

## **Comunicación organizacional 3.0**

### **Caso de estudio Blog de comunicación de la Universidad Técnica Particular de Loja-Ecuador**

Rosario, Puertas-Hidalgo. Universidad Técnica Particular de Loja-Ecuador.

[rjpuertas@utpl.edu.ec](mailto:rjpuertas@utpl.edu.ec)

Claudia, Rodríguez-Hidalgo. Universidad Técnica Particular de Loja-Ecuador.

[cvrodriguez@utpl.edu.ec](mailto:cvrodriguez@utpl.edu.ec)

Verónica, Altamirano. Universidad Técnica Particular de Loja-Ecuador.

[vpaltamirano@utpl.edu.ec](mailto:vpaltamirano@utpl.edu.ec)

Kruzkaya, Ordóñez. Universidad Técnica Particular de Loja-Ecuador.

[kordonez@utpl.edu.ec](mailto:kordonez@utpl.edu.ec)

Nelson Piedra, Universidad Técnica Particular de Loja

[nopiedra@utpl.edu.ec](mailto:nopiedra@utpl.edu.ec)

**Resumen:** La Web Social ha facilitado la incursión del usuario común en la Red de diversas maneras: blogs, redes sociales, redes para compartir imágenes y video etc., satisfaciendo las necesidades de información y comunicación entre las personas.

Sin embargo, las necesidades de comunicación no son estáticas y evolucionan a la par de los cambios que vive la sociedad. Estamos ante el surgimiento de formas de comunicación enmarcadas en redes de datos entrelazados. La Web 3.0 o Web Semántica facilita el desarrollo de herramientas que permiten enlazar datos, con la intención de que éstos crezcan y que la Internet se convierta en una base de datos global, donde toda la información se encuentre estructurada, ordenada y sea fácilmente accesible.

La web semántica, entendida como una extensión de la web actual, permite que la información sea entregada de manera definida, y que al mismo tiempo computadores y personas puedan trabajar mejor en cooperación” Berners-Lee T. (2001). En ese sentido se debe tomar en cuenta la ontología en la web semántica, un conjunto de datos con una sintaxis y semántica, leíble tanto por humanos y máquinas” (Gruber. 1993). Estas son utilizadas por personas, bases de datos, y aplicaciones que necesitan compartir información específica sobre un determinado asunto.

En este contexto la Universidad Técnica Particular de Loja (UTPL) plantea la utilización de un software basado en web semántica como parte de los

procesos de comunicación institucional. Una primera fase abarca el blog de noticias, con el fin de organizar y describir semánticamente la información relevante de la universidad. Al momento se ha logrado definir el alcance, las fuentes de información, el esquema semántico a utilizar (las llamadas ontologías).

Esta iniciativa constituye un proyecto pionero en Ecuador, al momento las investigaciones en torno al tema de web semántica son escasas, con pocos avances al respecto. El interés es mejorar el descubrimiento y uso de herramientas digitales a través de tecnologías de la Web Semántica.

Los usuarios son autodidactas, siempre estarán familiarizándose con el entorno comunicacional y tecnológico en virtud de sus propias necesidades, nuestra misión en el ámbito de la comunicación institucional es trabajar con esos usuarios para ayudarles a encontrar los recursos informativos que requieren de la organización. Ello implica conocer ampliamente las tecnologías de la web semántica e ir desechando paulatinamente la práctica de buscar contenidos y no significados.

**Palabras clave:** comunicación; comunicación organizacional; comunicación 3.0; web semántica; web 3.0; comunicación universitaria.

## 1. Comunicación organizacional digital

Internet se ha convertido en el principal instrumento de búsqueda de información por parte de los usuarios. De ahí la necesidad que las empresas y organizaciones de todo tipo resulten atractivas y eficaces a la hora de darse a conocer a través de las herramientas que ofrece la Web, en el marco de la comunicación organizacional.

Hasta hace poco una de las formas más eficaces y conocidas de hacerlo era a través de sitios web institucionales, en los cuales cada organización ofrecía información sobre sus principales fortalezas, servicios y atractivos; en el caso de las universidades lo relacionado con la oferta educativa, proyectos, investigación y vinculación con la colectividad.

Actualmente este proceso resulta mínimo frente a las innumerables herramientas de la Web que permiten comunicar diferentes tipos de información hacia un público cada vez más diverso. Las organizaciones construyen su página web pensando sólo en la organización, como un escaparate. Aunque debe transmitir la identidad organizacional, para que se conozca quién es y cómo es la empresa, es un espacio para recibir visitas de sus públicos tanto internos como externos, satisfacer sus necesidades,

gestionar contenidos y lograr a) que regresen y b) que compartan esa satisfacción con su red de contactos personales y virtuales (Túñez. 2012: 149). Las TICs ofrecen una inagotable fuente de sorpresas. En menos de 20 años, estas tecnologías transformaron la sociedad como la conocemos (Tironi & Caballo. 2004: 229). Permiten un modelo social basado en la facilidad de acceso a la información y a los nuevos modos de interrelación individual y grupal (Túñez. 2012: 9).

