text{NI} = (\text{Max} - \text{Suma}) / (\text{Max} - \text{Min}) = 0,75$$

Figura C.5 *Matriz de Distancias Convertidas entre nodos del grafo de la figura C.2 y valor calculado de Centralidad y Nivel de Interconexión*

C4 Consideraciones Finales

Nota: Ver el capítulo 3 de Fenton y Pfleeger [Fenton et al 97], para una lectura más amplia, de algunas de las siguientes consideraciones:

- ✓ Un proceso de medición debe identificar entes, características y atributos a medir. En Ingeniería de Software, hay tres entes o clases principales de interés: 1) Procesos; 2) Artefactos; 3) Recursos.
- ✓ Cualquier propiedad o fenómeno que se desee mensurar es una característica o atributo de alguno de esos entes (o subentes).
- ✓ En general, se pueden clasificar a los atributos en internos y externos (ver además en el glosario de esta tesis, la diferencia realizada, entre característica y atributo). Un atributo interno se puede medir generalmente de un modo directo, al examinar al ente sin considerar su comportamiento externo relacionado al entorno. Un atributo externo de un artefacto, proceso o recurso se puede medir al observar como se comporta dicho ente con respecto al entorno. Por ejemplo, para medir la calidad de un ente, generalmente estamos forzados a usar métricas indirectas (características de alto nivel) de atributos internos.

- ✓ El modelo GQM es un enfoque útil para decidir qué medir. Es un enfoque orientado a metas, por lo tanto, permite a los tomadores de decisión, elegir aquellas métricas que se relacionen a las metas más importantes de los problemas más urgentes.
- ✓ Fenton y Pfleeger relacionan al enfoque GQM con el modelo de madurez de procesos CMM [Humphrey 89, Paulk et al 93, 94]. Indican que se debe considerar la madurez de un proceso (en el contexto de una organización) cuando se decide qué medir. Si una entidad no es visible en el proceso de desarrollo, no puede ser medida apropiadamente. Los niveles de madurez del modelo CMM, yendo de procesos ad hoc a optimizado, se pueden asociar al tipo de medida seleccionada por el enfoque GQM. Sería efectivo en costos y esfuerzo medir atributos y características de acuerdo a la visibilidad de procesos específicos de una organización dada.

Particularmente, dicen: “*GQM and process maturity must work hand-in-hand. By using GQM to decide what to measure an then assessing the visibility of the entity, software engineers can measure an increasingly richer set of attributes*” [Fenton et al 97].

- ✓ El enfoque GQM no trata asuntos de criterios elementales para medir un atributo, ni cuestiones de selección, estructuración, y agregación de atributos y características en el proceso de evaluación.
- ✓ Finalmente, el enfoque GQM no trata cuestiones inherentes a sistemas de complejidad media o alta en el proceso de valoración y comparación, como cantidad de atributos; relaciones entre atributos y características como simultaneidad, reemplazabilidad, y neutralidad; efectividad en costo y esfuerzo de evaluación en relación a cantidad de elementos intervinientes, entre otros asuntos.

(Ver además las críticas de Hetzel [Hetzel 93], para el enfoque GQM).

APENDICE D

Aspectos del Modelo LSP

D.1 Función para el Cálculo de las Preferencias Parciales y Globales

A seguir, presentamos un algoritmo codificado en C++, que es parte del programa para obtener las preferencias parciales y globales en el proceso de evaluación. El mismo sirve para computar funciones lógicas simétricas, a partir de los operadores lógicos de LSP, introducidos en el Capítulo 7.

```
//-----
// Máximo y mínimo de x[0],...,x[n-1]
//-----

void MINMAX (int n, double x[], double& Xmin, double& Xmax)
{
    Xmin = x[0];
    Xmax = Xmin;
    for(int i=1; i<n; i++)
        if (x[i] < Xmin)
            Xmin = x[i];
        else
            if (x[i] > Xmax) Xmax = x[i];
}

double plog(double t)
{
    return t * ( 1. - t * ( 0.5 - t / 3. ) );
}

//-----
// FUNCION          WEIGHTED POWER MEAN
//
// Xmean   = [P[0] * x[0]**r + ... + P[n-1] * x[n-1]**r] ** (1/r)
// x[0 .. n-1] = Valores de las preferencias de entrada
// P[0 .. n-1] = Peso; P[0] +...+ P[n-1] = 1
// r         = Exponente (valor real, tomado de la tabla 7.2)
// n         = Cantidad de Entradas
//-----
```

```

double WPM (int n, double P[], double r, double x[])
{
    double Rmin = 1.E-16, Rmax = 1.E+16, small = 1.E-9;
    double Xmin, Xmax, Xminlog, Xmaxlog, h, Xmean;
    int i;

    MINMAX ( n, x, Xmin, Xmax );

    if ( r < -Rmax )
        Xmean = Xmin;
    else if ( r < -Rmin )
    {
        Xmean = 0.;
        if ( Xmin > 0. )
        {
            Xminlog = log(Xmin);
            for(i=0; i<n; i++)
            {
                h = r * ( log(x[i]) - Xminlog );
                Xmean += P[i] * (1. + exp(h)) * tanh(h/2.);
            }
            if ( fabs(Xmean) > small )
                Xmean = exp( Xminlog + log(1. + Xmean)/r );
            else
                Xmean = exp( Xminlog + plog(Xmean)/r );
        }
    }
    else if ( r <= Rmin )
    {
        Xmean = 0.;
        if ( Xmin > 0. )
        {
            for(i=0; i<n; i++) Xmean += P[i] * log(x[i]);
            Xmean = exp(Xmean);
        }
    }
    else if ( r <= Rmax )
    {
        Xmean = 0.;
        if ( Xmax > 0. )
        {
            Xmaxlog = log(Xmax);
            for(i=0; i<n; i++)

```

```
if (x[i] > 0.)
{
    h = r * ( log(x[i]) - Xmaxlog );
    Xmean += P[i] * (1. + exp(h)) * tanh(h/2.);
}
else Xmean -= P[i];

if ( fabs(Xmean) > small)
    Xmean = exp( Xmaxlog + log(1. + Xmean)/r );
else
    Xmean = exp( Xmaxlog + plog(Xmean)/r );
}
}
else
//  r > Rmax
    Xmean = Xmax;

return Xmean;
}
```