

GENXMLDC: Software para mostrar el uso de Tecnologías de la Web Semántica

GENXMLDC: Software to show the use of Semantic Web Technologies

Gustavo Rodríguez Bárcenas¹, Alex Santiago Cevallos Culqui², Segundo Humberto Corrales Beltrán³, Jorge Bladimir Rubio Peñaherrera⁴, Fausto Alberto Viscaino Naranjo⁵

RESUMEN:

En el presente trabajo se muestra una pequeña base teórica conceptual de la Web Semántica y las tecnologías que la componen, se hace énfasis en dos tecnologías que en la actualidad tienen gran difusión por las posibilidades que estas brindan en la descripción de documentos electrónicos como son el XML y el *Dublin Core*, así mismo se muestra un software desarrollado para generar documentos con este tipo de estructura de datos y etiquetado, permitiendo su posterior publicación, demostrando una de las muchas aplicaciones que tienen estos tipos de tecnologías.

Palabras claves: Dublin Core, metadatos, XML, software, SCRUM.

ABSTRACT:

In the present work a small conceptual theoretical base of the Semantic Web and the technologies that compose it are shown, it is emphasized in the technologies that have at the moment they have a great diffusion by the possibilities that these offer in the description of the electronic documents as are the XML and the Dublin Core, likewise it shows a software developed to generate documents with this type of data structure and labeled, allowing its subsequent publication, demonstrating one of the many applications that these types of technologies have.

Keywords: Dublin Core, metadata, XML, software, SCRUM.

Recibido 22 de febrero del 2017; revisión aceptada 22 de abril 2017

¹ Universidad Técnica de Cotopaxi, Latacunga, Cotopaxi, Ecuador, gustavo.rodriguez@utc.edu.ec

² Universidad Técnica de Cotopaxi, Latacunga, Cotopaxi, Ecuador, alex.cevallos@utc.edu.ec

³ Universidad Técnica de Cotopaxi, Latacunga, Cotopaxi, Ecuador, segundo.corrales@utc.edu.ec

⁴ Universidad Técnica de Cotopaxi, Latacunga, Cotopaxi, Ecuador, Jorge.rubio@utc.edu.ec

⁵ Universidad Regional Autónoma de Los Andes, Ambato, Ecuador, faustov_zh1@hotmail.com

1. INTRODUCCIÓN

La actual era de la información está caracterizada por un extraordinario crecimiento del volumen de datos, generados y almacenados generalmente en soportes electrónicos, sobre una gran variedad de disciplinas. Una creciente proporción de estos datos es almacenada en forma de bases de datos informáticas, para facilitar su acceso. La disponibilidad de tal cantidad de datos ha generado una demanda de nuevas y potentes herramientas para la extracción de información y conocimiento útil, orientado hacia una determinada aplicación.

Las tecnologías que también puntualizan esta era, así como sus propias técnicas para el tratamiento, representación y procesamiento de la información, las fronteras entre los modelos de información, se muestran difusos e indeterminados en muchas organizaciones. El modelo que hoy en día se potencian podría ser definido como un *metamodelo de usuario*. Es un *modelo de usuario* en la medida que representa una reducción de la base informativa, atendiendo a las consultas, necesidades o perfiles propios del usuario, por medio de procesos informáticos complejos [1, 2].

Las nuevas tecnologías propician cambios en la relación del individuo con el tiempo y con el espacio. Los tiempos tradicionales en el ciclo vital de las personas, dedicados al aprendizaje, al trabajo y al ocio sufren un proceso de transformación radical asociada a la desterritorialización de estas actividades.

De esta manera, la Visualización de Información y el Análisis Exploratorio de Datos están cobrando cada vez más importancia en el mundo empresarial y académico - científico. Su objetivo práctico consiste en encontrar nuevo conocimiento en bases de datos de dimensión, tamaño o complejidad demasiado elevados para un análisis manual. Se trata de procesos interactivos que requieren la intuición y el conocimiento de expertos, junto con la eficiencia computacional de la tecnología informática moderna. [3]

Cabe destacar la filosofía de trabajo, en la que una imagen vale más que mil palabras, y donde las técnicas deben generar imágenes cuyo valor más importante sea forzarnos a percibir lo que nunca esperábamos ver. La posibilidad de encontrar información oculta, en bases de datos de tamaño y dimensión elevada, pueden aportar una ventaja clave para el éxito de un proceso empresarial o de un estudio de investigación, en cualquier disciplina (medicina, ingenierías, investigación social, negocios, etc.). [3]

El manejo de datos e informaciones constituye uno de los aspectos más importantes para cualquier organización contemporánea. El manejo de información abarca diferentes actividades como la recolección, almacenamiento, recuperación y difusión [4] constituyendo esto elementos básicos en la actividad científica de cualquier organización.

Esta humanidad ha visto acumularse un caudal incalculable de conocimiento en el discursar de los tiempos. Su permanente transmisión, entre diversas generaciones y dentro de cada una de ellas, ha contribuido a acelerar en forma exponencial el desarrollo científico, tecnológico y el propio desarrollo de la humanidad.

La información se encuentra asociada invariablemente a toda obra humana, de hecho la mayoría de las actividades del ser humano está mediada por la información, devenido del propio proceso de interpretación de los distintos datos que nos rodean cotidianamente reproducidos en diversos soportes para su propia conservación y transmisión en el tiempo y en el espacio.

Desde tiempos muy remotos el hombre ha tenido la necesidad de comunicarse, y para esto ha transitado por varias etapas de la vida en que ha dejado representado en las líneas de la historia, innumerables concepciones que identifican claramente el proceso evolutivo de la información.