En el ámbito de la comunicación organizacional, entendida como el proceso estratégico que facilita y agiliza la interacción, directa o mediada, de los públicos internos y de la organización con su entorno, con internet se multiplica, lo que supone una transformación de la cultura organizacional y en la gestión de dicha comunicación (Túñez. 2012: 15, 147).

Aunque en la etapa inicial las organizaciones se preguntaban cuál era la utilidad de tener presencia web, hoy en día Internet se ha convertido en un territorio de vida o muerte, donde son los stakeholders reales, potenciales, o futuros quienes demandan que la empresa tenga presencia allí (Tironi & Caballo. 2004: 247). A esto se suma que el consumidor es cada vez más exigente y proactivo, puede estar en varios lugares y en varias actividades al mismo tiempo. Las organizaciones consideran imprescindible la utilización del Internet, ya no esperan que sus públicos los busquen sino que salen a su encuentro (Túñez. 2012: 9, 147).

Estar en Internet es relativamente sencillo, sin embargo el objetivo de la creación de un portal web de una organización es ser reconocida, recordada y valorada. (Túñez. 2012: 149, 151). Del mismo modo las organizaciones deben tener claros varios conceptos que se deben aprovechar para lograr que los usuarios regresen:

- **Arquitectura de la información**, es el diseño estructural de los contenidos y de la funcionalidad de la página de modo que resulte más fácil y organizado el acceso a los sistemas de etiquetado de navegación y búsqueda.

- **Hipertexto o hipermedia:** hipertexto se refiere a la interactividad con el sistema para recuperar o configurar la información deseada (Díaz Noci. 2002; cit en Túñez. 2012). La hipermedia amplía la definición de hipertexto porque entiende el relato construido con varios lenguajes o varios soportes interconectados entre sí.
- **Usabilidad:** basándose en la premisa de que el tiempo de permanencia en la web es muy corto, surge como la fiabilidad, su comodidad de manejo y su efectividad. Se refiere a la rapidez con que se aprende a utilizar algo, la eficiencia al utilizarlo y cuán memorable, grado de propensión al error y cuánto le gusta a los usuarios (Túñez. 2012: 149-151).

En ese sentido cabe recordar que en las sociedades modernas la “cultura está extremadamente influenciada por los fenómenos de la globalización y la digitalización” (Castells, 2009: 2), lo cual ha derivado en que cada vez más los ciudadanos estén familiarizados con las TICs.

Internet es un soporte de comunicación, es parte de la gestión de estrategias de comunicación interna y externa en entornos virtuales, es una herramienta de uso estratégico en las planificaciones de marketing y un nuevo escenario de negocio. Es un medio que canaliza y gestiona las relaciones con los públicos, se convierte en el punto de encuentro con periodistas proactivos que acuden a la entidad por información. En el mundo virtual no hay lugar para esconderse. Finalmente, marca un nuevo entorno y un nuevo soporte comunicativo entre los ciudadanos y la organización en un modelo de comunicación multidireccional (Túñez. 2012: 156-161).

Según Kotler (2009: 771. Cit. en Túñez.2012: 163) señala que uno de los diez errores más comunes de marketing es que la empresa que no use las tecnologías todo lo que podría. Esta premisa es también aplicable a la comunicación on line, por lo que es necesario conocer las herramientas de comunicación virtual más frecuentes:

