

Aplicación Web para la gestión semántica de una memoria de investigación

R. Carolina Medina-Ramírez

UAM Iztapalapa, Av. San Rafael Atlixco No 186, Vicentina, Iztapalapa, D.F., 09340, 01 (55) 5804-4636
cmed@xanum.uam.mx

Alejandra Martínez Méndez

UAM Iztapalapa, Av. San Rafael Atlixco No 186, Vicentina, Iztapalapa, D.F., 09340, 01 (55) 5804-4636
ayl0310@gmail.com

Verónica Longino Segundo

UAM Iztapalapa, Av. San Rafael Atlixco No 186, Vicentina, Iztapalapa, D.F., 09340, 01 (55) 5804-4636
veronicalons@gmail.com

José Luis García Cué

Colegio de Posgraduados, Carr. México-Texcoco km 36.5, Montecillo, Edo. de México, 01 (55) 5804-5900
jlgcue@colpos.mx

Resumen

El conocimiento que existe en una organización (memoria) está materializado en lo que se conoce comúnmente como recursos de información. Éstos son heterogéneos en formato y contenido de ahí que dependen del dominio de la organización. El gestionarlos considerando su contenido (significado o semántica) es una tarea importante en la búsqueda e integración de información. En este artículo se describe una propuesta para la gestión semántica de una memoria de investigación, se considera como caso de estudio la memoria del área de Redes y Telecomunicaciones de la UAM-Iztapalapa.

Palabras Claves: XML, memoria corporativa, XSLT, PhP.

1. Introducción

La información que produce una organización en particular una enfocada a la docencia e investigación es vasta y heterogénea tanto en contenido como en formato. El conocimiento que existe en una organización está materializado en diferentes recursos de información (documentos y/o personas) constituyendo así una memoria corporativa. Estos recursos pueden ser gestionados considerando su contenido o semántica para fines de búsqueda e integración de información, si la gestión de la memoria se realiza a través de tecnologías de la Web semántica se conoce como memoria semántica corporativa [3]. Las tecnologías de la Web semántica tienen por objetivo facilitar el enriquecimiento semántico (considerando su contenido y naturaleza) de los recursos de información existentes en la web o en webs corporativas, a través del empleo de estándares [1]. Lenguajes como XML (eXtensible Markup Language), RDF (Resource Description Framework), OWL (Ontology Web Language) son algunos ejemplos para alcanzar dichos objetivos [2]. Una de las ventajas de los sistemas que consideran el contenido de los recursos de información es el poder explotarlo, a través de mecanismos de razonamiento produciendo resultados más sensibles a las necesidades del usuario. Por otro lado, el contenido de los recursos (documentos y/o personas) es una fuente de conocimiento que puede ser explotado, capitalizado o difundido entre los miembros de una organización como al exterior de la misma. Existen trabajos que muestran las ventajas de capturar y gestionar el conocimiento que genera una organización no importando su dominio de aplicación [4, 5, 6, 7, 8]. Dentro de las ventajas están el ayudar a realizar el seguimiento de las experiencias adquiridas en procesos productivos, gestión de material didáctico considerando estilos de aprendizaje, difusión de la memoria de acuerdo a criterios como: el tipo de recurso de información, la productividad de los investigadores, reportes de actividades individuales o grupales por mencionar algunos. Además, permite establecer a corto o mediano plazo directrices para resolver problemas o realizar el mantenimiento de la memoria. La búsqueda de recursos de la memoria puede realizarse utilizando una ontología de dominio o a través de lenguajes asociados a la web semántica como XSLT o Xquery [2, 5].

La difusión de la imagen corporativa al interior como al exterior de la misma requiere de una estrategia para garantizar la visualización reciente de dicha información. Esa