La información de hecho, juega un rol fundamental en la propia adquisición de conocimientos, pues la analogía de lo tácito a lo explícito o de lo explícito a lo tácito los relacionan de una manera muy peculiar, formando una fusión conceptual de importante acontecimiento extendido en todos los tiempos, desde la comunidad primitiva hasta el hombre moderno que facilita su actividad humana a través de nuevas ciencias y tecnologías, permitiendo así un desarrollo vertiginoso, todo esto es permisible por la capacidad de entender los fundamentos teóricos y conceptuales que se encuentran en la actividad cognitiva del ser humano en el proceso de transformación de los datos que lo rodean en información y luego en conocimiento.

Tradicionalmente, la transferencia de información – conocimiento entre generador y usuario está estrechamente ligada al uso de fuentes y canales formales e informales; son modalidades interdependientes y complementarias, vinculadas con la estructura y organización social de la Ciencia y la Tecnología [5].

La irrupción de las Tecnologías de la Información y las Comunicaciones (TIC) en la vida del hombre y en su campo social, político, económico y cultural ha venido promoviendo cambios en la forma de realizar las tareas. Esto ha generado un conjunto de necesidades, sobre todo en el ámbito educativo y especialmente en la formación de las personas que conformarán la sociedad de los próximos siglos.

Está claro que las Tecnologías de la Información y las Comunicaciones forma parte ya de la propia actividad cotidiana del ser humano, en todas las áreas de conocimiento se ve reflejado este importante elemento tecnológico, a partir del propio surgimiento de la gran Red de

información (Internet) en su primera concepción se ha visto reflejado el proceso investigativo de las ciencias y las tecnologías con el uso de esta eficaz herramienta, está claro que hoy en día en Internet se comparte y se difunde un enorme cúmulo de información, muchas de ellas con un contenido bastante ruidoso y de muy mala calidad, donde los usuarios que dedican sus esfuerzos a la innovación e investigación sufre de la bien llamada “infoxicación” que se genera devenido de la propia libertad con que cualquier persona cuelga información en la Gran Red de Redes, es por tanto que aparecen nuevas necesidades de mostrar la información y que esta supla al menor costo de esfuerzo posible, en este sentido juegan un importante papel las distintas tecnologías en que están soportados los contenidos de la Web. El uso de Internet como herramienta educativa y de investigación científica ha crecido aceleradamente debido a la ventaja que representa el poder acceder a grandes bases de datos, la capacidad de compartir información entre colegas y facilitar la coordinación de grupos de trabajo.

La Internet de hoy en día ya no es una red académica, como en sus comienzos, sino que se ha convertido en una red que involucra, en gran parte, intereses comerciales y particulares. Esto la hace inapropiada para la experimentación y el estudio de nuevas herramientas en gran escala.

Adicionalmente, los proveedores de servicios sobre Internet "sobrevenden" el ancho de banda que disponen, haciendo imposible garantizar un servicio mínimo en horas pico de uso de la red. Esto es crítico cuando se piensa en aplicaciones que necesiten calidad de servicio garantizada, ya que los protocolos utilizados en la Internet actual no permiten eficientemente esta funcionalidad.

Esta nueva etapa va aparejada de la aparición de concepciones en cuanto al tratamiento de la información y su forma de mostrar, así como su almacenamiento, en este sentido podemos hacer referencia a los preceptos de la Web Semántica y las distintas tecnologías que la componen, dos de los ejemplos más conocidos lo destaca [7] son el servicio *Really Simple Syndication* (RSS), el cual es un vocabulario RDF (*Resource Description Framework*) basado en XML (*eXtensible Markup Language*) que permite realizar una catalogación de información, noticias, datos, eventos, etc., permitiendo encontrar información precisa adaptada a las preferencias de los usuarios dentro de una organización [6-10].

XML es un lenguaje de etiquetado que juega un papel fundamental en el intercambio de una gran variedad de datos, su desarrollo comienza en septiembre de 1996 dirigido por el W3C y con la participación de importantes empresas: Microsoft, IBM, Sun, Novell, ArborText, HP donde su propósito fue:

- diseñar un lenguaje de marcas optimizado para el WWW

- unir la simplicidad de HTML con la capacidad expresiva de SGML

Existen varios sistemas de Metadatos, entre ellos uno de los más populares es el *Dublin Core*, este nace como producto del trabajo cooperativo de ámbito internacional promovido por la OCLC (*Online Computer Library Center*) y el NCSA (*National Center for Supercomputing Applications*) cuyo objetivo principal fue la descripción de recursos electrónicos con el fin de facilitar su búsqueda y recuperación, así como apoyar la organización de la información en las organizaciones.

El propósito de *Dublin Core* es establecer un sistema normalizado para la descripción de documentos distribuidos en el WWW, cualquiera que sea su formato: HTML, PDF, PS, SGML, etc., facilitar a los autores y editores de documentos de una organización, la incorporación de elementos que identifiquen y describan sus aportaciones, corregir las deficiencias de los sistemas de recuperación basados en texto íntegro y evitar el uso de formatos complejos (MARC, TEI, etc.), que faculten la integración con sistemas de clasificación, indización y control de autoridades.

Los factores que han generado el éxito de las organizaciones en Internet, también han originado sus principales problemas: sobrecarga de información, heterogeneidad de fuentes y problemas consiguientes de interoperabilidad. La Web Semántica ayuda a resolver estos problemas, al permitir a los usuarios delegar tareas en herramientas de software [11-14].

El presente trabajo pretende mostrar algunos elementos teóricos conceptuales relacionados con las Tecnologías de la Web Semántica (TWC), así como usar sus potencialidades en el desarrollo de un software que permita entrar una serie de datos relacionados con los sistemas de difusión de información científica como son las revistas científicas, a través de una estructura de campos relacionados con el Autor, Título, entre otros, permitiendo su publicación en bases de datos o sitios web destinados a esta actividad. El uso de este tipo de tecnología permitirá una mejor descripción de estos recursos en cuanto a características, contenidos informativos, gestión de derechos y accesibilidad.