# Actas – V Congreso Internacional Latina de Comunicación Social – V CILCS – Universidad de La Laguna, diciembre 2013

---

**Tabla 1. Herramientas de comunicación organización on line**

HERRAMIENTA	DEFINICIÓN
<b>Atención al público</b>	Sistema de consultas online para resolver dudas o atender consultas sobre la organización o sus productos o servicios
<b>Aplicaciones</b>	Programas informáticos diseñados como herramientas para permitir a un usuario realizar una actividad concreta.
<b>Bases de datos</b>	La información de la organización en cifras. Datos sobre la propia actividad de la organización, sobre el sector en el que se desarrolla o sus propias actividades de investigación. Fomenta la transparencia institucional.
<b>Blogs y microblogs</b>	Los blogs en la web son usados como utensilios de comunicación personal y organizacional. Pretenden aumentar el sentido de pertenencia y favorecer la imagen interna que se divulgue al exterior.  Microblogs una forma de interacción con mensajes o relatos breves.
<b>Blogs corporativos</b>	Celaya (2008) diferencia entre blogs personales, temáticos, profesionales y corporativos (internos y externos). En los corporativos externos distinguen los orientados a ventas y marketing con enfoque relacional; de atención al cliente para escuchar a los usuarios y atender sus propuestas y reclamos.
<b>Buzón de sugerencias y quejas</b>	Una forma de participación y/o atención al cliente o público de la organización o al usuario de la web.
<b>Chats</b>	Espacio de encuentro y conversación textual, visual o sonora, de forma síncrona entre usuarios de la web. Instrumento de comunicación interno, externo o mixto.
<b>Correo electrónico</b>	Forma más habitual de comunicación, son envíos consentidos o solicitados por el usuario al acceder sus datos a la organización. Permite ajustar el mensaje a las características de los públicos a los que se dirige.
<b>Encuestas</b>	Fomenta la participación, permite conocer las tendencias de opinión de los usuarios que visitan la web y quieren dejar su opinión.
<b>Enlaces</b>	Enlaces a otras web de la organización u otras organizaciones. Mejora el posicionamiento en buscadores y crea un entorno de prestigio avalado por la reputación de las entidades enlazadas.
<b>Foros</b>	Herramientas de participación y debate
<b>Fotoblog</b>	Espacio de creación personal basado en imágenes
<b>Geoubicación</b>	Localizadores espaciales de la organización para facilitar el acceso del usuario a la sede o al establecimiento.
<b>Juegos y entretenimiento</b>	El entretenimiento es una forma habitual de los internautas. Para generar más tráfico a su página utilizan juegos.
<b>Mensajes instantáneos</b>	Sistema de alerta y se usa para transmitir información relevante sobre la organización a sus públicos
<b>Podcast</b>	Archivos de sonido (breve) que el usuario puede descargar para escuchar cuando quiera en un reproductor portátil.
<b>Radio y Televisión</b>	Las grandes organizaciones crean medios propios de radio y televisión que simulan con descargas las emisiones de los medios de comunicación tradicionales
<b>Wikis</b>	Espacios abiertos para ser construidos con aportaciones colectivas de los usuarios de la página que libremente construyen y modifican el contenido.

**Elaborado por:** Rosario Puertas-Hidalgo, tomado de Tuñez M. (2012) Gestión de la comunicación (pp.: 163-166).

## **2. Web Semántica y periodismo de datos**

La web es un enorme repositorio de información, formado de un conjunto de fragmentos que de alguna manera están integrados e interrelacionados, asociados a dominios particulares y a las empresas. La información como se almacena no tiene un claro significado, que facilite su recuperación y manipulación, de manera automatizada o manual. La idea principal de la Web semántica es tratar esta deficiencia (Barceló, Guzmán & Sánchez. 2006).

La Web Semántica dota a los contenidos de mayor significado, es decir de más semántica y aporta un camino para razonar en la Web, extendiendo así sus capacidades. La idea es permitir que cualquier usuario pueda encontrar respuesta a sus preguntas de forma rápida y sencilla, gracias a una información ordenada y definida.

En la web actual cuando un usuario hace una búsqueda sobre cualquier tema, en Google por ejemplo, encuentra una gran cantidad de páginas que contienen las palabras buscadas. En este caso se trata de una búsqueda basada en la similitud de palabras o sintáctica.

La web semántica, por otra parte, trabaja no en la similitud de palabras, sino con los significados, trabaja de acuerdo al contexto, recomienda información, facilita el encontrar información precisa y útil. Incluso recomienda información de acuerdo a la ubicación geográfica, perfil de consumo almacenado, entre otras. Son búsquedas mucho más inteligentes y eficientes, es decir con un sentido semántico.

Una de las herramientas que se maneja en el marco de la Web Semántica son las bases de datos utilizadas para compartir información específica incluyendo no solo las definiciones de conceptos básicos en diversas áreas, sino que los relaciona entre ellos.

Al aplicar estas herramientas semánticas en el periodismo, da lugar al periodismo de datos (data journalism), en un intento de contar las historias a través de números, obtenidos al procesar gran cantidad de datos para obtener información más amplia y objetiva. Lo interesante es lograr que estos datos sean fácilmente leíbles y asimilables por el usuario que busca información.

Se entiende por periodismo de datos a un tipo de aproximación de los medios para adaptarse y responder a los cambios del entorno de información: más interactivo, multi-dimensional y de datos enlazados, desde los que se pueden contar historias (Gray, J.; Bounegru, L. & Chambers, L. 2013). Los lectores tienen la posibilidad de explorar las fuentes de datos que se exponen en las noticias, participar en el proceso de creación y evaluación de reportajes, etc.

Algunos medios han volcado sus actividades hacia el periodismo de datos (data journalism): El Guardian, New York Times, Texas Tribune y Die Zeit han impulsado este concepto y están contando historias a través de datos. En Latinoamérica, La Nación de Argentina trabaja en temas de periodismo on line y visualización de datos, sin embargo, aún son pocas las experiencias que se pueden detallar en esta parte del mundo.

En este sentido, la aplicación de la web semántica para el desarrollo del trabajo de investigación periodística, ha cobrado importancia en los últimos años, adaptándose a los nuevos modelos de generación y transmisión de información. Tanto los medios de comunicación como los periodistas se enfrentan a un lector cada vez más preparado que accede fácilmente a la información en la red, la cual le permite convertirse en un generador de contenidos.

El periodismo ha dado un giro total con la inclusión de las diferentes plataformas tecnológicas con la producción y difusión de contenidos.

