

MODELO INTEGRAL PARA LA EVALUACIÓN DE LA CALIDAD DE LA ACCESIBILIDAD AL CONTENIDO WEB

Cecilia Elizabeth Gallardo¹, Ana Funes², Hernán César Ahumada¹

1: Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas, Universidad Nacional de Catamarca, Catamarca, Argentina

2: Departamento de Informática, Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis, San Luis, Argentina

COMPREHENSIVE MODEL FOR EVALUATING THE QUALITY OF WEB CONTENT ACCESSIBILITY

RESUMEN: El presente trabajo se enmarca en el área del Aseguramiento y Evaluación de la Calidad en Aplicaciones Web, dando respuesta a la necesidad de contar con un marco de referencia que permita conceptualizar y evaluar la accesibilidad al contenido web. Para ello, se propuso un modelo de calidad para la sub-característica “Accesibilidad” del modelo planteado en la norma ISO/IEC 25010, para lo cual se analizaron las Pautas de Accesibilidad al Contenido Web (WCAG) 2.0. El proceso completo de medición y evaluación se implementó siguiendo la estrategia integrada de Medición y Evaluación GOCAME (Goal-Oriented Context-Aware Measurement and Evaluation). A partir del modelo de conceptos de calidad elaborado, se desarrolló un modelo cuantitativo, basado en el método LSP (Logic Scoring of Preference), el cual permite evaluar la calidad de las aplicaciones web en cuanto a la accesibilidad de su contenido, entregando no solo un indicador global sino también indicadores parciales de acuerdo al modelo de conceptos de calidad propuesto. Otro aporte de esta tesis es un prototipo de aplicación web denominado M&ECalidadWeb, que brinda soporte y automatización al proceso de medición y evaluación de la calidad en aplicaciones web de acuerdo a las metodologías adoptadas en el trabajo de tesis.

Palabras Claves: Accesibilidad, Aplicaciones Web, Calidad, Métricas del Software, Métodos de Evaluación.

ABSTRACT: This work falls within the area of Quality Assurance and Evaluation in Web Applications, addressing the need for a reference framework to conceptualize and evaluate web content accessibility. To this end, a quality concept model corresponding to the “Accessibility” sub-characteristic of the ISO/IEC 25010 quality model was developed, for which the Web Content Accessibility Guidelines (WCAG) 2.0 were analyzed. The entire measurement and evaluation process was implemented following the guidelines of the Goal-Oriented Context-Aware Measurement and Evaluation (GOCAME) strategy. From the quality concept model, a quantitative model based on the application of the Logic Scoring of Preference (LSP) method was also developed, which allows for the evaluation of web applications’ content accessibility, providing not only a global indicator but also partial indicators according to the proposed quality concept model. Another contribution of this thesis is a web application prototype named M&ECalidadWeb, which supports and automates the quality measurement and evaluation process of web applications according to the methodologies adopted in the thesis.

KEYWORDS: Accessibility, Web Applications, Quality, Software Metrics, Evaluation Methods.

1 INTRODUCCIÓN

En la actualidad, las aplicaciones web se han convertido en una plataforma de comunicación e interacción esencial para muchas organizaciones e individuos, siendo cruciales para el comercio electrónico, el intercambio de información y una gran cantidad de actividades sociales, educativas, gubernamentales, entre muchas otras. Por esta razón, es importante que ofrezcan servicios de calidad; principalmente, un contenido accesible, que no presente barreras que dificulten o

imposibiliten la utilización del sistema web, proporcionando un acceso equitativo e igualdad de oportunidades a las personas con discapacidades o con capacidades disminuidas.

La Accesibilidad constituye un requisito de calidad, especificado mediante el Modelo de Calidad de Producto Software del estándar ISO/IEC 25010 [1] como una sub-característica de la Usabilidad. En este estándar, se define a la Accesibilidad como el “grado en que un producto o sistema puede ser utilizado por personas con la más amplia gama de características y capacidades

para lograr un objetivo determinado en un contexto de uso especificado". A su vez, existen otros trabajos provenientes del área de Ingeniería Web [2] [3] [4] [5] [6] donde, al igual que en el estándar ISO/IEC 25010, no identifican ni definen, de manera exhaustiva, los atributos que permitan medir y evaluar un concepto de alto nivel y tan complejo como la accesibilidad a nivel general, ni tampoco la accesibilidad al contenido web como un enfoque más específico.

Por otra parte, desde otro contexto y con un enfoque práctico, la iniciativa WAI (Web Accessibility Initiative) del Consorcio W3C ha desarrollado sus Pautas de Accesibilidad para el Contenido Web (WCAG) versión 2.0 [7], entendiéndose por contenido web la información dentro de una página web (texto, imágenes y sonidos, código o marcado, etc). WCAG 2.0 constituye un estándar internacional que proporciona pautas, criterios de éxito comprobables y técnicas, utilizados para evaluar los requerimientos de accesibilidad web de acuerdo a diversas necesidades.

Por ello, resulta necesario como punto de partida para la medición y evaluación de la Accesibilidad, la existencia de un modelo integral de atributos de calidad, que sirva de referencia para especificar, de manera precisa, los requisitos de accesibilidad al contenido web, tomando en cuenta las necesidades concretas de los usuarios según las aplicaciones objeto de evaluación, y que considere, además, el valioso aporte y recomendaciones de las pautas WCAG 2.0.

No obstante, para la medición y evaluación de la calidad de un producto de software, necesitamos modelos de calidad que consideren las características, atributos y relaciones que sean relevantes para un propósito dado y una necesidad de información de una categoría de entidad en concreto, como así también, es necesario contar con un enfoque transversal que guíe el proceso completo de medición y evaluación de la calidad. Es por ello que, para llevar a cabo este trabajo, se ha elegido adoptar la estrategia integrada de Medición y Evaluación (M&E) GOCAME (Goal-Oriented Context-Aware Measurement and Evaluation) [8].