estrategia por lo general recae en una persona que se encarga de hacer dicha actualización de forma manual. Ciertas vistas como: páginas web de profesores, listas de publicaciones individuales o por área de conocimiento, por tipo de evento o tipo de fuente (revistas, capítulos de libro, entre otros), lista de la formación de recursos humanos (tesistas de maestría, doctorado o licenciatura), lista de servicios sociales dirigidos, lista de proyectos de investigación financiados o institucionales, lista de laboratorios adscritos a un área de investigación son solo algunos ejemplos de información que se actualizan de manera frecuente. Por lo general, las instituciones dedicadas a la docencia e investigación cuentan con sistemas ad hoc a sus necesidades: control de la productividad de sus profesores y datos académicos de sus estudiantes, esta información por lo general se encuentra centralizada y es difícil que sea compartida abiertamente a profesores, coordinadores de licenciatura o incluso a jefes de área. La toma de decisiones a cualquier nivel depende fuertemente de contar con el mayor conocimiento y disponibilidad de la información. Por otro lado, la generación de reportes (individuales, por área de investigación, páginas web, etc.) apegados a formatos específicos son tareas que consumen mucho tiempo a los profesores quienes deben de generar diferentes reportes casi siempre con la misma información en diferentes formatos. El contar con un sistema para la gestión de una memoria de investigación es una necesidad imperante. En este artículo se describe una propuesta XML para la gestión de una memoria de investigación, se toma como caso de estudio el área de Redes y Telecomunicaciones (RyT) del Departamento de Ingeniería Eléctrica en la UAM-Iztapalapa.

2. Metodología

Se propuso una metodología de siete pasos para el diseño e implementación de una aplicación vía web para generar y gestionar una base de datos XML. Dentro de las vistas posibles tenemos: páginas web de profesores (formato de curriculum vitae), reportes de actividades de cada profesor en formato UAM-I, formatos de actividades del área de

Redes y Telecomunicaciones (RyT), listado de laboratorios asignados al área RyT, por mencionar algunas. En este artículo se describe el diseño de la aplicación haciendo énfasis en la vista páginas web de profesores para mostrar la viabilidad de la propuesta XML y tecnologías asociadas. A partir del análisis de fuentes variadas de información: reportes anuales de actividades individuales (profesores) así como el reporte anual del área RyT y formatos PROMEP, se procedió a proponer un esquema conceptual para almacenar los datos contemplados en cada formato analizado. A partir de ese modelo se propusieron diferentes hojas de estilo (XSLT +CSS) [2] para visualizar los datos respectivos correspondientes a la vista de páginas personales de profesores (contexto UAM). Los pasos de dicha metodología son:

1. Revisar los reportes anuales de las actividades individuales (de cada profesor) con el fin de obtener los metadatos necesarios para las páginas individuales.
2. Revisar páginas institucionales y personales con el fin de identificar patrones y estilos de visualización
3. Proponer un formato estándar para las páginas de profesores, incluyendo formato y contenido.
4. Proponer una estructura XML para capturar los datos correspondientes a la memoria de investigación capaz de soportar la generación de varias vistas. Por ejemplo, página web de profesores.
5. Proponer una interfaz (formulario) para capturar los datos correspondientes al reporte anual de actividades, así como los datos requeridos para generar la página personal del profesor.
6. Hacer las hojas de estilo correspondientes para generar la página HTML (index) y las correspondientes a *Thesis, Journals, Book_Chapters, Conferences, Technical_Reports*.
7. Diseñar e implementar un sistema vía web para generar las páginas personales de profesores bajo el contexto UAM, a partir de los metadatos del reporte anual de actividades (individual).

Es una propuesta modulable que puede ajustarse a otras áreas de investigación de la UAM o de otras instituciones que se dediquen a la docencia e investigación. Lo anterior

conlleva a considerar los requerimientos de los usuarios para agregar nuevos módulos (formulario, archivos XML y hoja de estilo).

3. Arquitectura y tecnologías utilizadas

En esta sección se describe la aplicación vía web desarrollada para generar y gestionar semánticamente una base de datos XML. Por el manejo semántico se debe entender el poder apegar estos metadatos a los lineamientos de los Datos Abiertos Enlazados [1].

Por otro lado, XML es un lenguaje de etiquetas utilizado no solo para el intercambio de datos sino también como un enfoque para gestionar memorias corporativas [3,4]. Este lenguaje puede ser utilizado también para el almacenamiento de datos dando pauta al enfoque de Bases de Datos XML. Una de las ventajas de este lenguaje es que se puede tener un esquema conceptual que guía el modelado de datos y de forma correspondiente el almacenamiento. La Fig. 1 ilustra el flujo de la aplicación web y cada una de las partes que la componen.