En la gestión de la comunicación organizacional se habla del marketing de relaciones o marketing relacional, sistema basado en herramientas tecnológicas que permite construir la lealtad de y hacia el cliente (Dvoskin. 2004: 429). Actualmente surge un nuevo concepto: periodismo de datos, asociado a la gestión de la comunicación organizacional.

### **3. El Blog de Noticias de la Universidad Técnica Particular de Loja, UTPL**

Según el autor Miguel Túñez (2012: 139) comunicar no es sólo difundir mensajes, para que se cumpla con la difusión es fundamental que el receptor entienda y atienda a esos mensajes. Las organizaciones difunden información directamente al público en acciones interpersonales o medios propios y ajenos. Al hablar de la organización los mensajes deben no sólo atender asuntos que interesan a la organización, sino enfocarse en lo que los públicos consideran interesante o importante.

Considerando que al informar sobre si misma está sobreentendido que no difundirá asuntos desfavorables para su imagen, sin embargo, si quiere construirse como referencia de veracidad y proteger su reputación gestionará los contenidos con enfoques positivos de forma honesta y equilibrada, evitando discursos publicitarios de autosatisfacción. Los medios más utilizados para relacionarse con los públicos son los boletines y revistas de papel y electrónicas; diarios y blogs on line; redes de seguimiento (Twitter) y de encuentro (Facebook); televisión y radios internas o en internet (Túñez. 2012: 139,140).

Entre los medios utilizados por las empresas para comunicarse con sus públicos en un soporte digital, está el blog corporativo. El blog utiliza la tecnología entre el correo electrónico y un sitio web, es una página web estructurada con bloques de texto y entradas, es frecuentemente actualizado, tiene un orden cronológico inverso (Túñez. 2012: 164).

La UTPL cuenta con varios canales de comunicación con sus públicos internos y externos. Los principales son el blog de comunicación [www.utpl.edu.ec/comunicacion](http://www.utpl.edu.ec/comunicacion) [on line], informativo UTPL [papel – on line], revista científica Perspectivas desde la ciencia [papel – on line], noticiero 7 Días en la Técnica [on line – televisora local], correo electrónico (Postmaster), portal y sitios web UTPL, entre otros.

En este apartado nos enfocaremos al blog corporativo utilizado para difusión interna y externa, tiene un enfoque relacional, busca fomentar proximidad con

los públicos. Se publica información de los eventos más relevantes, principalmente congresos, seminarios, proyectos e investigaciones.

El blog se creó en el 2005 como una herramienta de difusión de información relacionada a eventos de la Universidad, con el paso del tiempo se fueron incorporando temas relacionados a la investigación, proyectos y vinculación con la colectividad. Actualmente el blog constituye uno de los más importantes repositorios de información cronológica de la vida universitaria, utilizada como fuente de información tanto por la comunidad de la institución como por medios de comunicación interesados en temas específicos.

El blog es administrado por la Unidad de Gestión de la Comunicación (UGC). Su actualización es permanente, utiliza textos, fotografías y videos principalmente e incluye opciones para compartir la información a través de redes sociales.

#### **4. Caso UTPL**

Como parte de la experiencia con la web semántica, desde la academia, se desarrolló una herramienta para facilitar la búsqueda de datos precisos y en menor tiempo: Serendipity<sup>1</sup>, una plataforma que se la concibe como un punto de información, para conocer las investigaciones que se realizan en el mundo sobre contenidos abiertos (Tovar 2012). Este proyecto es desarrollado por un grupo de investigadores de la Universidad Politécnica de Madrid y de la UTPL, en colaboración con el OpenCourseWare Consortium. Su meta es mejorar la habilidad de investigación y descubrimiento de contenidos educativos abiertos (Piedra et al, 2011).

Siguiendo la experiencia de Serendipity se planteó un proyecto que pretende investigar y desarrollar herramientas semánticas que permitan enlazar, organizar y transmitir la información adaptando la comunicación a la tendencia mundial. El aporte de los Departamentos de Ciencias de la Comunicación y Ciencias de la Comunicación y Electrónica y Telecomunicaciones servirán para definir procesos de publicación de datos enlazados desde el contenido

---

<sup>1</sup> Prototipos de Visualización de Datos de Recursos Educativos Abiertos <http://serendipity.utpl.edu.ec/map/>

generado por la Unidad de Gestión de la Comunicación lo que permite la transferencia de contenidos.

Esta constituye una experiencia totalmente innovadora en el Ecuador, al momento la UTPL es la única universidad del país trabajando de la mano las tecnologías de la Web Semántica y la Comunicación Organizacional.

#### **4.1 Objetivos y Componentes**

Se estableció como objetivo del caso la aplicación de tecnologías de la Web Semántica al contenido de la comunicación institucional de UTPL. Las fases desarrolladas fueron: definición del alcance, creación de ontología, extracción de datos, conversión a formatos semánticos, enlace, publicación en formatos legibles por máquinas, y piloto sobre funcionalidad que permita mejorar la comunicación organizacional en UTPL.