GOCAME es una estrategia multipropósito que sigue un enfoque orientado a objetivos, en el cual todas las actividades son guiadas por un estado y necesidad de información específica [8]

[9]. Las actividades propuestas por la estrategia GOCAME siguen una secuencia determinada y se nutren de artefactos de entrada que surgen como salida de otra actividad. Así también, se involucran cálculos complejos, como en el caso de la obtención de valores de métricas e indicadores. Por esta razón, resulta fundamental contar con una herramienta software que gestione las distintas especificaciones que resultan de las actividades de GOCAME y que a la vez realice los cómputos de métricas e indicadores que permitan obtener los resultados finales de un proyecto de medición y evaluación de la calidad de un producto software. Teniendo en cuenta las necesidades formuladas, en este trabajo, se presenta una propuesta de soporte para la medición y evaluación de la accesibilidad al contenido web, consistente, por un lado, en un modelo de conceptos de calidad correspondiente a la sub-característica Accesibilidad del modelo de calidad ISO 25010, desarrollado analizando las Pautas de Accesibilidad al Contenido Web (WCAG) 2.0 y siguiendo los lineamientos de la estrategia integrada de Medición y Evaluación GOCAME [10]. Como parte de la estrategia adoptada, también se ha desarrollado un modelo cuantitativo, basado en el método LSP (Logical Scoring of Preference) [11] [12] [13], que permite reflejar la capacidad global del sistema evaluado para satisfacer todos los requisitos de Accesibilidad al Contenido Web. Por otro lado, como parte integral de la propuesta, se presenta como soporte a la estrategia GOCAME una aplicación web denominada M&ECalidadWeb, que permite la gestión y ejecución de artefactos para la medición y evaluación de aplicaciones web.

2 METODOLOGÍA

El presente trabajo se trata de una investigación aplicada de tipo exploratorio-descriptivo con aplicación en un contexto de estudio. Está especialmente orientado a describir, analizar y aplicar los principios y técnicas provenientes de la Ingeniería de Software e Ingeniería Web para realizar la conceptualización, medición y evaluación de la calidad del software.

Las unidades de estudio son la Ingeniería de Software e Ingeniería Web aplicadas al área de Aseguramiento de Calidad de Software y la unidad de análisis es la Accesibilidad al contenido

de aplicaciones web.

En las secciones 2.1 a 2.3, se describen las distintas etapas llevadas a cabo, de acuerdo a la metodología GOCAME, para la creación del modelo de calidad propuesto y las distintas métricas desarrolladas. En la sección 3 se describe la aplicación web desarrollada para dar soporte a la creación de modelos de calidad de acuerdo a GOCAME a la vez que se ejemplifica el proceso de implementación de la medición y evaluación usando la aplicación; en la sección 4 se describe la técnica aplicada para la validación del modelo propuesto y, finalmente, en la sección 5 se presentan las conclusiones y trabajo futuro.

2.1 Definición de Modelo de Calidad de Software

El principal artefacto resultante de la actividad “A1. Definir los requisitos no funcionales” de GOCAME es un modelo con estructura de árbol, que contiene los requerimientos no funcionales, es decir, las características, sub-características y atributos, considerados para el proceso de evaluación del aspecto de calidad planteado para el proyecto. En dicho modelo, se define como característica calculable a aquella característica que no puede ser medida directamente, sino que su valor se obtiene a partir de otras sub-características y y/o atributos que la definen; es decir, son aquellas que representan conceptos de mayor nivel de abstracción. Una sub-característica, es a su vez, una característica calculable, mientras que se habla de atributos cuando estas características pueden ser medidas en forma directa mediante una métrica y constituyen las hojas del árbol.

Para esta actividad, se consideró la sub-característica “Accesibilidad” de la característica de calidad “Usabilidad” de la norma ISO 25010 [1], enriquecida con sub-características, atributos y relaciones, que permitan evaluar la Accesibilidad al Contenido Web. Para esto se realizó un análisis exhaustivo de la documentación provista por las Pautas de Accesibilidad para el Contenido Web (WCAG) 2.0, considerando todas las pautas, principios, criterios de conformidad Nivel A y técnicas suficientes y de asesoramiento relacionadas con las tecnologías HTML (HyperText Markup Language), CSS (Cascading Style Sheets), y scripting del lado cliente y del servidor [14], obteniéndose un modelo integral

para la característica principal Accesibilidad al Contenido Web, compuesto por 28 sub-características y 60 atributos medibles (hojas del árbol). En la Tabla 1 se muestra un extracto del modelo resultante, hasta un tercer nivel de desagregación de características, con excepción de la sub-característica 1.2.4 que se muestra completamente expandida, con todas sus sub-características y atributos que intervienen en la especificación; estos últimos en letra cursiva para su mejor identificación.

2.2 Diseño de métricas

Siguiendo el proceso definido en GOCAME, la actividad “A2. Diseñar la medición”, consiste en seleccionar una métrica adecuada, desde un repositorio de métricas, para cada atributo del árbol de requerimientos, en caso de contar con tal tipo de repositorios.

Para el presente trabajo, resultó necesario definir y acordar cada una de las métricas utilizadas, debido a la inexistencia de tal repositorio. Las métricas propuestas han sido definidas para cumplir con los atributos del modelo, los cuales poseen una relación muy estrecha con las técnicas suficientes del estándar WCAG 2.0, es decir, reflejan las formas confiables de satisfacer los criterios de conformidad de la norma. Por ejemplo, para el atributo “1.2.4.3.2. Contenido dinámico convocado y ubicado adecuadamente”, hubiese bastado con definir una función binaria (100=cumple/0=no cumple). Sin embargo, dada la riqueza del método, hemos optado por la definición de una métrica que aplica una escala de proporción, que devuelve valores entre 0 y 100, donde 0 corresponde a la ausencia total de satisfacción del criterio, pasando por valores intermedios hasta llegar a 100, valor que corresponde con el cumplimiento total del atributo del árbol de requerimientos. En la Tabla 2 se puede observar la definición de la métrica que permite cuantificar al atributo 1.2.4.3.2.