2. METODOLOGÍA

Para el desarrollo de la investigación se siguieron una serie de etapas y se tuvieron en consideración un conjunto de aspectos que se describen a continuación:

a) Unidad de Análisis

Como unidad de análisis se toman específicamente los metadatos *Dublin Core*, y que permiten establecer un sistema normalizado para la descripción de documentos distribuidos en el WWW, cualquiera que sea su formato: HTML, PDF, PS, SGML, etc., facilitar a los autores y editores de documentos la incorporación de elementos que identifiquen y describan

sus aportaciones, corregir las deficiencias de los sistemas de recuperación basados en texto íntegro y evitar el uso de formatos complejos (MARC, TEI, entre otros), que faculten la integración con sistemas de clasificación, indización y control de autoridades, con los que la comunidad de las ciencias de la información se encuentran familiarizadas, referidos a las características intrínsecas de los documentos, que puedan implementarse utilizando distintos sistemas: HTML, SGML, que sean opcionales y además repetibles tantas veces como sea necesario.

b) Métodos

Se realizó una investigación partiendo de los siguientes principios:

- Descriptiva. Analiza y caracteriza el fenómeno de los metadatos para el desarrollo de un software orientado a demostrar la importancia del uso de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir recursos y potenciar sistemas más inteligentes en el descubrimiento de recursos.
- Revisión bibliográfica: se llevó a cabo a partir de la búsqueda de literatura científica consultando importantes fuentes de información primarias, tales como bases de datos de alto reconocimiento e impacto académico como SCOPUS, ScienceDirect, EBSCO, entre otras, a partir de la combinación de campos de búsqueda con terminologías relacionadas con los metadatos, específicamente *dublin core*, elementos relacionados principalmente con el contenido del recurso, cuando es visto como una propiedad intelectual, con la instanciación del recurso, entre otros.

Para el desarrollo del software que utiliza el estándar *Dublin Core* es usada la *Metodología SCRUM*:

Scrum es un marco de referencia para crear software complejo y entregarlo a tiempo de una forma mucho más sencilla, el cual lo habilitará para crear excelente Software mediante la aplicación de un conjunto de directrices a seguir por los equipos de trabajo y el uso de roles concretos [15, 16].

Los Roles de SCRUM son los siguientes:

- ✓ Product Owner (dueño del producto): Es el que está a cargo del proyecto y cerciora de que el equipo Scrum trabaje de forma adecuada.
- ✓ Development Team (el equipo de desarrollo): Es un equipo multidisciplinario y auto organizado, su cometido principal es construir el producto que el Product Owner especifica.

- ✓ Scrum Master: Es el responsable de que los miembros del equipo sigan el proceso como es debido y de remover los obstáculos que impiden que el equipo alcance el objetivo del sprint.

La metodología Scrum cuenta con diversa fases que son:

- ✓ Planificación del Sistema: la fase de planificación es un proceso continuo donde se establecen los objetivos, las estrategias, las actividades y los recursos necesarios para el logro de los objetivos establecidos. Antes de iniciar el proyecto de desarrollo de software es necesario realizar una serie de actividades previas como la recolección de requerimientos para que el equipo de trabajo pueda entender el contexto de negocio del software, las mismas que influirán en la finalización exitosa del proyecto.
- ✓ Análisis de Requerimientos: en esta fase se identifican a partir de técnicas de recolección de datos y métodos empíricos las necesidades del cliente o del usuario, para a partir de estos requisitos iniciar todo el proceso de desarrollo del sistema.
- ✓ Diseño: en esta fase es donde se realiza la estructura de la propuesta estableciendo los rangos definitivos de la misma, se determina los recursos necesarios para llevar a cabo este trabajo de investigación.
- ✓ Construcción y Pruebas: se desarrolla el sistema y se le realizan pruebas para comprobar el nivel de aceptación y los posibles fallos encontrados darles solución.
- ✓ Implementación: es la instalación y construcción del sistema con el uso de lenguajes de programación, gestores de bases de datos y otras herramientas necesarias.

3. RESULTADOS

Metadatos *Dublin Core*:

Los metadatos *Dublin Core* tratan de ubicar, dentro de Internet, los datos necesarios para describir, identificar, procesar, encontrar y recuperar un documento introducido en la red. Si este conjunto de elementos *Dublin Core* se lograra aceptar internacionalmente supondría que todos los procesos que indizan documentos en Internet encontrarían, en la cabecera de los mismos, todos los datos necesarios para su indización y además estos datos serían uniformes. Para la implementación a través de un software se hace necesario conocer que el sistema de *Dublin Core* se encuentra estructurado de la siguiente forma:

I. *Los Contenido:*

- a) Título: el nombre dado a un recurso, habitualmente por el autor.
 - Etiqueta: DC.Title

- b) Claves: los tópicos del recurso. Típicamente, Subject expresará las claves o frases que describen el título o el contenido del recurso. Se fomentará el uso de vocabularios controlados y de sistemas de clasificación formales.
- Etiqueta: DC.Subject
- c) Descripción: una descripción textual del recurso. Puede ser un resumen en el caso de un documento o una descripción del contenido en el caso de un documento visual.
- Etiqueta: DC.Description
- d) Fuente: secuencia de caracteres usados para identificar unívocamente un trabajo a partir del cual proviene el recurso actual.
- Etiqueta: DC.Source
- e) Idioma: Idioma/s del contenido intelectual del recurso.
- Etiqueta: DC.Language
- f) Relación: es un identificador de un segundo recurso y su relación con el recurso actual. Este elemento permite enlazar los recursos relacionados y las descripciones de los recursos.
- Etiqueta: DC.Relation
- g) Cobertura: es la característica de cobertura espacial y/o temporal del contenido intelectual del recurso. La cobertura espacial se refiere a una región física, utilizando por ejemplo coordenadas. La cobertura temporal se refiere al contenido del recurso, no a cuándo fue creado (que ya lo encontramos en el elemento Date).
- Etiqueta: DC.Coverage