Los componentes desarrollados, son

- Base de conocimiento sobre entidades universitarias para apoyar procesos de extracción de datos y conversión a formatos legibles por máquinas desde las noticias de UTPL. El objetivo fue la selección y descripción semántica de aquellas entidades o conceptos relevantes y que forman parte del diccionario de entidades inicial y las reglas semánticas.
- Crawler para extraer información de noticias de la UTPL.
- Proceso de publicación de datos enlazados a partir del contenido extraído.
- Prototipo de visualización de información

#### **4.2. Base de conocimiento**

La base de conocimiento creada almacena datos, objetivos, requerimientos, reglas y especificaciones. Un paso fundamental es la definición de ontología que represente el contexto de las noticias de la UTPL. La base de conocimiento que nos permita definir un diccionario de términos que posteriormente será empleado en el proceso de extracción de entidades. La alimentación inicial de la base de conocimiento es manual, sin embargo, el

sistema es capaz de obtener nuevos conocimiento a través de la interacción de usuarios.

### 4.3. Extracción de Información

El Scrapy de datos consiste en extraer datos desde contenidos almacenados en páginas HTML, blogs, pdfs, docs y otros formatos (ver figura 1). Este proceso permitirá separar datos de los formatos de presentación y facilitará el procesamiento de estos datos (ver figura 2). Los datos extraídos son almacenados en un repositorio para su posterior procesamiento. Se creo adicionalmente un crawler para ejecutar periódicamente los procesos de extracción.

Figura 1. Resultados de Scrapy de una noticia

predicado	objeto
url	http://www.utpl.edu.ec/comunicacion/2013/05/mensaje-del-rector-canciller-por-el-aniversario-42-de-la-utpl/
titulo	Mensaje del Rector-Canciller por el aniversario 42 de la UTPL
fecha	mayo 06, 2013
autor	Claudia Rodríguez Hidalgo
contenido	Queremos compartir con ustedes todas nuestras ilusiones y todos nuestros sueños para los próximos años.
contenidohtml	<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-

**Fuente:** Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

El crawler fue desarrollado usando tecnología de licencia libre.

- Lenguaje de desarrollo. Python.
- Librerías de extracción de información: Scrapy & BS4 de Python, empleadas para la extracción de Información desde páginas HTML.
- Librerías para procesamiento de lenguaje natural: NLTK, utilizada para procesamiento de texto, tokenización, y clasificación de Información.
- Base de datos relacional: MySql
- Almacenamiento de tripletes: Virtuoso Open Source

Figura 2. Extracción de entidades simples y compuestas

Sujeto	Predicado	Objeto
lud:UTPL	rdf:type	lud:university
lud:UTPL	dc:label	Universidad Técnica Particular de Loja
lud:UTPL	dc:label	Universidad Técnica Particular de Loja
lud:UTPL	dc:label	Universidad Técnica Particular de Loja
lud:Eventos	dc:Description	Suceso importante dentro del ámbito de la UTP
lud:UTPL	dc:label	UTPL
lud:UTPL	dc:label	Universidad Técnica
lud:UTPL	dc:label	Universidad Técnica Particular de Loja
lud:Evento	rdf:type	lud:Evento
lud:Evento	rdf:type	lud:Eventos
lud:Evento	rdfs:label	evento
lud:Evento	rdfs:label	eventos
lud:Proyecto	dc:Description	Conjunto de esfuerzos para fortalecer el ento
lud:Proyecto	rdf:type	lud:Proyectos
lud:Proyecto	rdfs:label	Proyecto
lud:Proyecto	rdfs:label	Proyectos
lud:Proyecto	rdfs:label	proyecto
lud:Proyecto	rdfs:label	proyectos
lud:Estudiante	dc:Description	Personal estudiantil de la UTP
lud:Estudiante	rdf:type	lud:Estudiantes
lud:Estudiante	rdfs:label	Estudiante
lud:Estudiante	rdfs:label	estudiante
lud:Estudiante	rdfs:label	estudiantes
lud:Estudiante	rdfs:label	Estudiantes
lud:Investigación	dc:Description	Conjunto de eventos y proyectos que generan u
lud:Investigación	rdf:type	lud:Investigaciones
lud:Investigación	rdfs:label	Investigación
lud:Investigación	rdfs:label	investigación
lud:Investigación	rdfs:label	Investigacion
lud:Investigación	rdfs:label	investigacion
lud:Investigación	rdfs:label	Investigaciones
lud:Investigación	rdfs:label	investigaciones

Fuente: Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

El contenido de la comunicación organizacional es expresado en lenguaje humano. Para que sea procesado computacionalmente se usó técnicas de procesamiento de lenguaje natural (PNL). El PNL es clave en el reconocimiento y utilización de la información expresada en lenguaje humano, permite resolver ambigüedades semánticas y sintácticas, e inferir la importancia de la información contextual.

