

**JISBD**  
2009  
500ª

# **XIV Jornadas de Ingeniería del Software y Bases de Datos**

San Sebastián, 8-11 de septiembre de 2009

Editores:

**Antonio Vallecillo**

Dept. Lenguajes y Ciencias de la Computación

ETSI Informática

Universidad de Málaga

Bulevar Louis Pasteur, 35. 29071 Málaga. Spain

e-mail: av@lcc.uma.es

**Goiuria Sagardui**

Departamento de Electrónica e Informática

Escuela Politécnica Superior

Mondragon Unibertsitatea

Loramendi, 4; 20500 Arrasate-Mondragón. Spain

e-mail: gsagardui@eps.mondragon.edu

Filmación e impresión: Gráficas Michelena

Depósito Legal: SS-988-2009

ISBN: 978-84-692-4211-7

## Prólogo

Las XIV Jornadas de Ingeniería del Software y Bases de Datos se celebraron en San Sebastián del 8 al 11 de septiembre de 2009, en el incomparable marco del Palacio Miramar, siendo organizadas por la Escuela Politécnica Superior de la Universidad de Mondragón. Como viene ocurriendo desde la edición del 2001, se han celebrado, en paralelo y compartiendo algunos actos, de las IX Jornadas de Programación y Lenguajes (PROLE). Ambos eventos son organizados bajo los auspicios de SISTEDES, la Sociedad de Ingeniería del Software y Tecnologías de Desarrollo de Software.

Este volumen recoge los trabajos seleccionados por el Comité de Programa de JISBD 2009. Este año se han incluido tres tipos distintos de contribuciones en las actas. En primer lugar tenemos los artículos de investigación originales, que describen resultados de investigación o experiencias industriales relativas a los campos de la Ingeniería del Software y de las Bases de Datos. Se recibieron un total de 81 resúmenes preliminares, de los cuales 75 se plasmaron finalmente en artículos. Entre ellos el Comité de Programa decidió seleccionar 22 como artículos largos. Esto ha supuesto un ratio de aceptación del 29 %, lo que demuestra el arduo proceso de revisión y selección al que fueron sometidos los artículos, así como la calidad de los finalmente seleccionados. Además de estos artículos, otros 5 fueron seleccionados para participar en la conferencia como artículos cortos, con la idea de favorecer y estimular el debate científico entre los asistentes y dar cabida a la presentación de trabajos incipientes. Todos los artículos fueron revisados siguiendo un sistema de revisión por pares, y discutidos entre los miembros del Comité.

Además de este tipo de artículos, este año también se ha incluido en las actas los resúmenes de las demostraciones de herramientas presentadas, y de los artículos relevantes ya publicados. Las demostraciones de herramientas son el camino para demostrar la viabilidad práctica de las propuestas teóricas y metodológicas formuladas por los equipos de investigación, y una muestra de lo que pueden aportar a la ciencia y a la industria. En esta ocasión se seleccionaron 23 herramientas para ser brevemente presentadas durante las sesiones, y expuestas durante las Jornadas.

Por otro lado, la madurez de la comunidad JISBD se plasma en un número creciente de trabajos publicados en revistas y congresos de primera línea. Muchos de estos trabajos pasan desapercibidos para la comunidad al no tener un reflejo en las propias Jornadas, más orientadas hacia trabajos emergentes. Por ello, JISBD acoge desde las últimas dos ediciones la presentación de este tipo de publicaciones, con un doble objetivo: por una parte, publicitarlos dentro de las propias jornadas; por otra, ofrecerlos como guía y estímulo al resto de la comunidad. En esta ocasión se incluyen en las actas los resúmenes de los 10 trabajos seleccionados por la organización en esta categoría.

El éxito de la conferencia JISBD también se refleja en el número de eventos que suceden a su alrededor. En primer lugar, JISBD 2009 contó con tres conferenciantes invitados de primer nivel: Don Batory (Department of Computer Sciences, University of Texas at Austin, EE.UU.) que impartió la charla “Stepwise Development of Streaming Software Architectures”; Jean Bézivin (University of Nantes, Francia) con “Advances

in Model Driven Engineering”; y Houari Sahraoui (Universidad de Montreal, Canada), cuya charla fue “Put the Horse before the Cart: Task-Driven Development of Software Visualization Tools”. Nuestro agradecimiento más sincero por su disponibilidad para aceptar la invitación y venir a San Sebastián a impartir sus conferencias.

El programa de JISBD 2009 también incluyó dos tutoriales sobre temas de candente actualidad, como son las líneas de producto y el desarrollo dirigido por modelos.

Otro de los puntos fuertes de JISBD son los Talleres, que conforman punto de reunión obligado para los investigadores que trabajan en algunos temas de especial interés, y que ofrecen un foro de discusión excelente para estimular el debate y la colaboración entre ellos. Este año se ha contado con 6 Talleres, algunos de amplia tradición como ADIS (9ª edición), DSDM (6ª edición) o PRIS (4ª edición) y otros más noveles como son PNIS, ZOCO o WASELF.

Nos gustaría expresar nuestro más sincero agradecimiento a todos los miembros del Comité de Programa por su tiempo y dedicación a la hora de revisar y seleccionar los artículos que fueron finalmente aceptados para su presentación, y que han permitido confeccionar un año más un programa de gran calidad y nivel. Por supuesto, queremos también agradecer a los autores que enviaron artículos a las Jornadas, fueran finalmente aceptados o no, por el esfuerzo realizado y por su contribución al evento.