II. *Propiedad Intelectual:*

- a) Autor o Creador: la persona o organización responsable de la creación del contenido intelectual del recurso. Por ejemplo, los autores en el caso de documentos escritos; artistas, fotógrafos e ilustradores en el caso de recursos visuales.
- Etiqueta: DC.Creator
- b) Editor: la entidad responsable de hacer que el recurso se encuentre disponible en la red en su formato actual.
- Etiqueta: DC.Publisher
- c) Otros Colaboradores: una persona u organización que haya tenido una contribución intelectual significativa, pero que esta sea secundaria en comparación con las de las personas u organizaciones especificadas en el elemento Creator. (por ejemplo: editor, ilustrador y traductor).

- Etiqueta: DC.Contributor
- d) Derechos: son una referencia (por ejemplo, una URL) para una nota sobre derechos de autor, para un servicio de gestión de derechos o para un servicio que dará información sobre términos y condiciones de acceso a un recurso.

- Etiqueta: DC.Rights

III. *Instanciación:*

- a) Fecha: una fecha en la cual el recurso se puso a disposición del usuario en su forma actual. Esta fecha no se tiene que confundir con la que pertenece al elemento Coverage, que estaría asociada con el recurso en la medida que el contenido intelectual está de alguna manera relacionado con aquella fecha.

- Etiqueta: DC.Date

- b) Tipo del Recurso: la categoría del recurso. Por ejemplo, página personal, romance, poema, diccionario, etc.

- Etiqueta: DC.Type

- c) Formato: es el formato de datos de un recurso, usado para identificar el software y, posiblemente, el hardware que se necesitaría para mostrar el recurso.

- Etiqueta: DC.Format

- d) Identificador del Recurso: secuencia de caracteres utilizados para identificar unívocamente un recurso. Ejemplos para recursos en línea pueden ser URLs y URNs. Para otros recursos pueden ser usados otros formatos de identificadores, como por ejemplo ISBN ("International Standard Book Number").

- Etiqueta: DC.Identifier

Resultados del desarrollo del Software (GenXMLDC):

Tomando como base las concepciones planteadas anteriormente en este trabajo relacionado con las dos tecnologías semánticas referidas se desarrolla un software que tiene como objetivo fundamental la generación de estructuras de datos en XML y Dublin Core y además la posibilidad de publicarlo en directorios locales de PC o sitios Web, así como en bases de datos, el software es desarrollado con Visual Basic, fueron seleccionados la mayoría de los campos que pueden ser usados para describir documentos electrónicos como artículos de revistas y otros, estos campos son:

- | | |
|----------------|------------------|
| 1. Título | 8. Colaboradores |
| 2. Autor | 9. Derechos |
| 3. Claves | 10. Fecha |
| 4. Descripción | 11. Recurso |

- | | |
|-----------|-------------------|
| 5. Fuente | 12. Formato |
| 6. Idioma | 13. Identificador |
| 7. Editor | |

Estos son los campos más frecuentes para describir documentos electrónicos, por lo cual se decidió usarse en el software Generador de XML y Metadatos en *Dublin Core*.



Figura 1. Ventana principal del software

En la figura 1 se muestra la ventana principal donde está compuesta por dos menús, el menú Principal y el menú Ayuda, donde se observan además opciones relacionadas con la Entrada de Datos, generar Metadatos (Dublin Core), Generar XML, Exportar y Salir, así mismo en el menú Ayuda aparecen las opciones de Contenido y Acerca de..., en la opción Entrada de datos aparece otra ventana como se muestra en la figura 2, con los distintos campos a los cuales se hizo referencia anteriormente.

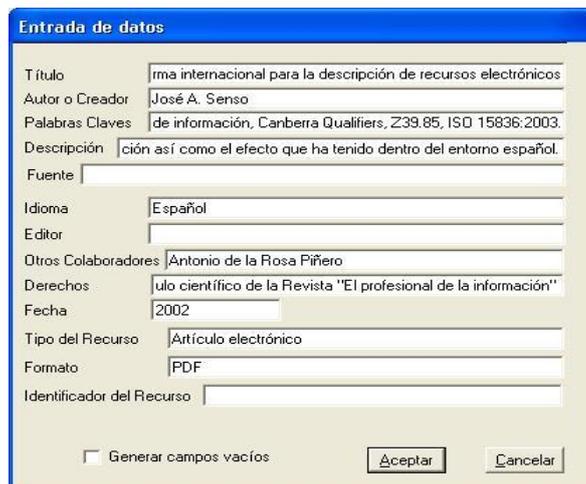


Figura 2. Ventana de entrada de datos con los distintos campos que describe el *Dublin Core*.

En la figura 2 cuando ejecutamos la acción de aceptar se generan las estructuras de datos del XML y Dublin Core, tomando para ello los contenidos de cada uno de los campos, el código que responde a esta acción en el lenguaje de programación Visual Basic es el que se muestra a continuación:

```

Private Sub cmdAceptar_Click()
    Dim x As Control
    Instanciacion = ""
    Body = ""
    XML = "<?xml version=" + Chr(34) + "1.0" + Chr(34) + " encoding=" + Chr(34) + "ISO-8859-1" + Chr(34) +
    "?>" + vbNewLine
    XML = XML + "<?xml-stylesheet type=" + Chr(34) + "text/css" + Chr(34) + " href=" + Chr(34) + "estilo.css"
    + Chr(34) + "?>" + vbNewLine
    XML = XML + vbNewLine + "<DOCUMENTO>" + vbNewLine
    For i = 0 To 12
        If (Check1.Value = 0 And Trim(txtTitulo(i).Text) <> "") Or Check1.Value = 1 Then
            Instanciacion = Instanciacion + P1 + txtTitulo(i).Tag + P2 + txtTitulo(i).Text + P3 + vbNewLine
            Body = Body + "<P>" + lbTitulo(i).Caption + ": " + txtTitulo(i).Text + "</P>" + vbNewLine
            XML = XML + " <" + List1.List(i) + ">" + txtTitulo(i).Text + "</" + List1.List(i) + ">" + vbNewLine
        End If
    Next
    XML = XML + "</DOCUMENTO>"
    SDIMain.mnGenerate.Enabled = True
    SDIMain.mnuXML.Enabled = True
    Unload Me
End Sub
  