<p>1. Accesibilidad al Contenido Web</p> <p>1.1. Perceptibilidad del contenido</p> <p>1.1.1. Alternativas textuales adecuadas para contenido no textual</p> <p>1.1.2. Alternativas adecuadas para medios basados en tiempo</p> <p>1.1.3. Adaptabilidad del contenido</p> <p>1.1.4. Contenido distinguible</p> <p>1.2. Navegación y componentes IU operables</p> <p>1.2.1. Funcionalidad operable completamente desde teclado</p> <p>1.2.2. Tiempo suficiente para usar el contenido web</p> <p>1.2.3. Diseño de contenido que no causa convulsiones</p> <p>1.2.4. Diseño navegable</p> <p>1.2.4.1. Acceso directo a Secciones principales del Contenido</p> <p>1.2.4.1.1. <i>Enlaces adecuados para acceder directamente a secciones principales</i></p> <p>1.2.4.1.2. <i>Mecanismo adecuado para omitir una sección complementaria</i></p> <p>1.2.4.2. Títulos adecuados de páginas web</p> <p>1.2.4.2.1. <i>Uso adecuado de título de página web</i></p> <p>1.2.4.2.2. <i>Identificación de la relación de página web actual con el conjunto de páginas a la que pertenece</i></p> <p>1.2.4.3. Orden de enfoque de componentes adecuado</p> <p>1.2.4.3.1. <i>Elementos del contenido posicionados o tabulados adecuadamente</i></p> <p>1.2.4.3.2. <i>Contenido dinámico convocado y ubicado adecuadamente</i></p> <p>1.2.4.3.3. <i>Orden de enfoque personalizado por el usuario</i></p> <p>1.2.4.4. Uso e identificación adecuada del propósito de enlaces</p> <p>1.2.4.4.1. <i>Combinación e identificación adecuada de texto del enlace con información de contexto</i></p> <p>1.2.4.4.2. <i>Determinación del valor de enlace</i></p> <p>1.2.4.4.3. <i>Personalización adecuada de texto de enlaces</i></p> <p>1.2.4.4.4. <i>Información adicional para enlaces</i></p> <p>1.3. Comprensibilidad del contenido</p> <p>1.3.1. <i>Legibilidad de contenidos textuales mediante correcta identificación del idioma</i></p> <p>1.3.2. Previsibilidad del contenido</p> <p>1.3.3. Asistencia en el ingreso de datos en formularios web</p> <p>1.4. Robustez del contenido y compatibilidad</p> <p>1.4.1. <i>Correctitud sintáctica</i></p> <p>1.4.2. <i>Validación satisfactoria de páginas web</i></p>
--

Tabla 1: Extracto del modelo de calidad para Accesibilidad al contenido web

<ul style="list-style-type: none"> • Atributo: 1.2.4.3.2. Contenido dinámico convocado y ubicado adecuadamente • Métricas relacionadas: #TA14 (Número total de páginas web); #CA12432 (Número de páginas web con contenido dinámico convocado y ubicado adecuadamente). • Método de Cálculo: $\%CA12432 = \begin{cases} CA12432 / TA14 * 100, & \text{Si } TA14 > 0 \\ 100, & \text{Si } TA14 = 0 \end{cases}$ <p>Notar que cuando #TA14 = 0, %CA12432 vale 100 ya que no se considera incumplimiento del atributo sino ausencia de ítems relacionados al mismo.</p> • Datos de Escala Numérica. Representación: continua; tipo de valor: real; tipo de escala: proporción; unidad: porcentaje.

Tabla 2: Métrica indirecta %CA12432:
Porcentaje de páginas web con contenido dinámico convocado y ubicado adecuadamente

2.3 Diseño de indicadores

En esta sección se describe, tal como lo establece el proceso de GOCAME, la actividad “A4. Diseñar la evaluación”. Notar que las actividades A3 y A5 de GOCAME no son tratadas aquí ya que corresponden a la etapa de evaluación del modelo y no a su creación.

En A4 se diseñan los indicadores, los cuales permiten interpretar los valores de los atributos a partir del resultado de cada métrica, así como de las características y sub-características de más alto nivel del árbol de requerimientos, permitiendo conocer de este modo no solo el grado de satisfacción global sino también de sub grupos de requisitos.

Se identificaron dos tipos de indicadores: 1) Indicador elemental, que interpreta el valor de un atributo mediante el uso de un modelo elemental y 2) Indicador parcial/global, cuyo valor deriva de otros indicadores de menor nivel y permite evaluar una (sub)característica calculable de nivel medio o alto de abstracción.

Las escalas de indicadores son interpretadas utilizando criterios de decisión acordes (también llamados niveles de aceptación), los cuales ayudan a analizar el nivel de satisfacción alcanzado por cada atributo, sub-característica o característica.

Respecto a los indicadores elementales, en nuestro análisis se determinó que la correspondencia entre métrica e indicador elemental es directa, ya que todas las métricas definidas poseen una escala porcentual que determina el nivel de cumplimiento del requisito. Esto es posible porque tanto la métrica y el indicador, poseen la misma escala y unidad. Por otra parte, para definir el modelo de agregación de los valores de indicadores parciales y global aplicamos el método cuantitativo de Puntuaciones Lógicas de Preferencias (LSP) [12], que se basa en técnicas de puntuación y lógica continua de preferencias y que ha sido incorporado por la metodología WebQEM [15], la que a su vez forma parte de la estrategia GOCAME.

Un modelo de agregación LSP permite computar indicadores parciales/global a partir de los indicadores elementales o parciales según el grado de anidamiento del árbol de requerimientos en el cual esté situada la sub-característica (o atributo) analizada. Cada indicador parcial hace uso de al menos otros 2 indicadores elementales

(o parciales) de menor jerarquía en el árbol de requerimientos. De esta manera, para los “n” valores de métricas correspondientes a los atributos del árbol de requerimientos, se obtienen, mediante el modelo elemental, “n” indicadores elementales. Luego, aplicando un mecanismo de agregación (o composición) paso a paso, los indicadores elementales se pueden ir agrupando convenientemente para producir una estructura de agregación final que permitirá obtener un único indicador global y múltiples indicadores parciales, que representan el grado de satisfacción de todos los requerimientos de calidad y de las distintas características y sub-características, respectivamente.