También queremos agradecer desde aquí al equipo organizador todo su esfuerzo y trabajo. Esto incluye a los organizadores locales de la Universidad de Mondragón, que han permitido hacer realidad esta conferencia, así como a los distintos Coordinadores que se han ocupado de organizar aspectos esenciales como los Talleres (Coral Calero), Demostraciones (Juan de Lara), Tutoriales (Ernest Teniente), Divulgación de Trabajos Relevantes ya Publicados (Belén Vela), Publicidad (Gentzane Aldekoa y José Raúl Romero), Actas (Leire Etxeberria) y Web (Ana Altuna). Nos gustaría también mostrar nuestro agradecimiento al Comité Permanente de las JISBD: primero, por depositar en nosotros su confianza a la hora de presidir el Comité de Programa y organizar la conferencia; y segundo, por su constante apoyo y soporte. Mención especial merece Oscar Díaz, cuyos consejos y ayuda han sido siempre inestimables. También mencionar el sistema de revisión de artículos que hemos utilizado, EasyChair, que fue de gran utilidad y ayuda durante todo el proceso de revisión y para la confección de estas actas.

Nuestro agradecimiento explícito a los patrocinadores del evento que hicieron posible que la conferencia fuera todo un éxito: las empresas Intersystems y Ulma, la Asociación de Técnicos en Informática (ATI), la revista Novática, las Universidades de Mondragón y del País Vasco, Cursos de Verano, el Gobierno Vasco, el Ministerio de Ciencia e Innovación, la caja de ahorros Caja Laboral y la Corporación Mondragón.

Muchas gracias a todos los asistentes a las JISBD 2009, y esperamos verles de nuevo en las próximas JISBD 2010.

San Sebastián, Septiembre 2009

Antonio Vallecillo  
Goiuria Sagardui

# Comité Ejecutivo

## **Presidente del Comité de Programa**

Antonio Vallecillo (Univ. Málaga)

## **Presidenta del Comité Organizador**

Goiuria Sagardui (Univ. Mondragón)

## **Coordinadora de Talleres**

Coral Calero (Univ. Castilla-La Mancha)

## **Coordinador de Demostraciones**

Juan de Lara (Univ. Autónoma Madrid)

## **Coordinador de Tutoriales**

Ernest Teniente (Univ. Polit. Cataluña)

## **Coordinadora de Divulgación de Trabajos Relevantes ya Publicados**

Belén Vela (Univ. Rey Juan Carlos)

## **Coordinadores de Publicidad**

Gentzane Aldekoa (Univ. Mondragón)

José Raúl Romero (Univ. Córdoba)

## **Coordinadora de actas**

Leire Etxeberria (Univ. Mondragón)

## **Coordinadora de la Web**

Ana Altuna (Univ. Mondragón)

## **Comité Organizador (Univ. Mondragón)**

Goiuria Sagardui

Joseba Andoni Aguirre

Gentzane Aldekoa

Ana Altuna

Merixell Armentia

Lorea Belategui

Itxaso Buruaga

Leire Etxeberria

Osane Lizarralde

Urtzi Markiegi

Xabier Sagarna

## Comité de Programa

Abrahão, Silvia (Univ. Polit. Valencia)  
Acuña, Cesar (Univ. Rey Juan Carlos)  
Aldana, José (Univ. Málaga)  
Álvarez, Bárbara (Univ. Polit. Cartagena)  
Aramburu, María José (Univ. Jaume I)  
Araujo, Joao (Univ. Nova Lisboa)  
Ávila, Orlando (Open Canarias)  
Bañares, Jose A. (Univ. Zaragoza)  
Berlanga, Rafael (Univ. Jaume I)  
Bertrand, Enrique (Software AG)  
Boronat, Artur (Univ. Leicester)  
Brisaboa, Nieves (Univ. Coruña)  
Cabot, Jordi (Univ. Oberta Cataluña)  
Cachero, Cristina (Univ. Alicante)  
Canal, Carlos (Univ. Málaga)  
Canos, José Hilario (Univ. Polit. Valencia)  
Carretero, Baltasar (T-Systems)  
Cavero, José María (Univ. Rey Juan Carlos)  
Corchuelo, Rafael (Univ. Sevilla)  
Costal, Dolors (Univ. Polit. Cataluña)  
Crespo, Yania (Univ. Valladolid)  
de la Riva, Claudio (Univ. Oviedo)  
Delgado Kloos, Carlos (Univ. Carlos III)  
Dolado, Javier (Univ. País Vasco)  
Duran, Francisco (Univ. Málaga)  
Fdez-Bertoa, Manuel (Univ. Málaga)  
Fdez-Medina, Eduardo (Univ. Castilla-La Mancha)  
Franch, Xavier (Univ. Polit. Cataluña)  
Fredlund, Lars-Ake (Univ. Polit. Madrid)  
Fuente, Pablo de la (Univ. Valladolid)  
Garbajosa, Juan (Univ. Polit. Madrid)  
García Molina, Jesús (Univ. Murcia)  
García, Félix (Univ. Castilla-La Mancha)  
Gaspar da Silva, Mario (Univ. Lisboa)  
Genero, Marcela (Univ. Castilla-La Mancha)  
Genova, Gonzalo (Univ. Carlos III)  
Gómez, Jaime (Univ. Alicante)  
Gonzalez, Daniel (Univ. La Laguna)  
Goñi, Alfredo (Univ. País Vasco)  
Guerra, Esther (Univ. Carlos III)  
Hernández, Juan (Univ. Extremadura)  
Hogdson, Peter (Procedimientos Uno)  
Irastorza, Arantza (Univ. País Vasco)  
Iribarne, Luis (Univ. Almeria)  
Iturrioz, Jon (Univ. País Vasco)  
Juristo, Natalia (Univ. Polit. Madrid)  
Laguna, Miguel A. (Univ. Valladolid)  
Llorens, Juan (Univ. Carlos III)  
Lopes, Antonia (Univ. Lisboa)  
Lopez Cobo, Jose M.(XimetriX)  
Lozano, Adolfo (Univ. Extremadura)  
Melia, Santiago (Univ. Alicante)  
Mena, Eduardo (Univ. Zaragoza)  
Moreira, Ana (Univ. Lisboa)  
Moreno, Ana María (Univ. Polit. Madrid)  
Pelechano, Vicente (Univ. Polit. Valencia)  
Pimentel, Ernesto (Univ. Málaga)  
Polo, Antonio (Univ. Extremadura)  
Quer, Carme (Univ. Polit. Cataluña)  
Riquelme, José (Univ. Sevilla)  
Rito, Antonio (Univ. Tec. Lisboa)  
Roda, José Luis (Univ. La Laguna)  
Ruíz, Francisco (Univ. Castilla-La Mancha)  
Ruíz-Cortés, Antonio (Univ. Sevilla)  
Sagardui, Goiuria (Univ. Mondragón)  
Samos, José (Univ. Granada)  
Sánchez, Fernando (Univ. Extremadura)  
Sánchez, Juan (Univ. Polit. Valencia)  
Sánchez, Victor (Open Canarias)  
Toro, Miguel (Univ. Sevilla)  
Toval, Ambrosio (Univ. Murcia)  
Trujillo, Juan Carlos (Univ. Alicante)  
Trujillo, Salvador (Ikerlan)  
Tuya, Javier (Univ. Oviedo)  
Urpí, Toni (Univ. Polit. Cataluña)  
Vicente, Cristina (Univ. Polit. Cartagena)