```

Donde P1, P2 y P3 fueron declaradas variables públicas en un módulo como se muestra a continuación:

```

Public Sub Main()
    P1 = "<meta name=" + Chr(34) + "DC."
    P2 = Chr(34) + " content=" + Chr(34)
    P3 = Chr(34) + " />"
    SDIMain.Show
End Sub
  
```

En las figuras 3 y 4 se muestran las acciones de generar los códigos correspondientes con esta opción, como se puede observar queda estructurada la información a partir de cada uno de los campos que la componen, como se muestra a continuación:

```

<meta name="DC.Title" content="Dublin Core Metadata Initiative. Norma internacional para la descripción de
recursos electrónicos" />
<meta name="DC.Creator" content="José A. Senso" />
<meta name="DC.Subject" content="Dublin Core, DC, Dublin Core Metadata Initiative, DCMI, Metadatos,
Recuperación de información, Canberra Qualifiers, Z39.85, ISO 15836:2003." />
<meta name="DC.Description" content="El presente trabajo trata de trazar cronológicamente la evolución de
uno de los sistemas de metadatos que más impacto y repercusión ha tenido dentro del ámbito bibliotecario: el
Dublin Core Metadata Initiative. Al mismo tiempo se describen algunos de los principales problemas que se
encuentra este sistema para lograr su implantación así como el efecto que ha tenido dentro del entorno
español." />
<meta name="DC.Language" content="Español" />
<meta name="DC.Contributor" content="Antonio de la Rosa Piñero" />
<meta name="DC.Rights" content="Es un artículo científico de la Revista "El profesional de la información" />
<meta name="DC.Date" content="2002" />
<meta name="DC.Type" content="Artículo electrónico" />
<meta name="DC.Format" content="PDF" />
  
```

Cada contenido de los campos se encuentra reflejados en cada metadato como hemos podido ver.

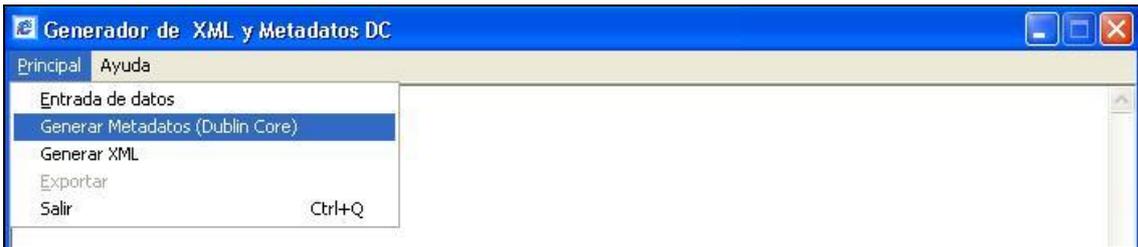


Figura 3. Ventana que muestra las opciones de Generar Metadatos (*Dublin Core*) activas por la acción de entrada de datos.

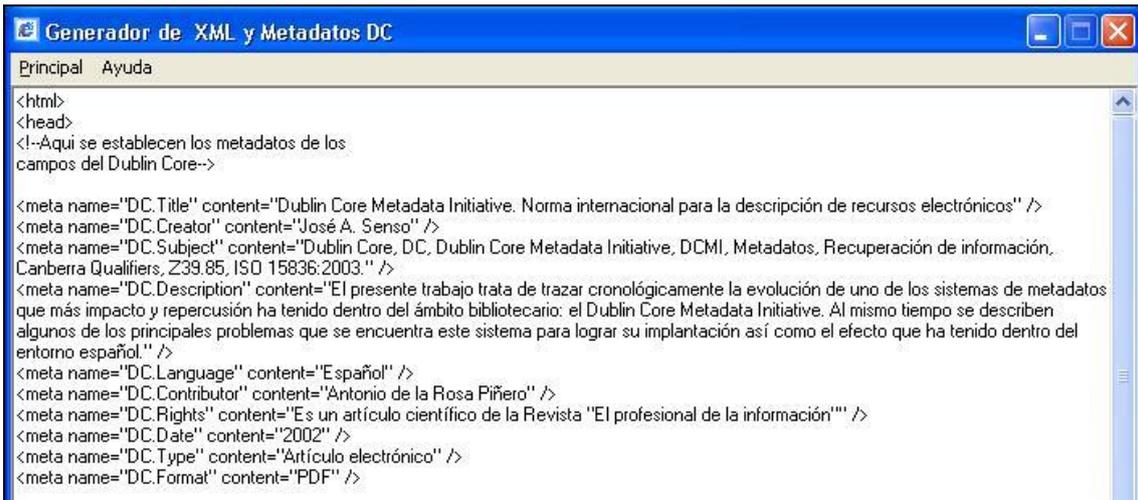


Figura 4. Ventana que muestra la estructura de datos usando las especificaciones de *Dublin Core*.



Figura 5. Ventana que muestra la opción de exportar.

La opción de exportar permite guardar en HTML la estructura generada, esto puede ser publicado en un sitio web o en una base datos, como se muestra en la figura 6 puede ser guardado en un directorio local de la PC.