El procesamiento contempló

- Análisis morfológico: para detectar la relación que se establece entre las unidades mínimas que forman una palabra o token. Con esta se extraen palabras simples de una noticia, que representan un concepto, lugar, persona o una entidad en particular, como pueden ser: Loja, UTPL, universidad, rector, evento, congreso, biología, José, entre otros
- Análisis sintáctico: que permitió etiquetar cada uno de los componentes sintácticos que aparecen en el texto y analizar cómo las palabras se combinan para formar construcciones gramaticalmente correctas,

especialmente útil para identificar conceptos formados por varias palabras. Con esta se extraen cadenas de texto que representan un concepto en específico, ejemplo: “Universidad Técnica Particular de Loja”, “José Barbosa”, “Universidad Politécnica de Madrid”, “Congreso Iberoamericano de Medicina”, y muchos más

- Análisis semántico: para determinar aspectos del significado, sentido o interpretación de las palabras extraídas, y su posterior representación en formatos que permitan procesar computacionalmente significados. Se pueden determinar entidades de tipo persona, ciudades, países, universidades, tipos de eventos.
- Análisis pragmático: que integra los resultados de análisis anteriores (Ver figura 3).

Figura 3. Ubicación de entidades en el contenido de una noticia

Sujeto	Predicado	Objeto
▶ <a href="http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/">http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/</a>	mentioned_entities	1
1	mentioned_entity	lud:UTPL
1	entity_location	26
<a href="http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/">http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/</a>	mentioned_entities	2
2	mentioned_entity	lud:UTPL
2	entity_location	395
<a href="http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/">http://www.utpl.edu.ec/comunicacion/2013/06/compania-de-teatro-de-la-utpl-estreno-%e2%80%9cpuebla-de-las-mujeres%e2%80%9d/</a>	mentioned_entities	3
3	mentioned_entity	lud:UTPL
3	entity_location	416
<a href="http://www.utpl.edu.ec/comunicacion/2013/06/firma-de-convenio-entre-utpl-y-universidad-tecnica-de-machala/">http://www.utpl.edu.ec/comunicacion/2013/06/firma-de-convenio-entre-utpl-y-universidad-tecnica-de-machala/</a>	mentioned_entities	4
4	mentioned_entity	lud:UTPL
4	entity_location	26

Fuente: Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

#### 4.4. Publicación de Datos Enlazados

La Web Semántica, mediante sus tecnologías, permite expresar datos, conocimiento y su significado mediante lenguajes formales, de esta manera es posible que agentes de software puedan identificar la semántica subyacente en estas representaciones. Esto significa que, la capacidad de entendimiento podrá ser asumida por las máquinas, para interpretar los datos que reciben del entorno, determinar su significado y generar nuevos datos mediante reglas lógicas. Según la W3C, la Web Semántica es una Web de datos. Estos datos pueden venir en toda clase de formatos, lenguajes, estilos, estructuras. Este enfoque de la Web Semántica está alineado con la visión original que tuvo Berners-Lee de la Web a finales de 1980 en la que el significado de la

información tiene un rol clave y la información está almacenada dentro de una base de datos global, distribuida y de datos enlazados a través de la Web.

Publicar Linked Data es compartir datos estructurados que pueden ser conectados a otros datos a través de enlaces que tienen asociado un significado semántico, usando la arquitectura general de la World Wide Web.

Linked Data permite conectar y navegar por los datos enlazados, igual que sucede actualmente con las páginas Web. Linked Data permite enriquecer y ampliar el conocimiento de un tema, a través del enlace de los datos con fuentes de datos heterogéneas disponibles desde la Web.

En la figura 4, se muestran algunos de los datos que tiene una noticia en UTPL. Todos estos datos están embebidos en el texto HTML del cuerpo de la noticia. Los humanos pueden ver el contenido e interpretarlo semánticamente. Sin embargo, los agentes máquinas no son capaces de decir qué parte del contenido HTML hace referencia por ejemplo al departamento en el que trabajan los investigadores mencionados en la noticia, o identificar sobre qué lugares geográficos se mencionan o en qué proyecto forman parte.

Figura 4. Prototipo en el que se visualiza la extracción de entidades de una noticia

Anterior | Siguiente

**NOTICIA**  
**url**  
http://www.utpl.edu.ec/comunicacion/2013/05/3ra-jornada-nacional-de-la-juventud/  
**título**  
3ra Jornada Nacional de la Juventud  
**fecha**  
mayo 09, 2013  
**autor**  
Claudia Rodríguez Hidalgo  
**contenido**  
Del 3 al 5 de mayo se desarrolló, en la ciudad de Ambato, la 3ra Jornada Nacional de la Juventud. El objetivo fue objetivo de celebrar la presencia viva de los jóvenes ecuatorianos en el caminar de la Iglesia que peregrina en el Ecuador, desde una profunda reflexión sobre el Espíritu Santo que les da fortaleza para vivir, testimoniar el Evangelio y la audacia de proclamarlo haciéndoles capaces de ir madurando el seguimiento a Jesús con profundidad y gozo para que así los jóvenes iluminen al mundo con la Luz de Cristo. En este contexto y por medio de un diálogo de Mons. Germán Pavón, Obispo de Ambato, con José Barbosa, Rector de la UTPL, se acordó el apoyo de la universidad para el desarrollo de la 3ra Jornada de la Juventud, la cual se hizo en tres líneas de acción: En ese sentido se realizó un taller con 36 personas de distintas parroquias y cantones de la provincia de Tungurahua. Esta actividad estuvo coordinada desde la UTPL por Juan Figueroa Director del Centro Regional Quito, y contó con la participación de Juan Carlos Ribadeneira, de Misiones Universitarias; Adolfo Nieto, Coordinador Zonal de sierra centro; Jimmy Aguirre, Coordinador del Centro Ambato y Orfa Quezada, asistente operativo. Por la Diócesis de Ambato estuvo el P. Franklin Tospanta, Coordinador General del Encuentro, Patricia Villalba y Cecilia Naranjo por el Obispado.