## Revisores

Miguel Ángel Martínez  
Luz Marina Moreno de Antonio  
Isabel Brito  
Nuno Cardoso  
Dante Carrizo  
Ana Cerdeira-Pena  
Antonio Corral  
Diego Alonso  
Javier Cámara  
Jose María García  
Irene Garrigós  
Anna Grimán  
Jorge García  
Ignacio García  
Sergio Ilarri  
Ernesto Jimenez  
Beatriz Bernárdez  
Joaquín Lasheras  
Manuel Llavador  
Esperanza Manso

Miguel Angel Martinez-Prieto  
Enric Mayol  
Jose-Norberto Mazon  
Fernando Molina  
Sonia Montagud  
M<sup>a</sup> Ángeles Moraga  
Isabel A. Nepomuceno-Chamorro  
Jesús Pardillo  
Juan Ángel Pastor  
Jennifer Perez  
Javier Pérez  
Manuel Resinas  
Roberto Rodríguez-Echeverría  
Roberto Ruiz  
Sergio Segura  
Manuel Serrano  
Pedro Sánchez  
Mari Carmen Otero  
Jose Zubcoff

# Patrocinadores





# Tabla de contenidos

---

## I Conferencias Invitadas

---

Advances in Model Driven Engineering . . . . .	3
<i>Jean Bézivin</i>	
Stepwise Development of Streaming Software Architectures . . . . .	4
<i>Don Batory</i>	
Put the Horse before the Cart: Task-Driven Development of Software Visualization Tools . . . . .	5
<i>Houari Sahraoui</i>	

---

## II Sesión 1. Verificación y pruebas

---

Generación de Pruebas Basada en Restricciones para Consultas SQL . . . . .	9
<i>María José Suárez-Cabal, Claudio de la Riva, Javier Tuya</i>	
Selección de Características para Mejorar los Modelos de Verificación de Información en EAI . . . . .	21
<i>Iñaki Fernández de Viana, José Luis Arjona, José Luis Álvarez, Pedro Abad</i>	
Análisis de los Efectos de las Políticas de Gestión de la Capacidad de los Servicios en el Cumplimiento de los SLAs utilizando Simulación . . . . .	33
<i>Elena Orta, Mercedes Ruiz, Miguel Toro</i>	
A tabu search algorithm for structural software testing ( <i>Artículo ya publicado</i> ) . .	45
<i>Eugenia Díaz, Javier Tuya, Raquel Blanco, José Javier Dolado</i>	
WebAVLTester: Una Herramienta de Pruebas Funcionales Automáticas para Formularios Web ( <i>demo</i> ) . . . . .	46
<i>Eugenia Díaz, Marta Fernández de Arriba, Roberto López</i>	
GAMera: una herramienta para la generación y selección mediante algoritmos genéticos de mutantes WS-BPEL ( <i>demo</i> ) . . . . .	50
<i>Antonia Estero, Inmaculada Medina, Juan José Domínguez, Lorena Gutiérrez</i>	

---

## III Sesión 2. Gestión de Proyectos

---

TUNE-UP: Seguimiento de proyectos software dirigido por la gestión de tiempos	57
<i>María Isabel Marante, Patricio Letelier, Francisco Suárez</i>	