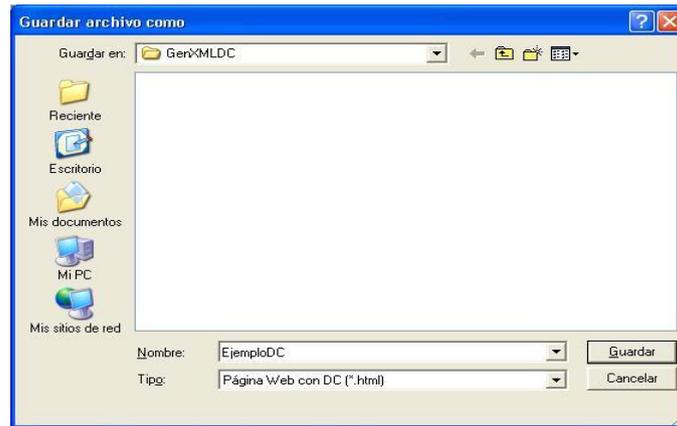


Figura 6. Opción de guardar en un directorio determinado.

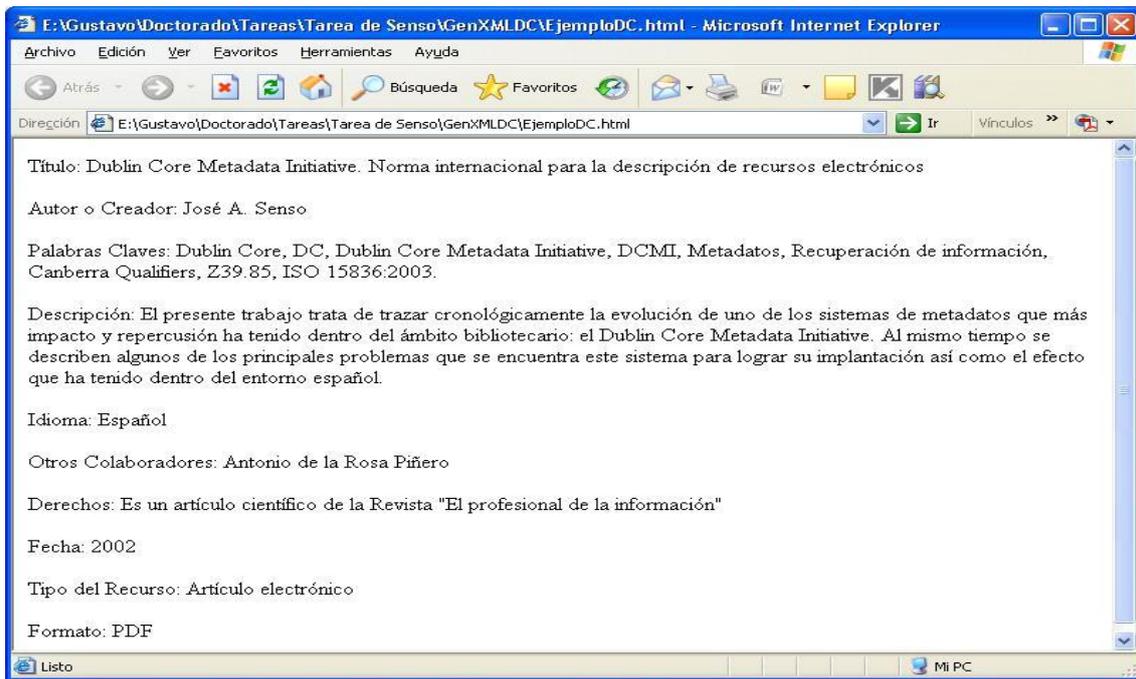


Figura 7. Resultados de la acción de guardar en HTML.

En la figura 7 se muestra en el navegador la composición de los contenidos reflejados en la acción de entrada de datos y luego generar el código utilizando las especificaciones que establece *Dublin Core*, en ambos casos de exportación los códigos que se relacionan con esta acción es el que se muestra a continuación:

```
Private Sub mnExport_Click()
'Exportar HTML
Dim sFile As String
'sFile almacenará el path completo que nos interesa
If Estado = 0 Then
'Guardar en html con los campos de Dublin Core
sFile = GuardarComo("html", "Página Web con DC")
Else
'Guardar en XML
sFile = GuardarComo("xml", "Documento XML")
End If
If sFile <> "" Then
Open sFile For Output As #1
Print #1, txtSalida.Text
```

```

Close
If Estado = 1 Then
    For i = 1 To Len(sFile)
        If Mid(sFile, i, 1) = "\" Then k = i
    Next
    'Coloca en el mismo directorio donde se guarda el XML l ahoja de Estilo.css
    camino = Left(sFile, k) + "Estilo.css"
    Open camino For Output As #2
    Print #2, txtStyle.Text
Close
End If
ShellExecute Me.hwnd, vbNullString, sFile, vbNullString, App.Path, SW_SHOWNORMAL
Else
End If
End Sub