**ENTIDADES**  
lud:UTPL -- 123  
lud:UTPL -- 179  
lud:Ambato -- 13  
lud:Ambato -- 114  
lud:Quito -- 187  
lud:Ambato -- 218  
lud:Ambato -- 229

**PERSONAJES**  
Jose Barbosa  
Juan Figueroa  
Adolfo Nieto  
Jimmy Aguirre

**Fuente:** Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

Para mejorar la capacidad de procesamiento de los agentes máquinas, utilizó un enfoque de Linked Data. Linked Data es un conjunto de principios para exponer datos en la Web (no solamente páginas Web que contienen información) de manera que estos sean directamente enlazables, interoperables, visibles y reusables Berners-Lee. T. (2006). Esto supone una evolución de la Web hacia un espacio global de información en el que la navegación se realiza a través de datos estructurados enlazados en vez de realizarse a través de documentos, como sucede ahora (Heath y Bizer, 2011).

El uso de Linked Data permite pasar de una Web basada en documentos HTML, en la que el usuario es el destinatario de la información publicada, a una Web de Datos Enlazados que están expresados en RDF un lenguaje para representar información sobre recursos, en la que agentes software pueden explotar estos datos de forma automática (recopilándolos, agregándolos, interpretándolos, publicándolos, etc.), potenciados por vocabularios (Gruber, 1993:1995) que usan especificaciones explícitas y formales de una conceptualización compartida en este caso globalmente. Implementar la Web de datos se basa en la idea de seguir cuatro principios de diseño propuestos por Tim Berners-Lee.

Para la aplicación de los principios de Linked Data se estableció un marco de trabajo (Piedra et al, 2014) que consta de cinco fases definidas:

- i. Identificación y selección de fuentes de contenidos. Entre las fuentes de información a evaluar, se seleccionó contenido de noticias en páginas web
- ii. Modelamiento de ontología / vocabularios;
- iii. Generación de datos en RDF,
- iv. Publicación y enlace de datos enlazados,
- v. Consumo y visualización de datos enlazados. Ver figura 3.

#### **4.5. Prototipo de Visualización**

La infografía es uno de los medios para transmitir información gráficamente. Como prototipo del proceso descrito anteriormente se elaboró un infografía para conseguir una representación visual de los datos extraídos y

descritos usando tecnologías de Linked Data. A continuación se presenta infogramas para mostrar información de síntesis basados en cálculos matemáticos sobre los datos extraídos (ver Figura 5.a y 5.b).

Figura 5.a **Infogramas calculados a partir de datos publicados en RDF desde Noticias UTPL**



**Fuente:** Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

En el contexto ecuatoriano son pocos los periodistas que han desarrollado o han escrito experiencias referentes al uso de técnicas de trabajo de datos en su ejercicio profesional. Lo que se ha expuesto por expertos en periodismo de investigación son las tácticas para realizar búsquedas avanzadas de recolección, destilamiento, filtración, y relación de datos.

Existen algunos medios pioneros en el tema de periodismo de datos. El Guardian, New York Times<sup>2</sup>, Chicago Tribune<sup>3</sup>, Texas Tribune y Die Zeit han impulsado el concepto de periodismo de datos (data journalism) y están contando historias a través de datos. En Latinoamérica, La Nación<sup>4</sup> de Argentina trabaja en temas de periodismo on line y visualización de datos. El diario londinense The Guardian lleva adelantes su iniciativa de periodismo de datos a través de su página Web Datablog. En sus cuatro años de vida, Datablog ha publicado numerosos reportajes escritos, que en la mayoría de los casos se han complementado con mapas y otras herramientas interactivas. Un ejemplo representativo de ellos apareció el 7 de agosto de 2012, con el título

2 El blog de The New York Times sobre periodismo de datos es Open Blog (<http://open.blogs.nytimes.com/>)

3 El espacio dedicado al periodismo de datos de Chicago Tribune se llama Maps & Apps (<http://www.chicagotribune.com/news/data/>)

4 El citado blog de La Nación se llama Data y esta es su dirección web: <http://blogs.lanacion.com.ar/data/>

“Could you be an athlete? Olympics 2012 by age, weight and height”<sup>5</sup> y elaborado por Rogers y Nicola Hughes, informática de The Guardian.

