

Ejemplos de casos en el Biga Data

Informe preliminar

Bases de Datos Avanzadas

Escuela Superior Informática (UCLM)

Mario Pérez Sánchez-Montañez y David Camuñas Sánchez

20 de abril de 2020

Resumen

Una *base de datos XML* constituye un sistema software que da persistencia a datos almacenados en formato *XML*. Estos datos pueden ser interrogados, exportados y serializados. Las bases de datos *XML* están generalmente asociadas con las bases de datos documentales.

En nuestro caso, crearemos una Base de Datos XML, para el almacenamiento sobre información del fondo de inversión de **BBVA BOLSA USA, FI** regulado por la Comisión Nacional del Mercado de Valores (*CNMV*). Centrándonos en el primer semestre del año 2019.

Keywords: Base de datos documental, XML, CNMV, Banco BBVA, USA

Introducción

El **Big Data** (*macrodatos, datos masivos, inteligencia de datos*), se conoce como el análisis masivo de datos. Esto es una cantidad de datos tan sumamente grande, que las aplicaciones que se dedicaban al procesamiento de datos que tradicionalmente se venían usando no son capaces de tratar y poner en valor en un tiempo razonable. Este término ha estado en uso desde la década de 1990.

Este término también hace referencia, a las nuevas tecnologías que hacen posible el almacenamiento y procesamiento de datos, además del uso que se hace de la información obtenida a través de ellas.

Tipos de datos que existen

A continuación, se dará una breve explicación de los tipos de datos existentes, y de los cuales hace uso esta tecnología (*Big Data*).

Datos estructurados

Estos se suelen usar en el tratamiento de datos. Ya que sus principales características son que pueden ser almacenados en tablas, y que tienen definido su formato.

Algunos datos de este tipo son:

- **Números.**
- **Cadenas de caracteres.**
- **Fechas.**

XML Schema

XML Schema es un lenguaje utilizado para describir la estructura y las restricciones de los contenidos de los documentos XML de una forma muy precisa y explícita, más allá de las normas sintácticas impuestas por el propio lenguaje XML [5].

Como se ha mencionado anteriormente el esquema XML lo hemos generado con la herramienta *Liquid Studio*. El resultado han sido un total de nueve archivos y dada la complejidad del esquema resultante no hemos incluido en la memoria los esquemas. Aunque, estos se pueden visualizar en forma de imagen y en formato *xsd* en los archivos de *XML-Schema* del entregable.

A continuación se mostrará un ejemplo sencillo, en este caso sobre el esquema del elemento *CommunicationWays*, el cual forma parte (es un subelemento) del elemento *InformeFIM*.

El elemento **CommunicationWays**, esta formado por dos subelementos, los cuales son: *CommunicationType* indica el tipo de comunicación que se lleva a cabo, y *CommunicationValue* indica en este caso que el valor será almacenado en un atributo de tipo String. Se puede observar su representación en la Figura 1

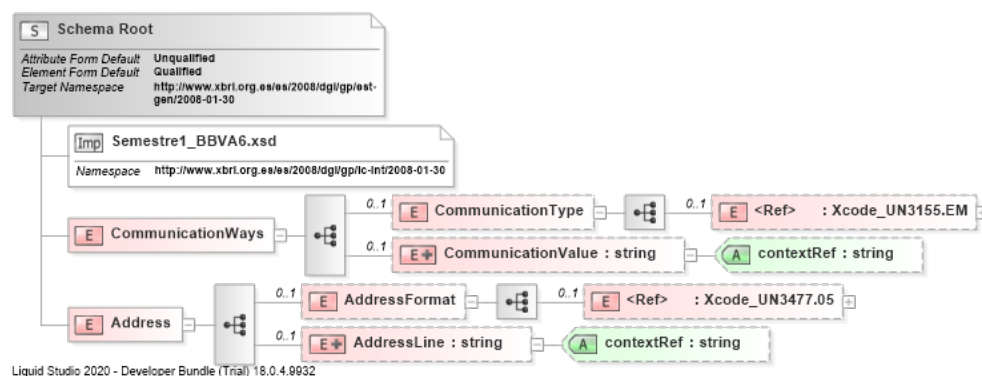


Figura 1: Representación Schema XML del elemento *CommunicationWays*

Donde el código fuente del respectivo archivo de tipo *XSD* sería el siguiente:

```
<xs:element name="CommunicationWays">
  <xs:complexType>
    <xs:sequence>
      <xs:element minOccurs="0" name="CommunicationType">
        <xs:complexType>
          <xs:sequence>
            <xs:element minOccurs="0"
              xmlns:q1="http://www.xbrl.org.es/es/
                2008/dgi/gp/lc-int/2008-01-30"
```


Transformación XSL

En esta sección vamos a ver un ejemplo sencillo de transformación *XSL*, (ya que la transformación de la base de datos de forma completa es bastante complejo).