```

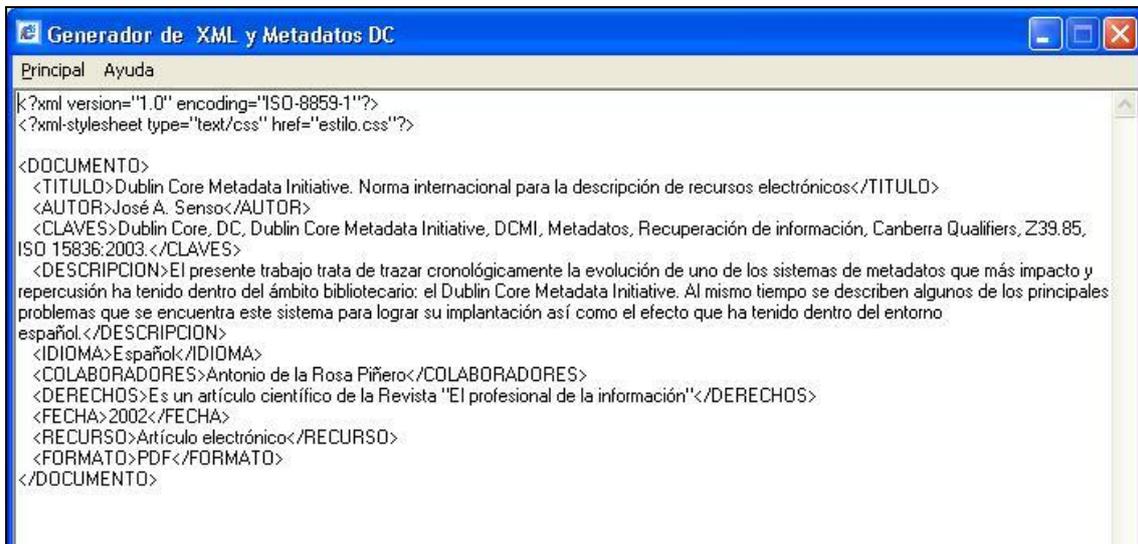


Figura 8. Ventana que muestra la estructura generada en XML.

Cuando se exporta el contenido generado en XML junto con el fichero Ejemplo.XML es guardado un fichero Estilo.CSS, que en este sentido es usado para mostrar el contenido de cada campo usado en XML.

Código generado en XML:

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/css" href="estilo.css"?>

<DOCUMENTO>
  <TITULO>Dublin Core Metadata Initiative. Norma internacional para la descripción de recursos electrónicos</TITULO>
  <AUTOR>José A. Senso</AUTOR>
  <CLAVES>Dublin Core, DC, Dublin Core Metadata Initiative, DCMI, Metadatos, Recuperación de información, Canberra Qualifiers, Z39.85, ISO 15836:2003.</CLAVES>
  <DESCRIPCION>El presente trabajo trata de trazar cronológicamente la evolución de uno de los sistemas de metadatos que más impacto y repercusión ha tenido dentro del ámbito bibliotecario: el Dublin Core Metadata Initiative. Al mismo tiempo se describen algunos de los principales problemas que se encuentra este sistema para lograr su implantación así como el efecto que ha tenido dentro del entorno español.</DESCRIPCION>
  <IDIOMA>Español</IDIOMA>
  <COLABORADORES>Antonio de la Rosa Piñero</COLABORADORES>
  <DERECHOS>Es un artículo científico de la Revista "El profesional de la información"</DERECHOS>
  <FECHA>2002</FECHA>
  <RECURSO>Artículo electrónico</RECURSO>
  <FORMATO>PDF</FORMATO>
</DOCUMENTO>

```

```
<FECHA>2002</FECHA>
<RECURSO>Artículo electrónico</RECURSO>
<FORMATO>PDF</FORMATO>
</DOCUMENTO>
```

Hojas de estilo CSS asociada al archivo generado en XML:

En una hoja de estilo CSS encontraremos, para cada elemento del XML, un bloque donde se especifican las características de presentación para ese tipo de elemento, tales como tamaño de letra, tipo de fuente, etc. a cada elemento le siguen sus propiedades, encerradas entre llaves. Para cada atributo se expresa el nombre y su valor separados por dos puntos.