En el modelo de estructura de agregación desarrollado, básicamente se utilizaron seis operadores GCD (Generalized Conjunction Disjunction) de la lógica continua de LSP: C+, C-, C--, A, DA y D++. En la mayoría de los casos, los atributos deseables de un subárbol, se consideraron con una relación de simultaneidad, es decir, todos ellos deben ser satisfechos, modelándose dicha relación mediante el operador C-, el cual representa una relación de cuasi-conjunción débil, o bien, mediante el operador C--, que modela una relación de cuasi-conjunción más débil que C- y más cerca del operador A (media aritmética). En ambos casos, la presencia de un 0 (cero) o sea ausencia total de una característica/ atributo deseable en alguna de las entradas no producirá un 0 (cero) en la salida, si bien castigará a la misma.

La Figura 1 muestra la estructura de agregación de los atributos de la sub-características “1.2.4.4. Uso e identificación adecuada del propósito de enlaces” (ver Tabla 1), donde se requiere que el texto proporcionado con el enlace o su información de contexto sea adecuada y que a la vez se pueda determinar el valor del enlace, y como ayuda adicional se pueda personalizar el texto de un enlace o brindar información extra para el mismo (atributos 1.2.4.4.1 a 1.2.4.4.4). En este caso, los dos primeros requisitos son considerados deseables y los dos últimos opcionales, es por ello que se ha aplicado una estructura de absorción parcial de LSP, donde intervienen los operadores C- y C-- que castigan la ausencia de atributos y el operador A, que modela una relación que no posee polarización y que está en el medio

entre los operadores de reemplazabilidad y de simultaneidad (para requisitos opcionales).

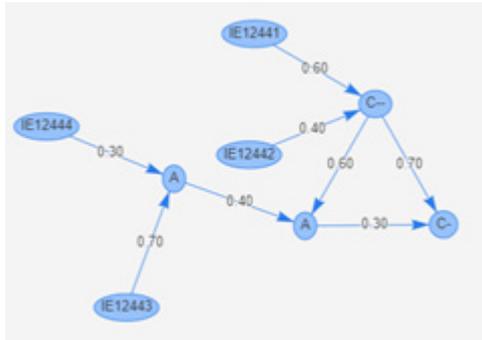


Figura 1. Estructura de agregación LSP para la subcaracterística “1.2.4.4. Uso e identificación adecuada del propósito de enlaces”.

3 Aplicación Web para la gestión y cálculo de métricas e indicadores

Con el fin de brindar soporte automatizado al evaluador durante el proceso de medición y evaluación de la calidad en Aplicaciones Web, dentro del contexto de este trabajo, se desarrolló el sistema web al que denominamos M&ECalidadWeb, el cual está basado en el marco conceptual C-INCAMI (Contextual-Information Need, Concept model, Attribute, Metric and Indicator) [8] y en las especificaciones del proceso de medición y evaluación de calidad, ambos componentes principales de la estrategia

GOGAME [10]. Así también, en algunas cuestiones, M&ECalidadWeb se ajusta al dominio del modelo desarrollado, tal como se comentará en los siguientes párrafos.

El objetivo de M&ECalidadWeb es gestionar toda la información que resulta del diseño de la medición y evaluación de la calidad en aplicaciones web, para luego calcular en forma automática los indicadores parciales y global, proceso que resulta muy complejo si se debe realizar en forma manual. M&ECalidadWeb se implementó mediante el framework de código abierto para desarrollo web Grails [16] y el lenguaje dinámico de programación Groovy [17], el cual también se ejecuta sobre la plataforma Java. Grails es un framework que se basa en el patrón arquitectónico Modelo-Vista-Controlador (MVC), que promueve la reutilización de código y la separación de conceptos, características que buscan facilitar la tarea de desarrollo de aplicaciones y su posterior mantenimiento.

3.1 Modelo de conceptos de la aplicación web

Siguiendo el marco conceptual C-INCAMI de GOCAME y las necesidades propias del modelo de calidad de la accesibilidad web se creó un modelo de clases del dominio del sistema M&ECalidadWeb, el cual refleja los principales conceptos de la aplicación y sus relaciones y que es mostrado en la Figura 2.

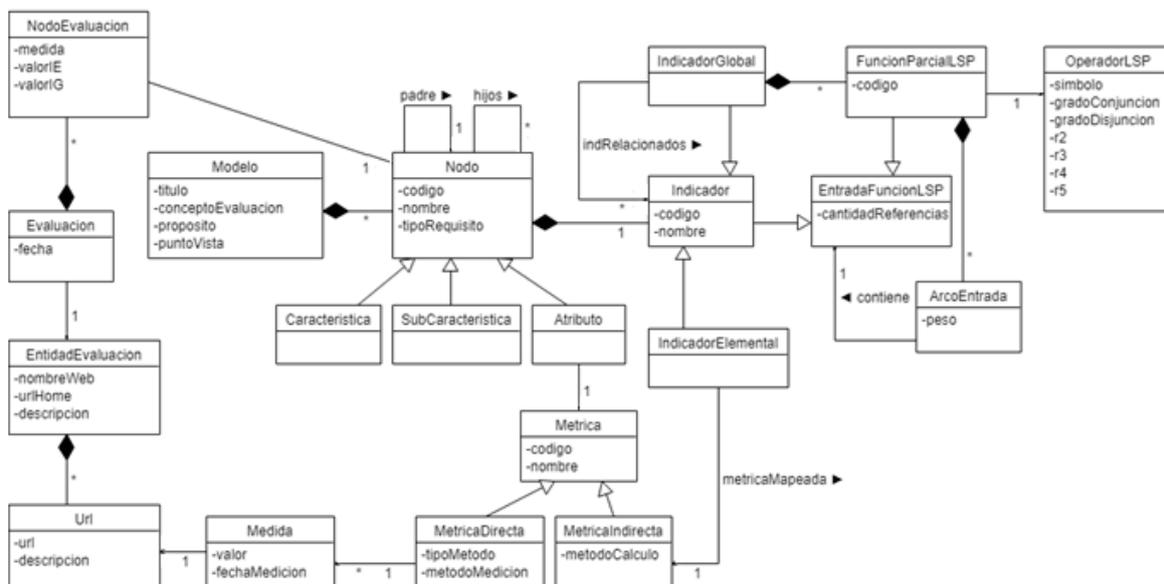


Figura 2. Modelo de clases del dominio de M&ECalidadWeb.