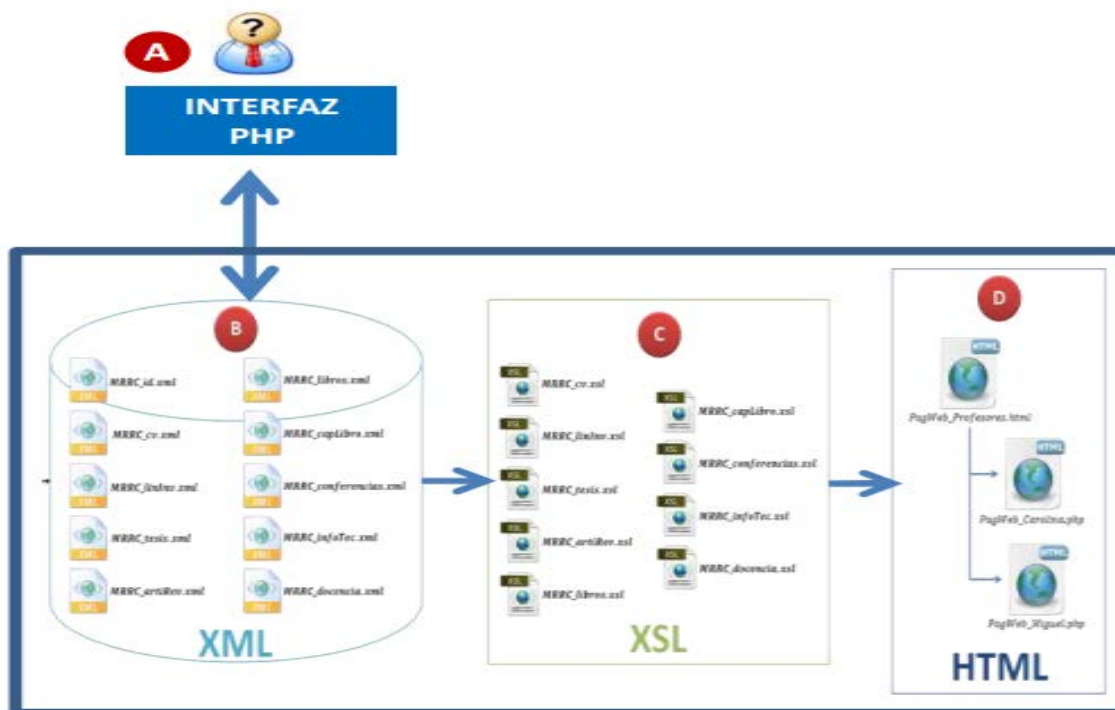


Fig. 1. Flujo de la aplicación web y cada una de las partes que la componen.

A. Interfaz - Gestión de la memoria corporativa del Área de Redes y Telecomunicaciones: Páginas Web de Profesores (Página de Inicio). La Fig. 2 muestra la página principal de la aplicación para la captura de datos (menú izquierdo) y consultas (visas) menú derecho.

Captura de Datos (formularios).

Contiene las opciones para diferentes formularios en **HTML** con formato y estilo donde el usuario introduce los datos correspondientes a los archivos: curriculum vitae <<cv.html>>, líneas de investigación <<linInv.html>>, tesis <<tesis.html>>, artículo de revista <<artiRev.html>>, libros <<libros.html>>, capítulos de libro <<capLibro.html>>, conferencias <<conferencias.html>>, informes técnicos <<infoTec.html>>, docencia <<docencia.html>>. Creamos el respectivo script de **PHP** que corresponde a cada formulario <<cv.php>>, <<linInv.php>>, <<tesis.php>>, <<artiRev.php>>, <<libros.php>>, <<capLibro.php>>, <<conferencias.php>>, <<infoTec.php>>, <<docencia.php>>.

B. Archivos XML (almacenar datos).

Por cada formulario que llenemos se creará un archivo XML instancia se tienen 10 archivos: <<cv.xml>>, <<id.xml>>, <<linInv.xml>>, <<tesis.xml>>, <<artiRev.xml>>, <<libros.xml>>, <<capLibro.xml>>, <<conferencias.xml>>, <<infoTec.xml>>, <<docencia.xml>>.

C. Archivos XSL (hojas de estilo).

Para cada XML instancia que generamos se crea una hoja de estilo para visualizar los

datos almacenados en los xml, a través de archivos XSLT: <<cv.xml>>, <<linInv.xml>>, <<tesis.xml>>, <<artiRev.xml>>, <<libros.xml>>, <<capLibro.xml>>, <<conferencias.xml>>, <<infoTec.xml>>, <<docencia.xml>>.