Figura 5.b Infogramas calculados a partir de datos publicados en RDF desde Noticias UTPL



Fuente: Proyecto Tecnologías de la web semántica aplicadas a la comunicación institucional UTPL

## 5. Conclusiones

El periodismo de datos tiene el potencial de adaptarse y responder a los cambios del entorno de información, que se caracteriza por ser más interactivo, multidimensional, de datos enlazados, en el que se usa información cuantificada, así como un entorno que adopta herramientas tecnológicas desde las que se pueden contar historias. Por otra parte, los lectores tienen la posibilidad de explorar las fuentes de datos que se exponen en las noticias, participar en el proceso de creación y evaluación de reportajes.

En el contexto de este artículo, la Web Semántica puede aportar a la comunicación organización, debido a que el contenido es descrito y almacenado en formatos procesables e interoperables por máquinas, de manera que los usuarios puedan encontrar recursos, acceder a servicios, recibir recomendaciones según sus necesidades de acceso, mejorar la participación, colaborar a través de la Web, y en definitiva entender los datos así como el contexto para ajustar el contenido, la estructura y presentación de los recursos según las necesidades específicas de los usuarios.

<sup>5</sup> Fuente: <http://www.guardian.co.uk/sport/datablog/2012/aug/07/olympics-2012-athletes-age-weight-height>

Lo semántico implica trabajar en significados de las cosas, y dar respuestas de acuerdo al contexto, recomendar información hacer que el usuario encuentre información más precisa y útil y adicionalmente recomendar información de acuerdo a nuestra ubicación geográfica, perfil de consumo y otros datos que van quedando registrados; hacer las búsquedas más inteligentes y eficientes.

## 6. Reconocimiento

El desarrollo de la investigación ha sido posible gracias al financiamiento de UTPL a través de la convocatoria interna de proyectos de investigación, 2013. Los autores desean agradecer el apoyo de los equipos técnicos de la Unidad de Gestión de la Comunicación y la Sección Departamental de Tecnologías Avanzadas de la Web y Sistemas Basados en el Conocimiento de UTPL.

## 7. Referencias bibliográficas

- Barceló, M.; Guzmán, G.; Sánchez A. (2006). **La Web Semántica como apoyo a la Gestión del Conocimiento y al Modelado Organizacional**. Revista Ingeniería Informática. Edición 12. [en línea]. Disponible en: <http://www.inf.udec.cl/~revista/ediciones/edicion12/articulo%2012-4.pdf> [consultado 5/11/2013].
- Berners-Lee, T., Hendler J., and Lassila O. (2001). Semantic web, a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, 284(5), 34-42, Online on: <http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2>.
- Berners-Lee. T. (2006) Linked Data. Design Issues, W3C, <http://www.w3.org/DesignIssues/LinkedData.html>
- Castells, M. (2009). **Creatividad, innovación y cultura digital. Un mapa de sus interacciones**. Revista Telos, 77, 5 de octubre de 2009. [en línea]. Disponible en: <http://sociedadinformacion.fundacion.telefonica.com/telos>. [consultado 9/11/2013].

- Dvoskin, R. (2004). **Fundamentos de Marketing: teoría y experiencia**. Ediciones Granica. Chile.
- Gray, J.; Bounegru, L. & Chambers, L. (2013) Data Journalism Handbook. [en línea]. Disponible en: <http://datajournalismhandbook.org/1.0/en/> [consultado 11/11/2013].
- Gruber, T. (1993) A translation approach to portable ontologies. Knowledge Acquisition, pp. 5(2): 199–220.
- Gruber, T. (1995) Towards principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies, pp. 43 (5/6): 907 – 928
- Heath T. and Bizer C. (2011) Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool.
- Piedra, N. Tovar, E. López, J. Chicaiza, J. (2014) "Consuming and producing linked open data: The case of Opencourseware", Emerald EarlyCite, in press.
- Piedra, N., Chicaiza, J., López, J., Martínez, O., Tovar, E., (2011) "Finding OERs with Social-Semantic Search". Education Engineering (EDUCON) IEEE, Amman, Jordan, 4-6 April 2011. [http://conftool.com/educon/index.php?page=browseSessions&form\\_session=32](http://conftool.com/educon/index.php?page=browseSessions&form_session=32)
- Tironi, E. & Cavallo, A. (2004). **Comunicación estratégica. Vivir en un mundo de señales**. Aguilar, Altea, Taurus, Alfaguara S.A. de Ediciones. Chile.
- Tovar, E. Piedra, N. Chicaiza, J. Lopez, J. Martinez, O. (2012) Development and promotion of OERs. Outcomes of an international research project under OpenCourseWare model, Journal of Universal Computer Science.
- Tuñez, J. M. (2012). **La gestión de la comunicación en las organizaciones**. Comunicación Social Ediciones y Publicaciones. Zamora – España.