Selección de Herramientas para la Gestión de Proyectos de Software en Pequeñas y Medianas Empresas . . . . .	69
<i>Lornel Rivas, María Pérez, Luis E. Mendoza, Anna Grimán</i>	
TEMPO: Una herramienta para la reutilización efectiva en la ingeniería de procesos software ( <i>demo</i> ) . . . . .	81
<i>Orlando Avila-García, Adolfo Sánchez-Barbudo, Víctor Roldán, Carlos González, Antonio Estévez</i>	
Una Aplicación Web de Gestión Empresarial Orientada a Proyectos para PYMEs ( <i>demo</i> ) . . . . .	85
<i>Leticia González, Miguel R. Luaces</i>	

---

## **IV Sesión 3. Requisitos / Ingeniería del software empírica**

---

Pautas para Agregar Estudios Experimentales en Ingeniería del Software . . . . .	91
<i>Enrique Fernández, Oscar Dieste, Patricia Pesado, Ramón García-Martínez</i>	
Adapting Software by Identifying Volatile and Aspectual Requirements . . . . .	103
<i>José M. Conejero, Juan Hernández, Ana Moreira, João Araújo</i>	
A Tool-supported Natural Requirements Elicitation Technique for Pervasive Systems centred on End-users . . . . .	115
<i>Francisca Pérez, Pedro Valderas</i>	

---

## **V Sesión 4. MDE y Transformaciones**

---

On the Refinement of Model-to-Text Transformations . . . . .	123
<i>Salvador Trujillo, Ander Zubizarreta, Josune de Sosa, Xabier Mendialdua</i>	
Towards Automatic Code Generation for EAI Solutions using DSL Tools . . . . .	134
<i>Hassan A. Sleiman, Abdul W. Sultán, Rafael Z. Frantz, Rafael Corchuelo</i>	
Reingeniería sobre Almacenes de Datos Seguros aplicando ADM . . . . .	146
<i>Carlos Blanco, Eduardo Fernández-Medina, Juan Trujillo</i>	
Wires* : A Tool for Orchestrating Model Transformations ( <i>demo</i> ) . . . . .	158
<i>José Eduardo Rivera, Daniel Ruiz-González, Fernando López-Romero, José María Bautista</i>	
Gra2MoL: Una Herramienta para la Extracción de Modelos en Modernización de Software ( <i>demo</i> ) . . . . .	162
<i>Javier Luis Cánovas, Jesús García</i>	

MoteGen: Una herramienta para el desarrollo de aplicaciones para redes de sensores ( <i>demo</i> ) . . . . .	166
<i>Fernando Losilla, Pedro Sánchez, Bárbara Alvarez, Diego Alonso</i>	
Generación de herramientas de modelado colaborativo independientes del dominio ( <i>demo</i> ) . . . . .	170
<i>Jesús Gallardo, Crescencio Bravo, Miguel Ángel Redondo</i>	
ModelSET Component Framework: Refinando el Ciclo de Vida de MDA ( <i>demo</i> )	174
<i>E. Victor Sanchez, Orlando Avila-García, Pablo J. Hernández, Salvador Martínez, Antonio Estévez</i>	
MOMENT2: EMF Model Transformations in Maude ( <i>demo</i> ) . . . . .	178
<i>Artur Boronat, José Meseguer</i>	

---

## **VI Sesión 5. BBDDs y Tecnologías de SGBD**

---

Philo: un Sistema Experimental de Gestión de Bases de Datos Distribuido en Memoria de Alto Rendimiento . . . . .	183
<i>Alejandro Bascuñana, Jesús Javier Arauz</i>	
Microsistemas de Información . . . . .	195
<i>Jordi Pradel, Jose Raya, Xavier Franch</i>	
Cache-aware load balancing for question answering ( <i>Artículo ya publicado</i> ) . . . .	207
<i>David Dominguez-Sal, Mihai Surdeanu, Josep Aguilar-Saborit, Josep-Lluis Larriba-Pey</i>	
An MDA approach for the development of data warehouses ( <i>Artículo ya publicado</i> ) . . . . .	208
<i>Jose-Norberto Mazón, Juan Trujillo</i>	
An Adaptive Mechanism to Protect Databases against SQL Injection . . . . .	209
<i>Cristian I. Pinzón, Juan F. De Paz, Javier Bajo, Juan M. Corchado</i>	
Un Conjunto de plugins de Eclipse para el Diseño Multidimensional de Almacenes de Datos Dirigido por Modelos ( <i>demo</i> ) . . . . .	215
<i>Octavio Glorio, Jesús Pardillo, Paul Hernández, Jose Quinto, Jose-Norberto Mazón, Juan Trujillo</i>	
OOH4RIA Tool: Una Herramienta basada en el Desarrollo Dirigido por Modelos para las RIAs ( <i>demo</i> ) . . . . .	219
<i>Santiago Meliá, Jose-Javier Martinez, Álvaro Pérez, Jaime Gómez</i>	

---

## **VII Sesión 6. Recuperación de Información, Indexación y BD en Web**

---

Indexación espacial de puntos empleando wavelet trees . . . . .	225
<i>Nieves R. Brisaboa, Miguel R. Luaces, Gonzalo Navarro, Diego Seco</i>	
Desarrollo de un compresor de textos orientado a palabras basado en PPM . . . . .	237
<i>Sandra Álvarez, Ana Cerdeira-Pena, Antonio Fariña, Susana Ladra</i>	
Reducción del Tamaño del Índice en Búsquedas por Similitud sobre Espacios Métricos . . . . .	249
<i>Luis G. Ares, Nieves R. Brisaboa, María F. Esteller, Óscar Pedreira, Ángeles S. Places</i>	
Reorganizing Compressed Text ( <i>Artículo ya publicado</i> ) . . . . .	261
<i>Nieves R. Brisaboa, Antonio Fariña, Susana Ladra, Gonzalo Navarro</i>	
Exploiting Pipeline Interruptions for Efficient Memory Allocation ( <i>Artículo ya publicado</i> ) . . . . .	262
<i>Josep Aguilar-Saborit, Mohammad Jalali, Dave Sharpe, Victor Muntés-Mulero</i>	