3.2 Principales funcionalidades de M&ECalidadWeb

A continuación, se explican las principales funciones de M&ECalidadWeb, aplicadas a un proyecto de medición y evaluación de la accesibilidad al contenido web. Se define la necesidad de información para el presente caso de estudio como “conocer, examinar y evaluar la característica de calidad externa Accesibilidad al Contenido Web de Aplicaciones Web, desde el punto de vista de usuarios con algún tipo de discapacidad total, parcial y/o personas mayores”. La entidad de evaluación seleccionada es el sitio web de ANSES, donde el subconjunto de páginas web a evaluar son las relacionadas con embarazo y nacimiento.

M&ECalidadWeb cuenta con una portada principal con un panel de opciones para acceder a toda la funcionalidad de la aplicación, tal como se muestra en la Figura 3 y se describe en las secciones 3.2.1 a 3.2.6.



Figura 3. Portada principal de la aplicación web M&ECalidadWeb

3.2.1 Gestión de modelos de calidad

Los principales atributos de un modelo de calidad, además de su título, concepto de evaluación, propósito y punto de vista de evaluación es el árbol de requerimientos, el cual se compone de la característica de evaluación principal, sub-características y atributos. En la Figura 4 se muestra un extracto del modelo de calidad desarrollado para la evaluación de la Accesibilidad al Contenido Web. M&ECalidadWeb permite especificar para cada atributo del árbol, el código

de la métrica que lo cuantifica y el código del indicador elemental correspondiente y, para el caso de las sub-características y características, mostrar el código del indicador parcial/global asociado.



Figura 4. Gestión de Modelo de Calidad en M&ECalidadWeb

3.2.2 Gestión de métricas directas e indirectas

Para las métricas directas, además de código y nombre, se puede especificar el procedimiento de medición con el cual se obtiene el valor de medida de un atributo en particular. En la mayoría de los casos, este tipo de métricas son auxiliares y se utilizan en la especificación de métricas indirectas. En el caso de las métricas indirectas, la aplicación M&ECalidadWeb permite registrar código, nombre, atributo del árbol de requerimientos a la que corresponde, métricas relacionadas y, además, un método de cálculo que consiste en un algoritmo de programación escrito en lenguaje Groovy. Esto permite realizar un cálculo automático para obtener el valor de medida del atributo correspondiente.

3.2.3 Gestión de indicadores elementales

Los indicadores elementales evalúan los atributos del árbol de requerimientos. Para el caso puntual del modelo de calidad de accesibilidad web desarrollado la correspondencia entre métrica e indicador elemental es directa, ya que todas las métricas definidas poseen una escala porcentual que se corresponde con el nivel de cumplimiento del requisito. Es por ello que, al registrarse un nuevo indicador elemental, solo se debe indicar, además de un código y nombre, a qué atributo del árbol de requerimientos corresponde, recuperándose

automáticamente la métrica asociada al atributo.

3.2.4 Gestión de indicadores parciales/global

Los indicadores parciales y global permiten evaluar una característica calculable de nivel medio o alto de abstracción, haciendo uso de un modelo especificado por el método LSP, tal como se explicó en secciones anteriores.

Al registrar un indicador parcial/global, se debe ingresar un código y nombre. M&ECalidadWeb presenta una lista con las sub-características y característica del modelo de calidad ya creado, para que el usuario seleccione una en particular y, de esta manera, el sistema pueda recuperar todos los indicadores elementales o parciales correspondientes a los atributos o conceptos anidados bajo la característica seleccionada. A continuación, se debe gestionar la estructura de agregación LSP correspondiente. En la Figura 5 se muestra la interfaz donde se ven los datos básicos de un indicador parcial/global, como así también, un grafo para representar la estructura de agregación de las funciones GCD asociadas al indicador.

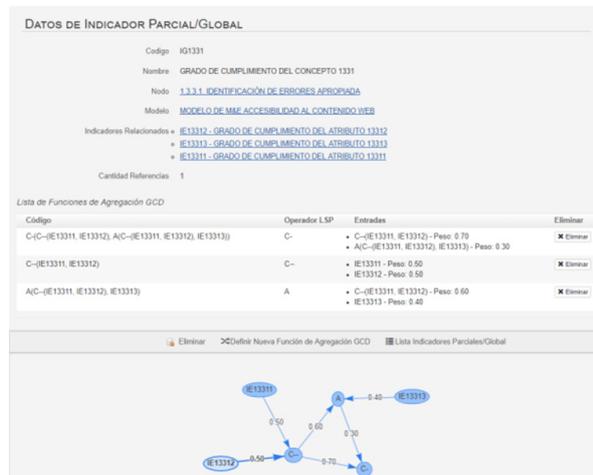


Figura 5. Gestión de indicador parcial/global en M&ECalidadWeb.

3.2.5 Gestión de operadores de agregación GCD

Las funciones u operadores de agregación GCD provistas por el método LSP permiten crear una estructura de agregación, indicando sus entradas, pesos y operadores GCD. Con el fin de establecer una entrada a una función GCD, M&ECalidadWeb muestra una lista desplegable que contiene los indicadores elementales o parciales anidados bajo

la sub-característica correspondiente al indicador seleccionado. Además, el sistema incorpora de forma dinámica a esta lista la salida de la estructura de agregación asociada al indicador en cuestión, de manera que quede disponible como entrada para un próximo paso de agregación de características. En la Figura 6, se muestra la correspondiente interfaz en M&ECalidadWeb, en donde se observa cuáles son las salidas de funciones o los indicadores elementales que pueden ser seleccionados como entrada de un nuevo operador, en este caso el operador A. Se muestra seleccionada C--(IE13311, IE13312), que representa la salida correspondiente a un operador C--, el cual ya ha agregado los indicadores elementales IE13311 y IE13312.