Transformando un elemento de los que esta formada nuestra base de datos que hemos visto en la sección anterior a un *HTML* sencillo.

Las transformaciones *XSL* o *XSLT* es un estándar de la organización W3C que presenta una forma de transformar documentos XML en otros archivos del mismo o distinto formato. Los archivos XSLT realizan la transformación del documento utilizando una o varias reglas de plantilla.

Estas reglas de plantilla unidas al documento fuente a transformar alimentan un procesador de XSLT, que realiza las transformaciones deseadas poniendo el resultado en un archivo de salida, o, como en el caso de una página web, las hace directamente en un dispositivo de presentación tal como el monitor del usuario [3].

Para realizar esta transformación hemos creado un nuevo archivo xml a partir del original, en el cual hemos incluido una parte de la etiqueta *InformeFIM*. Además de esto, hemos modificado las etiquetas, haciéndolas más sencillas, porque la transformación no se realizaba correctamente con las etiquetas originales. El código xml utilizado para la transformación se puede observar en el siguiente listado:

```
<?xml version="1.0" encoding="UTF-8"?>
<InformeFIM>
  <InformacionGeneralFIM>
    <DenominacionFondo>BBVA BOLSA USA, FI</DenominacionFondo>
    <DescripcionGeneral>
      <RegistroCNMV decimals="0" unitRef="pure">0974</RegistroCNMV>
      <XCode_TipoInforme_02>02</XCode_TipoInforme_02>
      <XCode_PeriodoInforme_02>02</XCode_PeriodoInforme_02>
      <AnoInforme>2019</AnoInforme>
      <Xcode_ISO4217.EUR>EUR</Xcode_ISO4217.EUR>
    </DescripcionGeneral>
    <RatingDepositario>A-</RatingDepositario>
    <FondoCompartimentos>false</FondoCompartimentos>
  </InformacionGeneralFIM>

  <DatosFondoCompartimentoFIM>
    <FechaRegistroFondo>1997-05-14</FechaRegistroFondo>
    <PoliticaDivisaFIM>
      <CategoriaFIM>
        <XCode_TipoFIM_08>08</XCode_TipoFIM_08>
        <XCode_VocacionInversora_09>09</XCode_VocacionInversora_09>
        <PerfilRiesgo>6 en una escala del 1 al 7</PerfilRiesgo>
      </CategoriaFIM>
    </PoliticaDivisaFIM>
  </DatosFondoCompartimentoFIM>
</InformeFIM>
```

```

</CategoriaFIM>
<PoliticaInversion>Fondo que invierte en las principales companas
de bolsa americana. La cartera del fondo est por lo tanto invertida
en USD dolar, sin que se realice cobertura a euros.</PoliticaInversion>
<OperativaDerivados>El Fondo ha realizado operaciones en instrumentos
financieros derivados con la finalidad de inversin y cobertura.
La metodologa aplicada para calcular la exposicin total al riesgo
de mercado es la metodologa del compromiso.Una informacin ms detallada
sobre la poltica de inversin del fondo se puede encontrar
en su folleto informativo.</OperativaDerivados>
<XCode_MetodologiaMedicionRiesgoMercado_01>01
</XCode_MetodologiaMedicionRiesgoMercado_01>
</PoliticaDivisaFIM>
</DatosFondoCompartimentoFIM>
</InformeFIM>

```

La transformación creada se puede apreciar en el siguiente listado:

```

<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:template match="/">
    <html>
      <body>
        <h2>Informe
        <xsl:value-of select=
        "InformeFIM/InformacionGeneralFIM/DenominacionFondo"/>
        </h2>
        <td>Ao Informe: <xsl:value-of select=
        "InformeFIM/InformacionGeneralFIM/DescripcionGeneral/
        AnoInforme"/>
        </td>
        <br/>
        <td>Divisa: <xsl:value-of select=
        "InformeFIM/InformacionGeneralFIM/DescripcionGeneral/
        Xcode_ISO4217.EUR"/>
        </td>
        <br/>
        <td>Calificacin: <xsl:value-of select=
        "InformeFIM/InformacionGeneralFIM/RatingDepositario"/>
        </td>
        <br/>
        <td>Fecha de registro: <xsl:value-of select=
        "InformeFIM/DatosFondoCompartimentoFIM/FechaRegistroFondo"/>
        </td>
        <br/>
        <td>Perfil de riesgo: <xsl:value-of select=