D. Consultas (Vistas)

Contiene una tabla con los nombres de las Páginas Web de los profesores <<profesores.html>>. Para este caso contamos con las Páginas Web de los profesores Reyna Carolina Medina Ramírez <<PagWeb_Carolina.php>> y Miguel López Guerrero <<PagWeb_Miguel.php>>.

Gestión de una memoria corporativa del Área de Redes y Telecomunicaciones: Páginas Web de Profesores

Captura de Datos

- Datos Generales
- Líneas de Investigación
- Tesis
- Artículos de Revista
- Libros
- Capítulos de libros
- Conferencias
- Informe Técnico
- Docencia
- Servicio Social
- Proyectos de Investigación (Alumnos)
- Laboratorios (RyT)
- Sitio Redes y

Casita abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

Departamento de Ingeniería Eléctrica

Área de Redes y Telecomunicaciones

ie departamento de ingeniería eléctrica

REDES Y TELECOMUNICACIONES
RYT

Versión 1.0 Estructura conceptual, captura y vistas (Página Web)

Consultas (Vistas)

Crear Página Web (Profesor)

Universidad Autónoma Metropolitana plantel Iztapalapa
San Rafael Atlixco No. 186, Col. Vicentina, Iztapalapa, 09340, México

Fig. 2. Página principal de la aplicación: Captura de datos, consultas (Vistas).

Se utilizaron para la implementación de la aplicación lenguajes como: hojas de estilo en cascada (Cascading Style Sheets) usadas para dar formato a las páginas webs escritas en lenguaje HTML provenientes de la consulta hecha a los datos almacenados en XML. PHP +xml: Herramientas SimpleXML y DomXML [2].

La estructura conceptual de la aplicación está formada por 10 archivos XML con metadatos que los describen. La tabla 1 muestra los metadatos de cada archivo XML que son considerados por el formulario correspondiente para la captura de datos.

Archivo	Descripción	Metadatos
<i>cv.xml</i>	Datos generales para conformar un Curriculum Vitae	<p>Datos personales: Nombre, apellido paterno, materno, celular, correo electrónico, foto, estado civil, nacionalidad, fecha de nacimiento, domicilio (calle, número, colonia, delegación).</p> <p>Datos académicos: Número de empleado, año de ingreso, máximo grado de estudios.</p>
<i>linInv.xml</i>	<i>líneas de investigación</i>	Nombre_línea de investigación, descripción.

<i>tesis.xml</i>	Metadatos correspondientes para tesis	Proyecto asociado, título de la tesis, nombre del alumno, grado de la tesis, año, estatus (en proceso, concluida), institución, asesores, descripción tesis, subir archivo de la tesis.
<i>artiRev.xml</i>	Metadatos correspondientes a las revistas	Título del artículo, estatus (publicado, en proceso) nombre de la revista, enlace a la revista, volumen, número de páginas, fecha publicación, ISSN, país, autores.
<i>libros.xml</i>	Libros	Título del libro, editorial, año, páginas, ISBN, enlace libro, autores.
<i>capLibro.xml</i>	Capítulos de libro	Título del libro, título capítulo, editorial, año, páginas, ISBN, número del capítulo, enlace al capítulo, autores, subir archivo.
<i>conferencias.xml</i>	Conferencias	Título, autores, nombre del evento, enlace evento, año, fecha, lugar, ciudad, país, subir archivo.
<i>infoTec.xml</i>	Informes técnicos	Nombre del informe, cargo, año, descripción, autores.
<i>docencia.xml</i>	Cursos impartidos	Nombre de la UEA, clave, grupo, trimestre, número de alumnos.

<i>En proceso</i>	Proyectos de investigación	Módulos actualmente en proceso de construcción.
	Servicios sociales	
	Página web del área de Redes y telecomunicaciones	

Tabla 1. Tamaño y estilo de fuente para los trabajos.

Para cada archivo XML se generó el formulario correspondiente para la captura de datos. El usuario tiene que llenar todos los campos, ya que son obligatorios. En la Fig. 3 se puede observar el formulario correspondiente para la captura de los metadatos de datos generales, el cual tiene una sección para datos personales como datos académicos. Para la parte de datos personales el usuario debe subir su foto personal, la cual debe ser .jpg, .png, .gif o .tiff, en caso de no ser alguno de estos archivos se le informa al usuario que lo que intenta subir no es una imagen. En el caso de datos académicos selecciona la opción del “input type radio” que corresponda. Al terminar de llenar el formulario el usuario manda a guardar la información capturada o si lo desea puede limpiar el formulario para llenarlo de nuevo.