---

## VIII Sesión 7. Líneas de Producto

---

Towards security requirements management for software product lines: a security domain requirements engineering process ( <i>Artículo ya publicado</i> ) . . . . .	267
<i>Daniel Mellado, Eduardo Fernández-Medina, Mario Piattini</i>	
A Software Product Line Definition for Validation Environments ( <i>Artículo ya publicado</i> ) . . . . .	268
<i>Belen Magro, Jennifer Perez, Juan Garbajosa</i>	
Realizing Feature Oriented Software Development with Equational Logic: An Exploratory Study . . . . .	269
<i>Roberto E. Lopez-Herrejon, José Eduardo Rivera</i>	
Revisión Sistemática de Métricas de Calidad para Líneas de Productos Software . . . . .	275
<i>Sonia Montagud, Silvia Abrahão</i>	
FMT: Una Herramienta de Modelado y Configuración de Líneas de Productos Software para MS Visual Studio ( <i>demo</i> ) . . . . .	281
<i>Rubén Fernández, Miguel A. Laguna, Jesús Requejo, Nuria Serrano</i>	
Moskitt FM and FAMA FW: Taking feature models to the next level ( <i>demo</i> ) . . . . .	285
<i>Carlos Cetina, Pablo Trinidad, Vicente Pelechano, Antonio Ruiz-Cortés, David Benavides</i>	

---

## IX Sesión 8. Ontologías, Web semántica

---

WEAPON: Modelo de Workflow con Ontologías para Procesos Administrativos <i>Álvaro E. Prieto, Adolfo Lozano-Tello</i>	291
Populating Data Warehouses with Semantic Data <i>Victoria Nebot, Rafael Berlanga</i>	303
Towards the Semantic Desktop: the seMouse approach ( <i>Artículo ya publicado</i> ) <i>Jon Iturrioz, Oscar Díaz, Sergio F. Anzuola</i>	315
Logic-based Ontology Integration using ContentMap ( <i>demo</i> ) <i>Ernesto Jiménez-Ruiz, Bernardo Cuenca, Ian Horrocks, Rafael Berlanga</i>	316
SPARQL Query Splitter: query translation between different contexts ( <i>demo</i> ) <i>Carlos R. Osuna, David Ruiz, Rafael Corchuelo, José Luis Arjona</i>	320
Annotator: Herramienta para la Anotación Semántica de Islas de Datos Amigables en la Web ( <i>demo</i> ) <i>José L. Álvarez, José L. Arjona, Agustín Domínguez, Nicolás Amador</i>	324

---

## **X Sesión 9. Validación/Modelado Conceptual**

---

Extensión UML para Casos de Uso Reutilizables en entornos Grid Móviles Seguros <i>David G. Rosado, Eduardo Fernández-Medina, Javier López</i>	331
Perfiles UML en el diseño de notaciones visuales <i>Jesús Pardillo, Cristina Cachero</i>	343
Decidable Reasoning in UML Schemas with Constraints ( <i>Artículo ya publicado</i> ) <i>Anna Queralt, Ernest Teniente</i>	354
Rigorous Software Development Using McErlang ( <i>demo</i> ) <i>Clara Benac, Lars-Åke Fredlund</i>	355
UMLtoSBVR: Una herramienta para la validación de modelos UML mediante SBVR ( <i>demo</i> ) <i>Raquel Pau, Jordi Cabot, Ruth Raventós</i>	359
Una Aplicación basada en Eclipse para la Personalización de Aplicaciones Web Dirigida por Modelos ( <i>demo</i> ) <i>Irene Garrigós, Octavio Glorio, Paul Hernández, Alejandro Mate</i>	363
Takuan: generación dinámica de invariantes en composiciones de servicios web con WS-BPEL ( <i>demo</i> ) <i>Manuel Palomo, Antonio García, Alejandro Álvarez, Inmaculada Medina</i>	367

---

## **XI Sesión 10. Calidad, Medición y Estimación de Productos y Procesos Software**

---

Un análisis de la Calidad en Uso de los Componentes Software utilizando Redes Bayesianas . . . . .	373
<i>Manuel F. Bertoa, María Ángeles Moraga, M. Carmen Morcillo, Coral Calero</i>	
Validación empírica de medidas para procesos ETL en almacenes de datos . . . . .	387
<i>Lilia Muñoz, Jesús Pardo, Jose-Norberto Mazón, Juan Trujillo</i>	
ECAPRIS: Metodología ágil de medición de calidad y productividad en PyMEs .	399
<i>Astrid Duque, Joaquín Lasheras, Ambrosio Toval</i>	
A Literature Review for Obtaining PDQM v.2.0 . . . . .	411
<i>Carmen Moraga, María Ángeles Moraga, Coral Calero, Angélica Caro</i>	
Towards a Catalogue of Patterns for defining Metrics over i* Models ( <i>Artículo ya publicado</i> ) . . . . .	417
<i>Xavier Franch, Gemma Grau</i>	
SMT: Software Measurement Tool ( <i>demo</i> ) . . . . .	418
<i>Beatriz Mora, Francisco Ruiz, Félix García</i>	

