

Fundamentos de la Web Semántica

Clase práctica

Modelos semánticos - XML



XML

Estándar W3C (World Wide Web Consortium)

Creado para intercambio de documentos semi-estructurados

Sintaxis definida, etiquetas anidadas → **documentos bien formados**

Formato y algunas restricciones de estructura y valores → **documentos válidos con respecto a un documento XML-Schema**

Escenario – COVID-19

El MEC necesita obtener diariamente información sobre los casos positivos de estudiantes y docentes de todas las universidades del país, para aplicar medidas que eviten el aumento de los contagios.

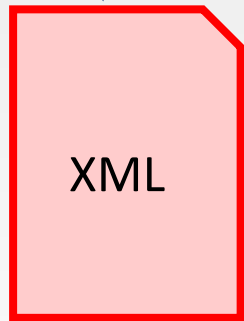
Para ello, solicita al MSP que diariamente le envíe los nuevos casos positivos que pertenecen a este grupo de la población, agrupados por universidad (discriminando si es pública o privada), facultad, carrera, curso. Para cada caso positivo de un curso dado, indicar si se trata de un estudiante o docente, y en caso que participe de otros cursos, se debe poder identificar fácilmente en qué otros cursos participa.

El MSP se compromete a enviar diariamente al MEC un documento XML con la información solicitada. Previamente al inicio de envío de la información, ambos organismos acuerdan un formato de intercambio, para evitar ambigüedades.

Escenario – COVID-19

MSP

Casos positivos



MEC



Escenario – COVID-19

El MEC necesita obtener diariamente información sobre los casos positivos de estudiantes y docentes de todas las universidades del país, para aplicar medidas que eviten el aumento de los contagios.

Para ello, solicita al MSP que diariamente le envíe los nuevos casos positivos que pertenecen a este grupo de la población, agrupados por universidad (discriminando si es pública o privada), facultad, carrera, curso. Para cada caso positivo de un curso dado, indicar si se trata de un estudiante o docente, y en caso que participe de otros cursos, se debe poder identificar fácilmente en qué otros cursos participa.

El MSP se compromete a enviar diariamente al MEC un documento XML con la información solicitada. **Previamente al inicio de envío de la información, ambos organismos acuerdan un formato de intercambio, para evitar ambigüedades.**

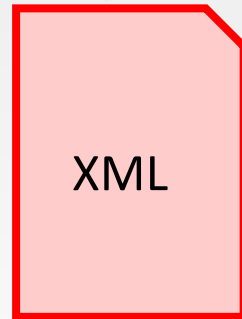
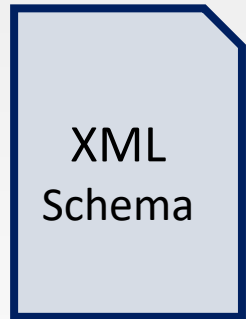
Escenario – COVID-19

MSP



Casos positivos

MEC



Escenario – COVID-19

El MEC necesita obtener diariamente información sobre los casos positivos de estudiantes y docentes de todas las universidades del país, para aplicar medidas que eviten el aumento de los contagios.

Para ello, solicita al MSP que diariamente le envíe los **nuevos casos positivos** que pertenecen a este grupo de la población, **agrupados por universidad (discriminando si es pública o privada), facultad, carrera, curso**. Para cada caso positivo de un curso dado, **indicar si se trata de un estudiante o docente**, y en caso que participe de otros cursos, se debe poder identificar fácilmente en qué otros cursos participa.

El MSP se compromete a enviar diariamente al MEC un documento XML con la información solicitada. **Previamente al inicio de envío de la información, ambos organismos acuerdan un formato de intercambio, para evitar ambigüedades.**

< covidUniversidades >

< universidad >

< nombreUniv > UDELAR < / nombreUniv >

< tipoUniv > PUB < / tipoUniv >

< facultad >

< carrera >

< curso >

< participante >

.....

< tipoParti > DOC < / tipoParti >

< / participante >

< / curso >

< / carrera >

< / facultad >

< universidad >

.....

< / covidUniversidades >

< xsd:element name=" covidUniversidades " >

< xsd:complexType >

< xsd:sequence >

.....

< xsd:element name=" universidad " minOccurs=" 1 " maxOccurs=" unbounded " >

< xsd:complexType >

< xsd:sequence >

.....

< xsd:element name=" tipoUniv " >

..... PUB, PRIV

< / xsd:element >

< xsd:element name=" facultad " minOccurs=" 1 " .. " unbounded " >

< xsd:complexType >

< xsd:sequence >

.....

< xsd:element name=" carrera " .. " 1 " .. " unbounded " >

< xsd:complexType >

< xsd:sequence >

.....

< xsd:element name=" curso " type=" TipoDatoCurso " .. " 1 " .. " unbounded " / >

.....

< xsd:sequence >

< xsd:complexType >

< / xsd:element >

< xsd:complexType name=" TipoDatoCurso " >

< xsd:sequence >

.....

< xsd:element name=" participante " type=" TipoDatoParticipante " minOccurs=" 1 " maxOccurs=" unbounded " / >

< / xsd:sequence >

< xsd:complexType >

< xsd:complexType name=" TipoDatoParticipante " >

< xsd:sequence >

.....

< xsd:element name=" tipoParti " >

..... DOC, EST

< / xsd:element >

< xsd:sequence >

< xsd:complexType >

Escenario – COVID-19

¿Qué podemos representar con las etiquetas?

¿Entidades? ¿Taxonomía de entidades?

¿Atributos?

¿Relaciones entre entidades?