```

```

"InformeFIM/DatosFondoCompartimentoFIM/PoliticaDivisaFIM/
CategoriaFIM/PerfilRiesgo"/>
</td>
<br/>
<td>Poltica de inversin: <xsl:value-of select=
"InformeFIM/DatosFondoCompartimentoFIM/PoliticaDivisaFIM/
PoliticaInversion"/>
</td>
<br/>
<td>Operativa en Derivados: <xsl:value-of select=
"InformeFIM/DatosFondoCompartimentoFIM/PoliticaDivisaFIM/
OperativaDerivados"/>
</td>
<br/>
</body>
</html>
</xsl:template>
</xsl:stylesheet>

```

La idea de este código es construir un *HTML* para mostrar algunos elementos relevantes del informe del fondo de inversión, para ello se hace uso de elementos *HTML* y de la función *XSL*, *value-of select=*. Esta función selecciona un elemento del XML que se desea convertir, indicándole la ruta a seguir dentro de las etiquetas [9]. El resultado en *HTML* de la transformación se puede apreciar en la figura 3.

Informe BBVA BOLSA USA, FI

Año Informe: 2019

Divisa: EUR

Calificación: A-

Fecha de registro: 1997-05-14

Perfil de riesgo: 6 en una escala del 1 al 7

Política de inversión: Fondo que invierte en las principales compañías de bolsa americana. La cartera del fondo está por lo tanto invertida en USD dólar, sin que se realice cobertura a euros.

Operativa en Derivados: El Fondo ha realizado operaciones en instrumentos financieros derivados con la finalidad de inversión y cobertura. La metodología aplicada para calcular la exposición total al riesgo de mercado es la metodología del compromiso. Una información más detallada sobre la política de inversión del fondo se puede encontrar en su folleto informativo.

Figura 3: Resultado de la transformación XSL

Ejemplos de consulta

Uno de los principales objetivos de estas consultas, es la comprobación de que nuestra base de datos funciona de forma correcta, debido a que con una serie de consultas, se puede observar que nuestra base de datos ha almacenado de forma correcta (estructura, formato...) los datos que componen la información.

En nuestro caso cada nodo o elemento por el que esta formada la base de datos es distinta información sobre el primer semestre de este banco, la cual los atributos que la forman son los explicados anteriormente en el apartado anterior.

A continuación se mostrarán una serie de consultas realizadas a nuestra base de datos, junto con su resultado.

Devolver toda la información de los datos económicos del FIM

Consulta

```
for $i in xbrl/InformeFIM/DatosFondoCompartimentoFIM
return $i/DatosEconomicosFIM'
```

Resultado

```
<iic-fim:DatosEconomicosFIM
  xmlns:iic-fim="http://www.cnmv.es/iic/fim/1-2009/2009-03-31"
  xmlns:xbrli="http://www.xbrl.org/2003/instance"
  xmlns:iic-ges="http://www.cnmv.es/iic/ges/1-2009/2009-03-31"
  xmlns:dgi-lc-es="http://www.xbrl.org.es/es/2008/dgi/gp/lc-es/2008-01-30"
  xmlns:dgi-est-gen="http://www.xbrl.org.es/es/2008/dgi/gp/est-gen/2008-01-30"
  xmlns:xbrll="http://www.xbrl.org/2003/linkbase"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:iic-com="http://www.cnmv.es/iic/com/1-2009/2009-03-31"
  xmlns:dgi-lc-int="http://www.xbrl.org.es/es/2008/dgi/gp/lc-int/2008-01-30"
  xmlns:iic-com-fon="http://www.cnmv.es/iic/mep/1-2009/2009-03-31"
  xmlns:dgi-types="http://www.xbrl.org.es/es/2008/dgi/gp/types/2008-01-30"
  xmlns:iso4217="http://www.xbrl.org/2003/iso4217">
  <iic-com:IndiceRotacionCartera
    contextRef="Anexo1_4190_S1_2019_da"
    decimals="2"
    unitRef="pure">0.46
  </iic-com:IndiceRotacionCartera>
  <iic-com:IndiceRotacionCartera
    contextRef="Anexo1_4190_S1_2019_dpp"
    decimals="2"
    unitRef="pure">1.59
  </iic-com:IndiceRotacionCartera>
  <iic-com:IndiceRotacionCartera
    contextRef="Anexo1_4190_S1_2019_daa"
```

```

        decimals="2"
        unitRef="pure">0.46
    </iic-com:IndiceRotacionCartera>
    .
    .
</iic-fim:DatosEconomicosFIM>