Gestión de una memoria corporativa del Área de Redes y Telecomunicaciones: Páginas Web de Profesores

Captura de Datos

Datos Generales

Lineas de Investigación

Tesis

Artículos de Revista

Libros

Capítulos de libro

Conferencias

Informe Técnico

Docencia

Servicio Social

Proyectos de Investigación (Alumnos)

Laboratorios (RyT)

Sito Redes y Telecomunicaciones

Créditos

Datos personales:

Nombre:

Apellido Paterno:

Apellido Materno:

Estado civil:

Nacionalidad:

Fecha de nacimiento:

Domicilio

Ciudad:

Calle:

Número:

Colonia:

Delegación:

Datos académicos:

No. de empleado:

Año de ingreso:

Máximo grado académico:

Contacto

Celular:

Subir una foto:

Seleccionar archivo Ningún archivo seleccionado

Correo:

Consultas (Vistas)

Crear Página Web (Profesor)

Universidad Autónoma Metropolitana plantel Iztapalapa

San Rafael Atlixco No. 186, Col. Vicentina, Iztapalapa, 09340, México.

Fig. 3. Formulario Datos Generales (Datos personales y académicos).

De manera similar existe un formulario para cada archivo XML de la Tabla 1 con los metadatos mencionados.

4. Resultados

Después de establecer la estructura de la Aplicación Web para la gestión de la memoria corporativa del área de Redes y Telecomunicaciones y llenado cada formulario que corresponde a la *Captura de datos* de la *Página Principal* el usuario (profesor) puede

crear su *Página Web* (botón del lado derecho de la página principal). Se implementó la vista de Páginas Web de profesores, la cual contiene información correspondiente a un Currículum vitae (CV): información personal, educacional, laboral, profesional, etc., Se Hace una consulta sobre el archivo cv.xml. Al elegir la opción de crear la página web de un profesor, la aplicación nos redirecciona a otra página en la cual aparecerán los nombres de los profesores de los cuales se han creado su Página Web automáticamente. Se utiliza un mismo estilo para visualizar la página del profesor. La Fig. 4 muestra la página de la profesora Reyna Carolina Medina Ramírez.

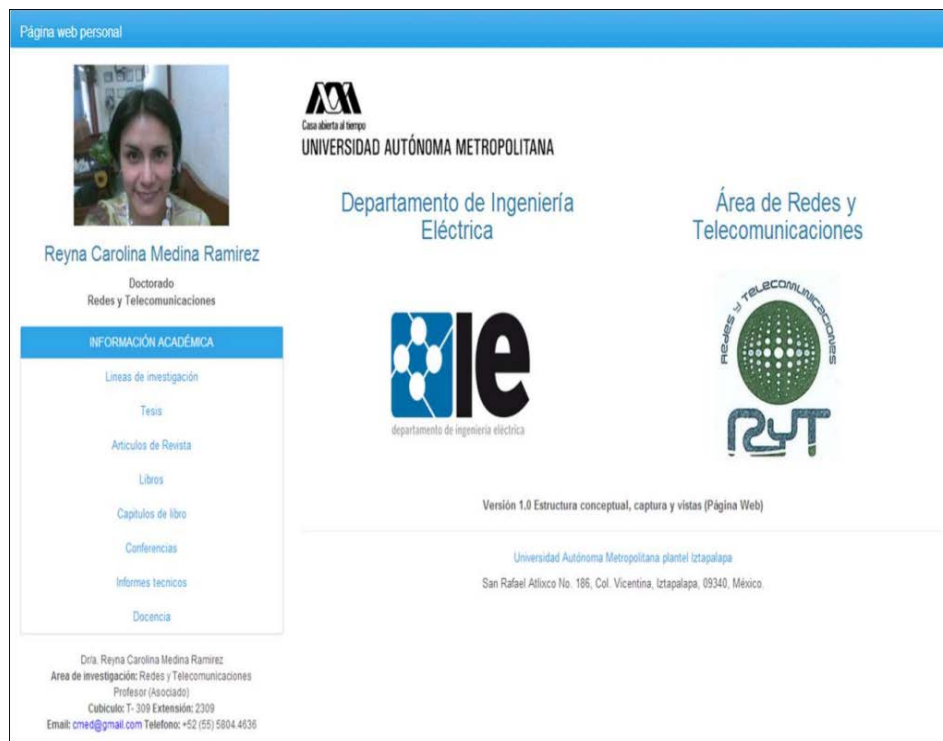


Fig. 4. Página Web de un Profesor.