---

## **XII Talleres**

---

Integración de Aplicaciones e Información Empresarial (ZOCO, 7ª edición) . . . .	425
<i>Rafael Corchuelo, David Ruiz, José Luis Álvarez, José Luis Arjona</i>	
Apoyo a la Decisión en Ingeniería del Software (ADIS, 9ª edición) . . . . .	426
<i>José C. Riquelme, Roberto Ruiz, Daniel Rodríguez</i>	
Pruebas en Ingeniería del Software (PRIS, 4ª edición) . . . . .	427
<i>Claudio de la Riva, Peter Hodgson, Ewout van Driel, Fergus Flaherty, Juan Garbajosa, Luis Fernández, Macario Polo, Javier Tuya</i>	
Procesos de Negocio e Ingeniería de Servicios (PNIS, 2ª edición) . . . . .	428
<i>Antonio Ruiz-Cortés, Manuel Resinas, Francisco Ruiz, Félix García</i>	
Desarrollo de Software Dirigido por Modelos (DSDM, 6ª edición) . . . . .	429
<i>José Raúl Romero, Orlando Avila-García, Vicente Pelechano</i>	
Autonomic and SELF-adaptive Systems (WASELF, 2ª edición) . . . . .	430
<i>Javier Cámara, Carlos E. Cuesta, Miguel Ángel Pérez-Toledano</i>	

---

## **XIII Tutoriales**

---

Análisis en líneas de productos: avances, desafíos y lecciones aprendidas . . . . .	433
<i>David Benavides, Antonio Ruiz-Cortés, Pablo Trinidad</i>	
Herramientas Eclipse para el Desarrollo de Software Dirigido por Modelos . . . .	434
<i>Cristina Vicente, Diego Alonso</i>	
<b>Índice de autores</b> . . . . .	435

# A Literature Review for Obtaining PDQM v.2.0

Carmen Moraga<sup>1</sup>, M<sup>a</sup> Ángeles Moraga<sup>1</sup>, Coral Calero<sup>1</sup>, Angélica Caro<sup>2</sup>,

<sup>1</sup>Alarcos Research Group – Institute of Information Technologies & Information Systems.  
Dept. Information Technologies & Systems ., Univ. of Castilla-La Mancha, Spain  
[Carmen.Moraga@alu.uclm.es](mailto:Carmen.Moraga@alu.uclm.es), {MariaAngeles.Moraga, [Coral.Calero](mailto:Coral.Calero@uclm.es)}@uclm.es

<sup>2</sup>Department of Computer Science and Information Technologies, Univ. of Bio Bio, Chile  
[mcaro@ubiobio.cl](mailto:mcaro@ubiobio.cl)

**Abstract.** Nowadays the use of Web portals allows to obtain a vast amount of data. Those users use to carry out their daily tasks. In this context, the quality of the data is a important factor. The objective of this paper is to describe the systematic literature review (SLR) carried out in order to discover the relevant attributes in data quality for Web portals, and to evaluate the interest in data quality, in this context, since 2006 (when another SLR was carried out and a Portal Data Quality Model (PDQM) was defined). As a result, we have defined PDQM v.2.0, composed by 39 relevant attributes, for the assessment of data quality in Web portals.<sup>1</sup>

**Keywords:** Data/Information quality, Web portals

## 1 Introduction

The Internet is currently an important communication environment containing interesting data and useful information. However, unnecessary, out of date or erroneous data are also included. Data Quality (DQ) is therefore an actual factor in competitiveness. One means of accessing information obtained from the Internet is through a Web portal, a site which assembles information from multiple sources.

Bearing in mind the importance of DQ, the main goal of this paper is to discover the state-of-the-art in Web portal DQ through a SLR based on a previous one [1], which covered the years 1996 to 2005, and in which 33 attributes were chosen to define PDQM (Portal Data Quality Model), a model for the assessment of Web portal DQ and focused on the data consumer point of view [2]. The current SLR covers 2006 to the end of 2008, and its objective is to identify new relevant attributes.

This paper details the results and explains the process. As original contribution, we present, on the one hand, some figures about web DQ community interests in the last years and, on the other hand, the identification of new DQ attributes to update PDQM.

---

<sup>1</sup> This work is part of the projects: INCOME (PET2006-0682-01) from Ministerio de Educación and IVISCUS (PAC08-0024-5991) from the Consejería de Educación y Ciencia (JCCM) and DQNet (TIN2008-04951-E) supported by the Spanish Ministerio of Educación y Ciencia.



The rest of this paper is organized as follows. In Section 2, the SLR process, including the planning and conduction phases, is presented. The main results obtained are reported in Section 3. Section 4 explains the analysis of the obtained attributes. The interest in data quality in Web portals is described in Section 5 and conclusions and future works are outlined in Section 6.

## 2 Review Process

This section summarizes the activities (reader can find more information in [3]) performed in each of the two main phases of the procedure for performing an SLR, as proposed by [5]:

**a) Planning the review:** The most important pre-review activities are identified by the research questions(s) that the systematic review will address, and by producing a review protocol (i.e. plan) which defines the basic review procedures. In this phase, the following steps have been carried out: (1) *Identification of the need for a review*; (2) *Specifying the research questions* and (3) *Developing a review protocol, establishing the source selection, the search terms and the inclusion and exclusion criteria*.

**b) Conducting the review:** Once the protocol has been agreed, the review can begin. In this phase, the following steps have been made:

(1) *Selection of primary studies*. The search process was completed on 31/12/2008, and 4105 papers were found. Many papers were eliminated because they were duplicated, remaining 1332 papers. These papers were then analyzed and a total of 173 papers were selected, once the inclusion and exclusion criteria had been applied, 69 papers were obtained.