```

Devolver todas los datos sobre las Comparativas de rentabilidad media

Consulta

■ En XQuery:

```

for $i in xbrl/ComparativaRentabilidadMedia
return $i

```

■ En XPath:

```

xbrl/ComparativaRentabilidadMedia

```

Resultado

```

<iic-ges:ComparativaRentabilidadMedia
  xmlns:iic-ges="http://www.cnmv.es/iic/ges/1-2009/2009-03-31"
  xmlns:xbrli="http://www.xbrl.org/2003/instance"
  xmlns:iic-fim="http://www.cnmv.es/iic/fim/1-2009/2009-03-31"
  xmlns:dgi-lc-es="http://www.xbrl.org/es/2008/dgi/gp/lc-es/2008-01-30"
  xmlns:dgi-est-gen="http://www.xbrl.org/es/2008/dgi/gp/est-gen/2008-01-30"
  xmlns:xbrll="http://www.xbrl.org/2003/linkbase"
  xmlns:iic-com="http://www.cnmv.es/iic/com/1-2009/2009-03-31"
  xmlns:dgi-lc-int="http://www.xbrl.org/es/2008/dgi/gp/lc-int/2008-01-30"
  xmlns:iic-com-fon="http://www.cnmv.es/iic/mep/1-2009/2009-03-31"
  xmlns:dgi-types="http://www.xbrl.org/es/2008/dgi/gp/types/2008-01-30"
  xmlns:iso4217="http://www.xbrl.org/2003/iso4217">
  <iic-ges:ComparativaVocacionMonetarioCortoPlazoPatrimonio
    contextRef="Anexo1_4190_S1_2019_da"
    decimals="0" unitRef="EUR">0
  </iic-ges:ComparativaVocacionMonetarioCortoPlazoPatrimonio>

  <iic-ges:ComparativaVocacionMonetarioCortoPlazoParticipes
    contextRef="Anexo1_4190_S1_2019_da"
    decimals="0" unitRef="pure">0</
  <iic-ges:ComparativaVocacionRentaFijaEuroPatrimonio
    contextRef="Anexo1_4190_S1_2019_da"
    decimals="0" unitRef="EUR">4095230000
  </iic-ges:ComparativaVocacionRentaFijaEuroPatrimonio>
  <iic-ges:ComparativaVocacionRentaFijaEuroParticipes

```

```

        contextRef="Anexo1_4190_S1_2019_da"
        decimals="0" unitRef="pure">107911
    </iic-ges:ComparativaVocacionRentaFijaEuroParticipes>
    .
    .
</iic-ges:ComparativaRentabilidadMedia>

```

Devolver el valor de las inversiones financieras al exterior

Consulta

```

for $i in xbrl/InversionesFinancieras/InversionesFinancierasValorEstimado/
InversionesFinancierasExterior
return $i/InversionesFinancierasValor