```
TITULO {display: block; margin-left:15 pt; font-weight: bold; font-size: 12 pt; font-style:normal; color: red;}
AUTOR {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
CLAVES {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
DESCRIPCION {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
FUENTE {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
IDIOMA {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
EDITOR {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
COLABORADORES {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
DERECHOS {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
FECHA {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color:
black;}
RECURSO {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
FORMATO {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
IDENTIFICADOR {display: block; margin-left:15 pt; font-size: 10 pt; font-style: italic; color: black;}
```

Por ejemplo, los títulos deben mostrarse en negrita (bold). El resultado que se obtiene cuando el navegador aplica esta hoja de estilo al documento de la figura 8 es el que se muestra en la figura 9. Para que el navegador sepa cuál es la hoja de estilo que debe aplicar, el documento debe incluir una directiva. En el ejemplo de la figura 8, se trata de la directiva `<?xml-stylesheet type="text/css" href="estilo.css"?>`, en la figura 9 se encuentra todo el contenido de los distintos campos utilizando la estructura de datos usados en XML, la hoja de estilo *estilo.css* de cierta manera formatea este contenido, donde es interpretado por el navegador visualizándolo como se muestra en la figura.

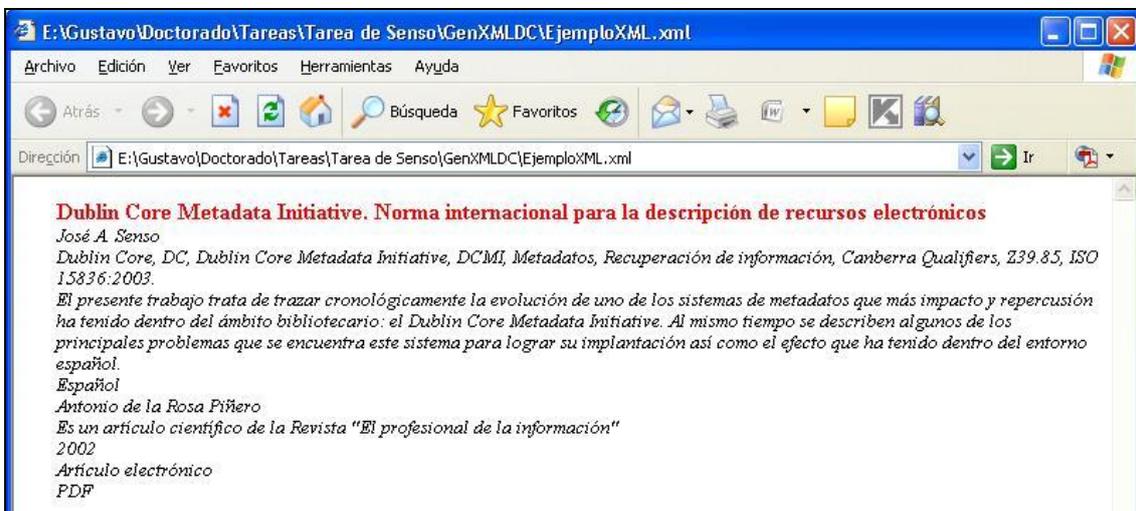


Figura 9. Ventana que muestra al navegador el con el contenido del XML asociado al fichero estilo.css.

En este sentido ha sido demostrado a través del desarrollo de un software como usar las especificaciones del *Dublin Core*, así también con la estructura que describe el XML, estas tecnologías semánticas son de gran difusión en la actualidad, tomándose como base para la descripción de contenidos y documentos electrónicos en la ciencia de la información, los autores del presente trabajo ha querido mostrar una pequeña aplicación de este tipo de tecnología, en la actualidad se pueden encontrar numerosas aplicaciones bien concebidas y funcionales en un entorno donde la interoperabilidad semántica forma una importante cadena en el nuevo contexto donde inevitablemente el mundo se está adentrando hoy en día.

4. CONCLUSIONES

Actualmente las organizaciones que tengan la capacidad para generar y diseminar conocimiento presentarán una de las principales ventajas competitivas en el actual escenario internacional. Es por esto que es prioritario contar con una infraestructura que apoye la creación y compartición de conocimiento.

Se han presentado algunos aspectos teóricos conceptuales de las especificaciones XML y *Dublin Core* y a su vez se muestran estos como una solución para describir documentos electrónicos.

Alrededor de la Web Semántica existen multitud de tecnologías y estándares, como lo son XML y *Dublin Core*, excelentes para modelar información que tiene por sí misma una semántica que puede resultar de vital importancia para las organizaciones tomar decisiones acertadas.

No cabe duda, de que las capacidades del XML y el *Dublin Core* son prometedoras, y así es reconocible por muchas empresas que comercializan aplicaciones Internet, con el cercano seguimiento de la evolución de todos los estándares relacionados con XML y *Dublin Core*.

Se desarrolló un software utilizando como lenguaje de programación el Visual Basic y la metodología SCRUM aplicable para su uso en un entorno de gestión y representación de información científica, la codificación utilizando las especificaciones del XML y el *Dublin Core*, muestra de esta forma una de las innumerables aplicaciones de este tipo de tecnología semántica.

BIBLIOGRAFÍA

- [1] L. Codina Bonilla, "Internet invisible y web semántica:¿ el futuro de los sistemas de información en línea?," *Tradumática*, 2003.
- [2] I. Subirats-Coll, "La web semántica y su aplicación en servicios de información: El caso de la Organización de las Naciones Unidas para la Alimentación y la Agricultura," 2013.
- [3] M. R. Sánchez, "Herramientas para Visualización de Información y Análisis Exploratorio de Datos," 2007.
- [4] G. Ponjuán, M. M. Mugia, M. d. C. V. Álvarez, M. L. Santos, and Y. M. Lahera, "Sistemas de información: principios y aplicaciones," *Editorial Felix Varela*, p. 18, 2007.
- [5] M. C. Pérez and M. Sabelli, "Uso de información electrónica por los académicos de la universidad de la república (Uruguay)," *Información, Cultura y Sociedad*, vol. 9, pp. 29-52, 2003.

- [6] A. Pérez, R. Francisco, G. Sánchez, and M. Barceló, "Avances hacia la Aplicación de las Tecnologías de la Web Semántica en las Organizaciones," *Revista Ingeniería Informática*, 2007.
- [7] T. Saorín, F. Peset, and A. Ferrer-Sapena, "Factores para la adopción de linked data e implantación de la web semántica en bibliotecas, archivos y museos," *Information research*, vol. 18, 2013.
- [8] W. Vega and H. Umaña, "Diseño de Servicios Web Semánticos utilizando el desarrollo de software dirigido por modelos [Semantic Web Services design using model-driven software development]," *Ventana Informática*, 2014.
- [9] D. C. M. Initiative, "Dublin core metadata initiative," 2014.
- [10] D. C. M. Initiative, "DCMI Metadata Terms, 2012," ed, 2015.
- [11] A. E. S. Cantor, "Uso de Ontologías y Web Semántica para apoyar la Gestión del Conocimineto," *Ciencia e Ingeniería Neogranadina*, vol. 17, 2007.
- [12] D. C. M. Initiative, "Dublin Core Metadata," *URL: purl. ocl. org/metadata/dublin_core (2016-10-10)*, 2016.
- [13] A. C. C. Nascimento, R. T. Madeira, and M. B. Da Silva, "Dublin core: a metadata standard in the 3 Marys," in *Proceedings of the 2015 International Conference on Dublin Core and Metadata Applications*, 2015, pp. 247-248.
- [14] E. Porter, "Normalizing decentralized metadata practices using business process improvement methodology: a data-informed approach to identifying institutional core metadata," in *Proceedings of the 2014 International Conference on Dublin Core and Metadata Applications*, 2014, pp. 207-209.
- [15] Y. D. A. Balaguera, "Metodologías ágiles en el desarrollo de aplicaciones para dispositivos móviles. Estado actual," *Revista de Tecnología*, vol. 12, 2015.
- [16] M. C. G. Mantilla, L. L. C. Ariza, and B. M. Delgado, "Metodología para el desarrollo de aplicaciones móviles," *Revista Tecnura*, vol. 18, pp. 20-35, 2014.