Figura 6. Gestión de función de agregación GCD para indicador parcial/global en M&ECalidadWeb

3.2.6 Gestión de otros parámetros

El sistema permite también gestionar las distintas entidades a evaluar, donde se debe ingresar el nombre de la aplicación web a evaluar, su URL principal y la colección de todas las URL de las páginas web que se desean revisar.

Así también, permite el registro de operadores LSP, los cuales intervienen en la creación de una estructura de agregación LSP. Para registrar un nuevo operador, M&ECalidadWeb solicita el ingreso del símbolo que lo representa, su nombre, el grado de conjunción y de disyunción, y los valores del exponente r de la función según la cantidad de entradas (de 2 a 5).

Antes de una evaluación, se deben registrar en el sistema, las medidas que corresponden a los valores de las métricas directas en base a la observación y evaluación de una página web en particular. Acá el sistema lista todas las métricas directas asociadas al modelo de calidad, para que el usuario ingrese/modifique los valores de la medición.

3.2.7 Implementación de la evaluación

Para realizar la implementación de la evaluación

propiamente dicha, el sistema solicita la selección de un modelo de calidad y una entidad a evaluar para realizar los cálculos correspondientes. Con los valores registrados de las medidas de las métricas directas del modelo, M&ECalidadWeb realiza un recorrido del árbol de requerimientos desde las hojas a la raíz, para ir obteniendo los valores de las métricas indirectas de atributos (hojas del árbol) en base al método de cálculo especificado en cada una; luego, asigna este número al valor de los indicadores elementales. Finalmente, realiza el cálculo de los indicadores parciales de las sub-características y de la característica del árbol (raíz), devolviendo el valor de preferencia de cada indicador parcial/global.

Esta funcionalidad de M&ECalidadWeb brinda un valor agregado a la fase de implementación de la evaluación, ya que de otra manera el evaluador debería realizar un sinnúmero de cálculos manuales, con altas probabilidades de cometer errores.

En la Figura 7 se muestra un fragmento del resultado de la evaluación del sitio de ANSES en base a las medidas registradas y al modelo de evaluación propuesto, exponiendo los valores de indicadores elementales y parciales/global para cada ítem del árbol de requerimientos. M&ECalidadWeb muestra, además, al lado de cada indicador, los niveles de aceptabilidad, como criterio de decisión, de acuerdo al siguiente criterio:

- Rojo (Insatisfactorio): deben tomarse acciones de cambio con una alta prioridad, identificando aquellos aspectos más críticos ($0 \leq \text{Indicador Parcial/Global} \leq 35$).
- Amarillo (Regular): necesidad de acciones de mejora ($35 < \text{Indicador Parcial/Global} \leq 70$).
- Verde (Satisfactorio): calidad satisfactoria de la característica analizada, no requiriendo medidas correctivas ($70 < \text{Indicador Parcial/Global} \leq 100$).

Requisitos de calidad para el Concepto: ACCESIBILIDAD AL CONTENIDO WEB

IE-Indicador Elemental - IPG-Indicador Parcial/Global

Características y Atributos	Medida	Valor IE	Valor IPG
1. ACCESIBILIDAD AL CONTENIDO WEB			41.5318
1.1. PERCEPTIBILIDAD DEL CONTENIDO			51.5318
1.1.1. ALTERNATIVAS TEXTUALES ADECUADAS PARA CONTENIDO NO TEXTUAL			100%
1.1.1.1. ALTERNATIVAS TEXTUALES CORTAS ADECUADAS PARA CIERTOS CONTENIDOS NO TEXTUALES			100%
1.1.1.1.1. ALTERNATIVA TEXTUAL ADECUADA PARA UNA IMAGEN	100%	100%	
1.1.1.1.2. ALTERNATIVA TEXTUAL ADECUADA PARA UN GRUPO DE IMÁGENES ADYACENTES	100%	100%	
1.1.1.1.3. ALTERNATIVA TEXTUAL ADECUADA PARA COMPOSICIONES DE CARACTERES ASCII	100%	100%	
1.1.1.1.4. ALTERNATIVA TEXTUAL ADECUADA PARA CONTENIDO DE AUDIO Y VIDEO EN DIRECTO	100%	100%	
1.1.1.1.5. ALTERNATIVA TEXTUAL ADECUADA PARA REGIONES SELECCIONABLES DE MAPA DE IMAGEN	100%	100%	
1.1.1.2. ALTERNATIVA TEXTUAL LARGA ADECUADA PARA CONTENIDO NO TEXTUAL	100%	100%	
1.1.1.3. PERCEPTIBILIDAD DE PRUEBAS CAPTCHA			100%
1.1.1.3.1. ALTERNATIVA TEXTUAL ADECUADA PARA CAPTCHA	100%	100%	
1.1.1.3.2. DIFERENTES ALTERNATIVAS PARA CUMPLIR PRUEBA CAPTCHA	100%	100%	
1.1.1.4. MECANISMO ADECUADO DE OCULTAMIENTO PARA IMÁGENES DECORATIVAS	100%	100%	
1.1.2. ALTERNATIVAS ADECUADAS PARA MEDIOS BASADOS EN TIEMPO			100%
1.1.2.1. TRANSCRIPCIÓN TEXTUAL ADECUADA PARA MEDIOS BASADOS EN TIEMPO	100%	100%	

Figura 7. Fragmento de la evaluación del sitio de ANSES en M&ECalidadWeb

4 Validación de las métricas propuestas

Al proponerse una nueva métrica de software existe el problema de demostrar que la métrica propuesta es adecuada para medir el software [18] [19], lo cual debe realizarse a través de un proceso de validación. Al respecto, durante las últimas cuatro décadas, los investigadores han debatido sobre qué constituye una métrica “válida”, centrándose en los criterios de validación de métricas de software [20]. Aun así, de la revisión de la literatura, se observa que no existe un consenso común en el concepto de validación de métricas de software, el cual especifique, taxativamente, los criterios que se deberían cumplir para considerar válida una métrica dada.