```

Resultado

```

<iic-com:InversionesFinancierasValor
  xmlns:iic-com="http://www.cnmv.es/iic/com/1-2009/2009-03-31"
  xmlns:xbrli="http://www.xbrl.org/2003/instance"
  xmlns:iic-fim="http://www.cnmv.es/iic/fim/1-2009/2009-03-31"
  xmlns:iic-ges="http://www.cnmv.es/iic/ges/1-2009/2009-03-31"
  xmlns:dgi-lc-es="http://www.xbrl.org.es/es/2008/dgi/gp/lc-es/2008-01-30"
  xmlns:dgi-est-gen="http://www.xbrl.org.es/es/2008/dgi/gp/est-gen/2008-01-30"
  xmlns:xbrll="http://www.xbrl.org/2003/linkbase"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dgi-lc-int="http://www.xbrl.org.es/es/2008/dgi/gp/lc-int/2008-01-30"
  xmlns:iic-com-fon="http://www.cnmv.es/iic/mep/1-2009/2009-03-31"
  xmlns:dgi-types="http://www.xbrl.org.es/es/2008/dgi/gp/types/2008-01-30"
  xmlns:iso4217="http://www.xbrl.org/2003/iso4217"
  contextRef="Anexo1_4190_S1_2019_ipy"
  decimals="2"
  unitRef="EUR">13031000
</iic-com:InversionesFinancierasValor>
.
.
<iic-com:InversionesFinancierasValor
  xmlns:iic-com="http://www.cnmv.es/iic/com/1-2009/2009-03-31"
  xmlns:xbrli="http://www.xbrl.org/2003/instance"
  xmlns:iic-fim="http://www.cnmv.es/iic/fim/1-2009/2009-03-31"
  xmlns:iic-ges="http://www.cnmv.es/iic/ges/1-2009/2009-03-31"
  xmlns:dgi-lc-es="http://www.xbrl.org.es/es/2008/dgi/gp/lc-es/2008-01-30"
  xmlns:dgi-est-gen="http://www.xbrl.org.es/es/2008/dgi/gp/est-gen/2008-01-30"
  xmlns:xbrll="http://www.xbrl.org/2003/linkbase"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dgi-lc-int="http://www.xbrl.org.es/es/2008/dgi/gp/lc-int/2008-01-30"
  xmlns:iic-com-fon="http://www.cnmv.es/iic/mep/1-2009/2009-03-31"

```

```

    xmlns:dgi-types="http://www.xbrl.org/es/es/2008/dgi/gp/types/2008-01-30"
    xmlns:iso4217="http://www.xbrl.org/2003/iso4217"
    contextRef="Anexo1_4190_S1_2019_ia"
    decimals="2"
    unitRef="EUR">61101000
</iic-com:InversionesFinancierasValor>

```

Devolver el valor de una inversión financiera en el exterior (la obtendremos por su identificador)

Consulta

```

for $i in xbrl/InversionesFinancieras/InversionesFinancierasValorEstimado/
InversionesFinancierasExterior
where $i/CodigoISIN = "US912828Q525"
return $i/InversionesFinancierasValor

```

Resultado

```

<iic-com:InversionesFinancierasValor
  xmlns:iic-com="http://www.cnmv.es/iic/com/1-2009/2009-03-31"
  xmlns:xbrli="http://www.xbrl.org/2003/instance"
  xmlns:iic-fim="http://www.cnmv.es/iic/fim/1-2009/2009-03-31"
  xmlns:iic-ges="http://www.cnmv.es/iic/ges/1-2009/2009-03-31"
  xmlns:dgi-lc-es="http://www.xbrl.org/es/es/2008/dgi/gp/lc-es/2008-01-30"
  xmlns:dgi-est-gen="http://www.xbrl.org/es/es/2008/dgi/gp/est-gen/2008-01-30"
  xmlns:xbrll="http://www.xbrl.org/2003/linkbase"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dgi-lc-int="http://www.xbrl.org/es/es/2008/dgi/gp/lc-int/2008-01-30"
  xmlns:iic-com-fon="http://www.cnmv.es/iic/mep/1-2009/2009-03-31"
  xmlns:dgi-types="http://www.xbrl.org/es/es/2008/dgi/gp/types/2008-01-30"
  xmlns:iso4217="http://www.xbrl.org/2003/iso4217"
  contextRef="Anexo1_4190_S1_2019_ipy"
  decimals="2"
  unitRef="EUR">13031000
</iic-com:InversionesFinancierasValor>

```

Referencias

1. <http://basex.org/>. Basex.
2. https://developer.mozilla.org/es/docs/Web/XML/Introduccion_a_XML. Introducción a xml.
3. https://es.wikipedia.org/wiki/Extensible_Stylesheet_Language_Transformations. Transformaciones xsl.
4. https://es.wikipedia.org/wiki/Fondo_de_inversi3n. Definición fondo de inversión.
5. https://es.wikipedia.org/wiki/XML_Schema. Xml schema.
6. <https://www.cnmv.es/portal/Consultas/IIC/Fondo.aspx?nif=V81726200&vista=1&fs=12/04/2020>. Bbva bolsa usa, fi.
7. <https://www.liquid-technologies.com/>. Liquid studio 2020.
8. https://www.w3schools.com/xml/tryxslt.asp?xmlfile=cdcatalog&xsltfile=cdcatalog_ex2. Generar xslt.
9. https://www.w3schools.com/xml/xsl_value_of.asp. Xsl value of.
10. <https://www.xml.com/>. Xml.