(2) *Data extraction and monitoring*. Once the primary studies were chosen, the following relevant information was recovered: a) Data of the paper, including the search engine, title, year, type of publication and authors; b) Data for their classification, considering the following dimensions: DQ attributes, point of view, context, application domain, quality model, measure and tool.

## 3 Results

This section presents the “Reporting the review” phase. The papers were classified in order to answer our research questions that are presented next.

### **RQ1: “Which Web data quality attributes are addressed by researchers?”**

A total of 130 attributes were initially obtained. From them, 63 were defined for web or web site DQ and 67 were specific to Web portal. Bearing in mind that our objective is to select the most relevant attributes, those without description were discarded (20 of them). Next, we analyzed the attributes specific to Web Site or Web that could be applicable for Web portals selecting only 22 attributes. Finally, we analyzed the definition of the attributes, detecting that there were attributes with different names and related to the same concept, as a solution we grouped them, as is shown in Table 1., removing other 50 attributes. Therefore, a total of 39 attributes (67 for Web Portal + 22

for Web site or Web but applicable to Web portal – 50 synonymous) related to data quality attributes were eventually selected (see Table 2).

**Table 1:** Attributes with different name but equal concept

Attribute name	Synonymous
Accessibility	Access
Accuracy	Accurate
Amount of data	Amount of information – Appropriate Amount
Attractiveness	Detail or Appropriate level of detail or Right level of detail – Presentation
Believability	Credibility – Trustworthiness or Trustworthy– Informativeness – Reasonableness – Temporality – Veracity
Completeness	Adequacy – Sufficiency – Exactly what you need – Precision
Concise Representation	Conciseness – Association between values – Length of value
Consistent representation	Logical Consistency or Consistency
Currency	Currentness – Up-to-date or Up-to-dateness
Duplicates	Duplicity
Ease of operation	Ease of manipulation
Effectiveness	Effectively
Efficiency	Efficiently
Interactivity	Interactive
Organization	Navigability or Navigation
Readability	Human-readable – Complexity
Relevancy	Relevance
Reliability	Reliable – Free-of-error
Reputation	Authority – Authorship – Provenance
Security	Secure
Specialization	Personalization
Timeliness	Permanence – Right Time – Speed
Understandability	Ease of understanding or Ease to understand – Well-written – Comprehensiveness or Comprehensive-Coverage – Naturalness
Usability	Ease of use or Ease to use
Usefulness	Useful – Entertainment
Value-added	Quality of values or Value

**Table 2:** Attributes for Web portals Eventually Selected

Accessibility	Completeness	Ease of operation	Novelty	Response Time	Usefulness
Accuracy	Concise Representation	Effectiveness	Objectivity	Security	Validity
Amount of data	Consistent representation	Efficiency	Organization	Specialization	Value-added
Applicability	Currency	Expiration	Readability	Timeliness	Verifiability
Attractiveness	Customer Support	Flexibility	Relevancy	Traceability	
Availability	Documentation	Interactivity	Reliability	Understandability	
Believability	Duplicates	Interpretability	Reputation	Usability	

### **RQ2: “From what point of view is the Web data quality analyzed?”**

The majority of the papers found are related to the data consumer point of view. This means that researchers are more concerned about the DQ with which the consumer is provided. We can therefore affirm that it is necessary to study the DQ from other perspectives, since all the perspectives are obviously related.

### **RQ3: “In which context the Web data quality is evaluated?”**

As was previously stated, in order to attain a wide knowledge of DQ attributes the following contexts were included: Web portal, Web site and the Web in general. It must be stressed that Web sites are studied more frequently than Web portals.

### **RQ4: “Is a quality model defined?”**

A 60% of the works defined a DQ model. This means that the works are not limited to the simple definition of attributes.

**RQ5: “Do measures for Web data quality exist?”**

Only 22% of the papers do not include measures unlike the situation of some years ago. It then seems that researchers have now realized the importance of measurement.

**RQ6: “Does a tool with which to support the proposed approach exist?”**

Although the majority of the proposals define measures, only 8 of the 69 selected papers provided a support tool, which represents 12%. This reveals the difficulty of automating the proposed measures.

## 4 Obtaining PDQM v.2.0

This section provides an analysis of the PDQM attributes, the attributes obtained in our SLR and how these can be used to obtain PDQM v.2.0.

### 4.1 Comparative between PDQM and our SLR

PDQM has 33 DQ attributes [1], selected in 2005. We have made a comparison between the PDQM attributes and the attributes detected in our SLR. Thus, a set of attributes was only selected in PDQM, first column in Table 3. The second column shows the attributes chosen from SLR and in PDQM. Finally, the last column presents 6 new attributes obtained from the SLR.