En este trabajo, para validar tanto las métricas definidas para medir los atributos del árbol de requerimientos, como las estructuras de agregación del modelo cuantitativo LSP propuesto, se ha adoptado la metodología propuesta por Meneely et al. en [21], a la que se arribó mediante una revisión sistemática de la literatura. En [21] se realiza una categorización de criterios, relaciones y ventajas para cada uno de ellos y, además, se presenta un proceso paso a paso para seleccionar los criterios apropiados de validación de métricas, basado en la intención de uso de la métrica, y observando las ventajas que brindan los diferentes criterios de acuerdo al proyecto en cuestión. En la Tabla 3 se observa la articulación de las ventajas de cada criterio de validación de métricas, enmarcados dentro de dicha metodología.

Utilizando la Tabla 3, la metodología propuesta por Meneely sugiere seleccionar las ventajas que se consideren apropiadas para el uso previsto de las métricas, que se deseen presentar a la comunidad de investigadores o al equipo de desarrollo de software.

En nuestro trabajo, el criterio adoptado de validación fue demostrar que cuando se implementan las métricas definidas, éstas permitan obtener valores que reflejen con “correctitud” (correctness) el estado del sistema evaluado. En consecuencia, se seleccionaron los criterios asociados con la correctitud, los cuales, como se ve en la tabla, son: (#10) Validez de constructo, (#20) Validez de instrumento, (#26) Fiabilidad de métrica, (#30) Validez de notación, (#35) Validez del protocolo y (#41) Estabilidad, demostrándose

#	Criterio	Solidez matemática	Practicidad	Correctitud	Eficiencia	Fortalecimiento de hipótesis	Significancia	Información para toma de decisión	Centrado en la calidad	Construcción de teoría	Contribución del consenso	Detección de diferencias
1	Validez a priori					X						
2	Accionabilidad		X					X				
3	Continuidad adecuada	X					X					
4	Granularidad adecuada											X
5	Asociación							X		X		
6	Validez de atributo						X				X	
7	Validez de modelo causal		X					X		X		
8	Validez de relación causal		X					X		X		
9	Validez de contenido						X				X	
10	Validez de constructo			X							X	
11	Constructividad							X		X		
12	Validez de la definición										X	
13	Poder discriminativo		X					X		X		
14	Consistencia dimensional	X					X					
15	Validez económica		X					X				
16	Validez empírica					X				X		
17	Validez externa							X	X	X		
18	Independencia de factores					X	X					
19	Validez de mejora				X							
20	Validez del instrumento	X		X								
21	Validez de crecimiento creciente	X										X
22	Sensibilidad a la interacción	X										X
23	Consistencia interna						X					
24	Validez interna						X					
25	Monotonidad	X										
26	Fiabilidad de métrica			X								
27	No colinealidad					X			X			
28	No explotabilidad		X									
29	No uniformidad	X					X					X
30	Validez de notación			X								X
31	Validez de permutación						X					
32	Predictibilidad		X						X	X		
33	Validez de sistema de predicción		X					X	X			
34	Relevancia de producto o proceso		X									
35	Validez de protocolo			X								X
36	Consistencia de rango					X		X	X			
37	Insensibilidad a cambio de nombre						X					
38	Repetibilidad					X				X		
39	Condición de representación	X					X					
40	Validez de escala	X										
41	Estabilidad			X								X
42	Validez teórica						X					
43	Rastreabilidad					X		X				
44	Invariancia de transformación											X
45	Validez de teoría subyacente									X		
46	Validez de unidad						X					
47	Usabilidad		X	X					X			

Tabla 3. Mapeo de criterios de validación hacia ventajas [21]

que las métricas y estructuras LSP desarrolladas cumplieran con dichos criterios de validación. No obstante, también se podría llevar a cabo una validación empírica, realizando experimentos para determinar si un grupo de personas con discapacidad 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. Luego, estos estudios podrían ser corroborados mediante coeficientes que midan el grado de consistencia y homogeneidad entre las respuestas de los encuestados, tal como el coeficiente ALPHA para la estructura interna de pruebas [22].

5 CONCLUSIONES

Un primer aporte de este trabajo consiste en un modelo de calidad compuesto por características, sub-características y atributos que permiten identificar, cuantificar y evaluar el concepto Accesibilidad al contenido en aplicaciones web. Para ello se analizaron exhaustivamente las pautas y criterios de conformidad de la norma WCAG 2.0, las cuales brindan un enfoque práctico de los aspectos que se deben cumplir para alcanzar el objetivo de la accesibilidad. Consideramos que el modelo de calidad desarrollado permite tanto a los diseñadores de sitios web como a los evaluadores de calidad, contar con criterios de evaluación de calidad, sistemáticamente definidos para satisfacer la accesibilidad al contenido web.

Otra contribución de este trabajo es un modelo cuantitativo para evaluar los atributos y características de más alto nivel que hacen a la accesibilidad del contenido web, el cual permite obtener como resultado final un indicador global del grado de satisfacción de dicha característica, a la vez que ofrece indicadores parciales de la misma. Dicho modelo de evaluación ha sido desarrollado siguiendo las pautas del método LSP, reflejando con precisión las relaciones entre atributos (simultaneidad, reemplazabilidad, neutralidad) y todas las necesidades de los usuarios finales que, en este caso, representan a aquellas personas con algún tipo de discapacidad o dificultad.