5. Conclusiones

La caracterización es una fase importante en la gestión del conocimiento. El representar éste a través de estándares es un factor importante para explotarlos a posteriori, vincularlos, procesarlos e incluso hacer razonamiento sobre ellos. En este artículo se

describió una aplicación web para la gestión semántica de los recursos de información que componen una memoria de investigación. Se capturó la información de los 13 profesores adscritos al área de Redes y Telecomunicaciones (RyT). Dicha información es vasta e histórica. Se hicieron pruebas del sistema para generar el Curriculum Vitae (CV) y páginas web del profesor, así como se puso a consideración de los profesores del área RyT de quienes se obtuvieron recomendaciones sobre el formato para visualizar los datos del CV. El lenguaje XML y tecnologías asociadas fueron considerados para la implementación de la misma, mostrando la viabilidad del enfoque propuesto para la gestión de información. Como trabajo futuro, se considera ampliar la aplicación incorporando los módulos para generar los demás reportes que requiere la memoria de investigación como reportes de área, control de laboratorios, por mencionar algunos. También se vislumbra migrar hacia la vinculación de datos apegándose a los lineamientos del movimiento de datos abiertos enlazados [1]. Explotar los metadatos utilizados en el XML para aprovechar así (inferencias), el contenido de los recursos de información y mostrar la utilidad del enfoque propuesto.

6. Referencias

- [1] T. Berners-Lee, J. Hendler, O Lassila. Linked Data.- The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3):1-22. 2009.
- [2] Consorcio del Web tutoriales (w3 schools). <http://www.w3schools.com/xml/> [Último acceso: septiembre, 2014].
- [3] R. Dieng-Kuntz. Corporate Semantic Webs. *Special Issue Semantic Web. Ercim* No. 51, October 2002.
- [4] Auguste Rabarijoana, Rose Dieng, Olivier Corby. Building and Search an XML-Based Corporate Memory. *Intelligent Systems*, Volume 15 Issue 3, May 2000.
- [5] Auguste Rabarijoana, Rose Dieng, Olivier Corby. Exploitation of XML for Corporate Knowledge Management. *Knowledge Acquisition, Modeling and*

Management.Lecture Notes in Computer Science Volume 1621, pp 373-378. 1999.

- [6] Judith Hernández Ramón, José Luis García Cué, Yolanda M. Fernández Ordóñez, Reyna Carolina Medina Ramírez. Propuesta de una memoria de recursos didácticos de gestión vía web considerando su naturaleza semántica y los estilos de aprendizaje. *Review of Learning Styles*, nº9, Vol 9, abril de 2012.
- [7] Esmeralda Ramos, Iván Flores, Haydemar Núñez. Una memoria organizacional para gestionar información y conocimiento de proyectos de investigación de instituciones venezolanas. En: *Ventana Informática*. No. 28 (ene.-jun.). Manizales (Colombia): Facultad de Ciencias e Ingeniería, Universidad de Manizales. p. 117-131. ISSN: 0123-9678. 2013.
- [8] Smain Bekhti, Nada Matta, Chaker Djaiz. Knowledge representation for an efficient re-use of project memory. *Applied Computing and Informatics*, volume 9, p. 119–135. 2011.

7. Autores

Dra. Reyna Carolina Medina Ramírez. Obtuvo su grado de Doctor en Ciencias (especialidad Computación) de la Universidad de Nice-Sophia Antipolis, Francia en 2003. Desde septiembre de 2005 es Profesora Investigadora adscrita al Departamento de Ingeniería Eléctrica en la UAM unidad Iztapalapa. Actualmente es jefa del Área de investigación de Redes y Telecomunicaciones. Sus intereses de investigación incluyen web semántica y datos abiertos enlazados.

C.Verónica Longino Segundo. Estudiante de la Licenciatura en Computación de la Universidad Autónoma Metropolitana Unidad Iztapalapa.

C.Alejandra Martínez Méndez. Estudiante de la Licenciatura en Computación de la Universidad Autónoma Metropolitana Unidad Iztapalapa.

Dr. José Luis García Cué Profesor Investigador Adjunto Orientación en Estadística, Campus Montecillo Colegio de Postgraduados Sistema Nacional de Investigadores (Nivel I). Sus intereses de investigación incluyen TIC, Estilos de Aprendizaje, representación de conocimiento.