**Table 3:** Attributes by origin

Attributes in PDQM not obtained in our SLR	Attributes obtained in our SLR also included in PDQM			New attributes obtained only in our SLR
Customer Support	Accessibility	Consistent Representation	Reliability	Effectiveness
Documentation	Accuracy	Currency	Reputation	Efficiency
Duplicates	Amount of data	Ease of operation	Security	Readability
Expiration	Applicability	Interactivity	Specialization	Usability
Flexibility	Attractiveness	Interpretability	Timeliness	Usefulness
Response Time	Availability	Novelty	Understandability	Verifiability
Traceability	Believability	Objectivity	Validity	
	Completeness	Organization	Value-added	
	Concise Representation	Relevancy		

### 4.2 PDQM v.2.0

Bearing in mind that PDQM was organized into four categories: *DQ intrinsic*, *DQ operational*, *DQ contextual* and *DQ representational*, [1]. We have classified the new DQ attributes in the PDQM’s categories:

- *Verifiability* is related to the references to original sources, so it must be classified into the DQ operational category. [6]
- *Effectiveness* is related to the use of adequate analytical techniques in the design of the Web Portals [7], *Efficiency* is referred to the level of performance using appropriate amounts and types of resources [8, 9], *Usability* is focused on the ease

of use and *Usefulness* is related to the content, the use of appropriate language in a specific context, and so on. Thereby, they can be classified into the DQ contextual category [8].

- *Readability* is focused on the ease of read, so it is related to DQ representational category [10].

In Fig. 1, we show PDQM with the 6 new attributes included in boldface.

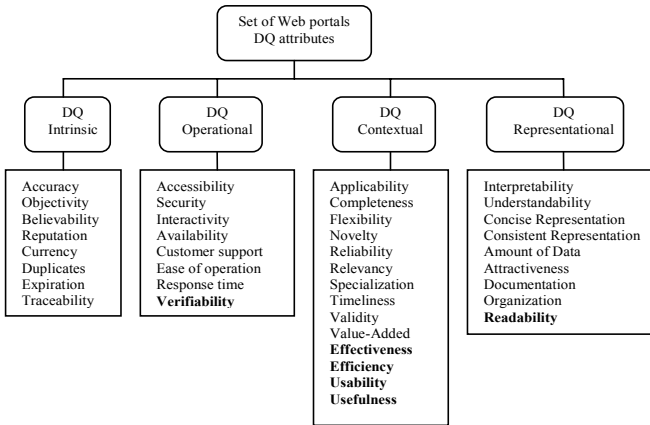


Fig. 1: PDQM v.2.0

## 5 Interest in data quality for Web portals

From the data obtained from the SLR, we prepared also an overview of the interest in data quality for web. For this, we have made a study of the selected papers per year. The number of papers per year that have been selected is shown in Fig. 2. This figure shows that researchers’ interest in data quality in the Web is increasing.

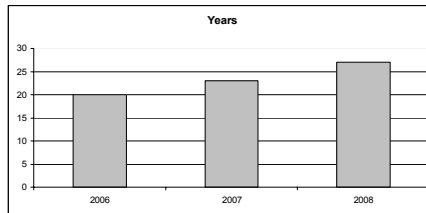


Fig. 2: Papers in each reviewed year

After analysing the types of publication of the chosen papers we can conclude that: 47 papers were published in journals, 19 in conferences, and 3 in workshops. 70% of the papers were therefore published in journals. This is perhaps due to the lack of DQ specific conferences and of topics related to DQ in non-specific conferences..

From those data we think that the interest on portals DQ is increasing.

## 6 Conclusions and future works

In this paper, a SLR has been carried out in order to obtain new DQ attributes to define a new version of PDQM (v.2.0) that includes 6 new attributes detected in this SLR together with the previous ones.

Moreover, realization of this SLR has led us to certain conclusions. Firstly, the majority of the papers study Web DQ from the data consumer's perspective. Secondly, it should be noted that Web sites are studied more frequently than Web portals. Thirdly, more than half the proposals define a DQ model. This means that the works are not limited to the simple definition of attributes. Fourthly, almost all the proposals define measures with which to assess the DQ level. Without measures it is obviously not possible to evaluate DQ. However, these measures are not easy to calculate automatically since only 12% of the proposals have developed a tool for their assessment.

As future work, we will study the new DQ standard of SQUARE [4], to align PDQM v.2.0 with the standard. With this, we will determine the final set of DQ attributes for PDQM v.2.0. The next step will be to validate the model and define the measures to evaluate the DQ in web portals using our model.

## References

1. Caro, A., Calero, C., Caballero, I., Piattini, M.: A proposal for a set of attributes relevant for Web portal data quality. *Software Quality Journal*.2008; 16(4): 513-542
2. Caro, A., Calero, C., Caballero, I., Piattini, M.: Defining a data quality model for web portals. 7th International Conference on Web Information System Engineering (WISE 2006) *Lecture Notes in Computer Science*.2006; Vol. 4255363-374
3. Moraga, C., Moraga, M., Calero, C., Caro, A.: Towards the Discovery of Data Quality Attributes for Web Portals. In: 9th International Conference on Web Engineering. (ICWE 2009). LNCS 5648, 2009. 251-259.
4. [ISO/IEC-FDIS-25012]: Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model.2008;
5. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, School of Computer Science and Mathematics, Keely University.2007;
6. Stvilia, B., Twidale, M.B., Smith, L.C., Gasser, L.: Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and Technology*.2008; 59(6): 983-1001
7. Yen, B., P.J-H., H., Wang, M.: Toward an analytical approach for effective Web site design: A framework for modeling, evaluation and enhancement. *Electronic Commerce Research and Applications*.2007; 6(2): 159-170
8. Grigoroudis, E., Litos, C., Moustakis, V.A., Politis, Y., Tsironis, L.: The assessment of user-perceived web quality: Application of a satisfaction benchmarking approach. *European Journal of Operational Research*.2008; 187(3): 1346-1357
9. ISO/IEC-FDIS-25012: Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model.2008;
10. De Wulf, K., Schillewaert, N., Muylle, S., Rangarajan, D.: The role of pleasure in web site success. *Information & Management*.2006; 43(4): 434-446