Para acompañar el proceso de medición y evaluación de la calidad en aplicaciones web siguiendo la estrategia GOCAME, se desarrolló la aplicación web denominada M&ECalidadWeb, cuyo objetivo es gestionar todos los artefactos resultantes de las distintas etapas del proceso mencionado. M&ECalidadWeb registra el modelo de calidad, diseño de métricas e indicadores y la medición de páginas web como entradas, y realiza el cálculo automático de métricas e indicadores para cada nivel del árbol de requerimientos de calidad, ahorrando al evaluador un sinfín de

cálculos manuales propensos a errores.

Tanto el modelo de conceptos de calidad, el modelo cuantitativo de evaluación y la aplicación web M&ECalidadWeb desarrollados en esta tesis, se aplicaron en el proceso de medición y evaluación de una sección del sitio web de ANSES. Esto permitió detectar varios incumplimientos de la accesibilidad al contenido web, como así también, se pudo obtener una evaluación integral de las páginas analizadas, examinando en conjunto aquellos aspectos de valoración subjetiva que no son considerados por herramientas automáticas, ya que requieren del análisis humano. Si la evaluación solo se hubiera limitado a las características de medición objetiva, la puntuación final de la evaluación del sitio hubiese estado sesgada y no sería del todo fidedigna, debido a la incidencia que tienen los atributos excluidos, que pueden disminuir o aumentar el indicador global en un valor considerable.

Finalmente, cabe mencionar que el proceso completo de medición y evaluación de la calidad del software, referido al aspecto accesibilidad al contenido web, ha sido implementado siguiendo los lineamientos de diferentes estrategias, metodologías y estándares, como lo son la estrategia GOCAME, el método LSP, la Ingeniería Web, donde se contextualiza este trabajo, y los estándares ISO 25010 y WCAG2.0. Esto ha permitido evidenciar los beneficios de utilizar disciplinas y metodologías consolidadas para formalizar el proceso de evaluación del software en general y de la característica *accesibilidad* en particular.

En cuanto al trabajo futuro, consideraremos abordar otras formas de validación, como por ejemplo la validación empírica y otras que pudieran surgir, a fines de comparar esos resultados con los resultados teóricos obtenidos.

6 REFERENCIAS

- [1] ISO/IEC 25010, "Systems and software engineering - System and software product Quality Requirements and Evaluation (SQuaRE)" - System and software quality models, 2011
- [2] G. Brajnik, "Towards Valid Quality Models for Websites", 7th Conference on Human Factors and the Web. Madison, Wisconsin, 2001

- [3] J. Offutt, "Quality Attributes of Web Software App", IEEE Software, Vol. 19, 2002, pp. 25-32
- [4] L. Hasan, E. Abuelrub, "Assessing the quality of web sites", Applied Computing and Informatics, Vol. 9, N° 1, 2011, pp. 11-29
- [5] T. Orehovacki, "Proposal for a set of quality attributes relevant for Web 2.0 application success". Information Tech. Interfaces (ITI), 2010. 32nd Int. Conf. Cavtat, Dubrovnik, 2010
- [6] R. Polillo, "Quality Models for Web 2.0, Sites: A Methodological Approach and a Proposal", LNCS. Current Trends in Web Engineering, Vol. 7059, 2011, pp. 251-265
- [7] WAI, W3C, "Web Content Accessibility Guidelines 2.0", 2008. [Online]. Disponible en: <https://www.w3.org/TR/WCAG20/>
- [8] L. Olsina, F. Papa, H. Molina, "How to Measure and Evaluate Web Applications in a Consistent Way", Web Engineering: Modelling and Implementing Web Applications. Springer-Verlag, London, 2008, pp. 385-420
- [9] L. Olsina, P. Lew, A. Dieser, B. Rivera, "Updating Quality Models for Evaluating New Generation Web Applications", Journal of Web Engineering, 2012
- [10] P. Becker, F. Papa, L. Olsina, "Enhancing the Conceptual Framework Capability for a Measurement and Evaluation Strategy". Current Trends in Web Engineering. Springer, 2013, pp. 104-116
- [11] Y. S. Stanley, J. Dujmovic, D. Batory, S. Navathe, R. Elnicki, "A cost-benefit Decision Model: Analisis, Comparison, and Selection of Data Management Systems". ACM Transactions on Database Systems, Vol. 12, N° 3, 1987, pp. 472-520
- [12] J. Dujmovic, "Continuous Preference Logic for System Evaluation". IEEE Transactions on Fuzzy Systems, Vol. 15, N° 6, 2007, pp. 1082 – 1099
- [13] J. Dujmovic, "A Method for Evaluation and Selection of Complex Hardware and Software Systems". The 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems, CMG 96 Proceedings, 1996
- [14] C. E. Gallardo, A. Funes, "Un Modelo para la Evaluación de la Calidad de la Accesibilidad al Contenido Web". CONAIISI 2015. Buenos Aires, Argentina, 2015
- [15] L. Olsina, G. Rossi, "Measuring Web Application Quality with WebQEM". IEEE Multimedia, Vol. 9, N° 4, 2002, pp. 20-29
- [16] G. Smith, P. Ledbrook, "Grails in Action". Manning Publications Co. Greenwich, 2009
- [17] A. L. Davis, "Learning Groovy". Apress, 2016
- [18] K. Srinivasan y T. Devi, "A Novel Software Metrics and Software Coding Measurement in Software Engineering", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, n° 3, 2014, pp. 303-308
- [19] K. Srinivasan y T. Devi, "A Complete and Comprehensive Metrics Suite for Object Oriented Design Quality Assessment", International Journal of Software Engineering and Its Applications, vol. 8, n° 2, 2014, pp. 173-188
- [20] K. Srinivasan y T. Devi, "Software metrics validation methodologies in software engineering", International Journal of Software Engineering & Applications, vol. 5, n° 6, 2014, pp. 87
- [21] A. Meneely, B. Smith y L. Williams, "Validating software metrics: A spectrum of philosophies", ACM Transactions on Software Engineering and Methodology (TOSEM), vol. 21, n° 4, 2013, pp. 1-28
- [22] L. J. Cronbach, "Coefficient alpha and the internal structure of tests", psychometrika, vol. 16, n° 3, 1951, pp. 297-334