< covidUniversidades >

< univPublica >

< nombreUniv > UDELAR < / nombreUniv >

< facultad >

< carrera >

< curso >

< participante >

< docente >

.....

< / docente >

< estudiante >

.....

< / estudiante >

< / participante >

< / curso >

< / carrera >

< / facultad >

< / univPublica >

< univPrivada >

< nombreUniv > ORT < / nombreUniv >

< facultad >

.....

< / facultad >

< / univPublica >

< / covidUniversidades >

< xsd:element name=" covidUniversidades " >

< xsd:complexType >

< xsd:sequence >

< xsd:element name=" univPublica " >

< xsd:complexType >

< xsd:sequence >

.....

< xsd:element name=" facultad " type=" TipoDatoFacultad " minOccurs=" 1 " maxOccurs=" unbounded " >

< xsd:sequence >

< xsd:complexType >

< / xsd:element >

< xsd:element name=" univPrivada " minOccurs=" 1 " maxOccurs=" unbounded " / >

< xsd:element name=" facultad " type=" TipoDatoFacultad " minOccurs=" 1 " maxOccurs=" unbounded " >

< / xsd:element >

< xsd:sequence >

< xsd:complexType >

< / xsd:element >

< xsd:complexType name=" TipoDatoFacultad " >

...carrera...

< xsd:element name=" participante " type=" TipoDatoCurso " minOccurs=" 1 " maxOccurs=" unbounded " / >

< xsd:complexType >

< xsd:complexType name=" TipoDatoCurso " >

< xsd:element name=" participante " type=" TipoDatoParticipante " minOccurs=" 1 " maxOccurs=" unbounded " / >

< xsd:complexType >

< xsd:complexType name=" TipoDatoParticipante " >

.....

< xsd:element name=" docente " >

.....

< / xsd:element >

< xsd:element name=" estudiante " >

.....

< / xsd:element >

< xsd:complexType >

Escenario – COVID-19

La semántica de las etiquetas es ambigua

Aunque se utilice un vocabulario de un namespace, no distinguimos entre categorías, relaciones, atributos

Escenario – COVID-19

El MEC necesita obtener diariamente información sobre los casos positivos de estudiantes y docentes de todas las universidades del país, para aplicar medidas que eviten el aumento de los contagios.

Para ello, solicita al MSP que diariamente le envíe los nuevos casos positivos que pertenecen a este grupo de la población, agrupados por universidad (discriminando si es pública o privada), facultad, carrera, curso. Para cada caso positivo de un curso dado, indicar si se trata de un estudiante o docente, y **en caso que participe de otros cursos, se debe poder identificar fácilmente en qué otros cursos participa.**

El MSP se compromete a enviar diariamente al MEC un documento XML con la información solicitada. **Previamente al inicio de envío de la información, ambos organismos acuerdan un formato de intercambio, para evitar ambigüedades.**

Escenario – COVID-19

```
<xsd:complexType name="TipoDatoCurso">  
  <xsd:sequence>  
    <xsd:element name="idCurso" type="xsd:positiveInteger"/>  
    <xsd:element name="nombreCurso" type="xsd:string"/>  
    <xsd:element name="participante" type="TipoDatoParticipante" minOccurs="1" maxOccurs="unbounded"/>  
  </xsd:sequence>  
</xsd:complexType>
```

```
<xsd:complexType name="TipoDatoParticipante">  
  <xsd:sequence>  
    <xsd:element name="idParti" type="xsd:positiveInteger"/>  
    <xsd:element name="nombreParti" type="xsd:string"/>  
    <xsd:element name="idOtroCurso" type="xsd:positiveInteger" minOccurs="0" maxOccurs="unbounded"/>  
  </xsd:sequence>  
</xsd:complexType>
```

```
<xsd:complexType name="TipoDatoCurso">
  <xsd:sequence>
    <xsd:element name="idCurso" type="xsd:positiveInteger"/>
    <xsd:element name="nombreCurso" type="xsd:string"/>
    <xsd:element name="participante" type="TipoDatoParticipante" minOccurs="1" maxOccurs="unbounded"/>
  </xsd:sequence>
</xsd:complexType>
```

```
<xsd:complexType name="TipoDatoParticipante">
  <xsd:sequence>
    <xsd:element name="idParti" type="xsd:positiveInteger"/>
    <xsd:element name="nombreParti" type="xsd:string"/>
    <xsd:element name="idOtroCurso" type="xsd:positiveInteger" minOccurs="0" maxOccurs="unbounded"/>
  </xsd:sequence>
</xsd:complexType>
```

```
.....
<xsd:key name="claveCurso">
  <xsd:selector xpath="./curso"/>
  <xsd:field xpath="idCurso"/>
</xsd:key>
<xsd:keyref name="refClaveCurso" refer="claveCurso">
  <xsd:selector xpath="./participante"/>
  <xsd:field xpath="idOtroCurso"/>
</xsd:keyref>
```

```
.....
```

Escenario – COVID-19

<curso>

<idCurso>**1**</idCurso>

<nombreCurso>**Fundamentos de Base de Datos**</nombreCurso>

<participante>

<idParti>46176532</idParti>

<nombreParti>**Juan Perez**</nombreParti>

</participante>

<participante>

<idParti>33109061</idParti>

<nombreParti>**Sandra Medina**</nombreParti>

<idOtroCurso>**2**</idOtroCurso>

</participante>

.....

.....

</curso>

<curso>

<idCurso>**2**</idCurso>

<nombreCurso>**Introducción a la ingeniería de software**</nombreCurso>

.....

XML - Limitaciones

XML describe documentos, no recursos (personas, libros, etc.)

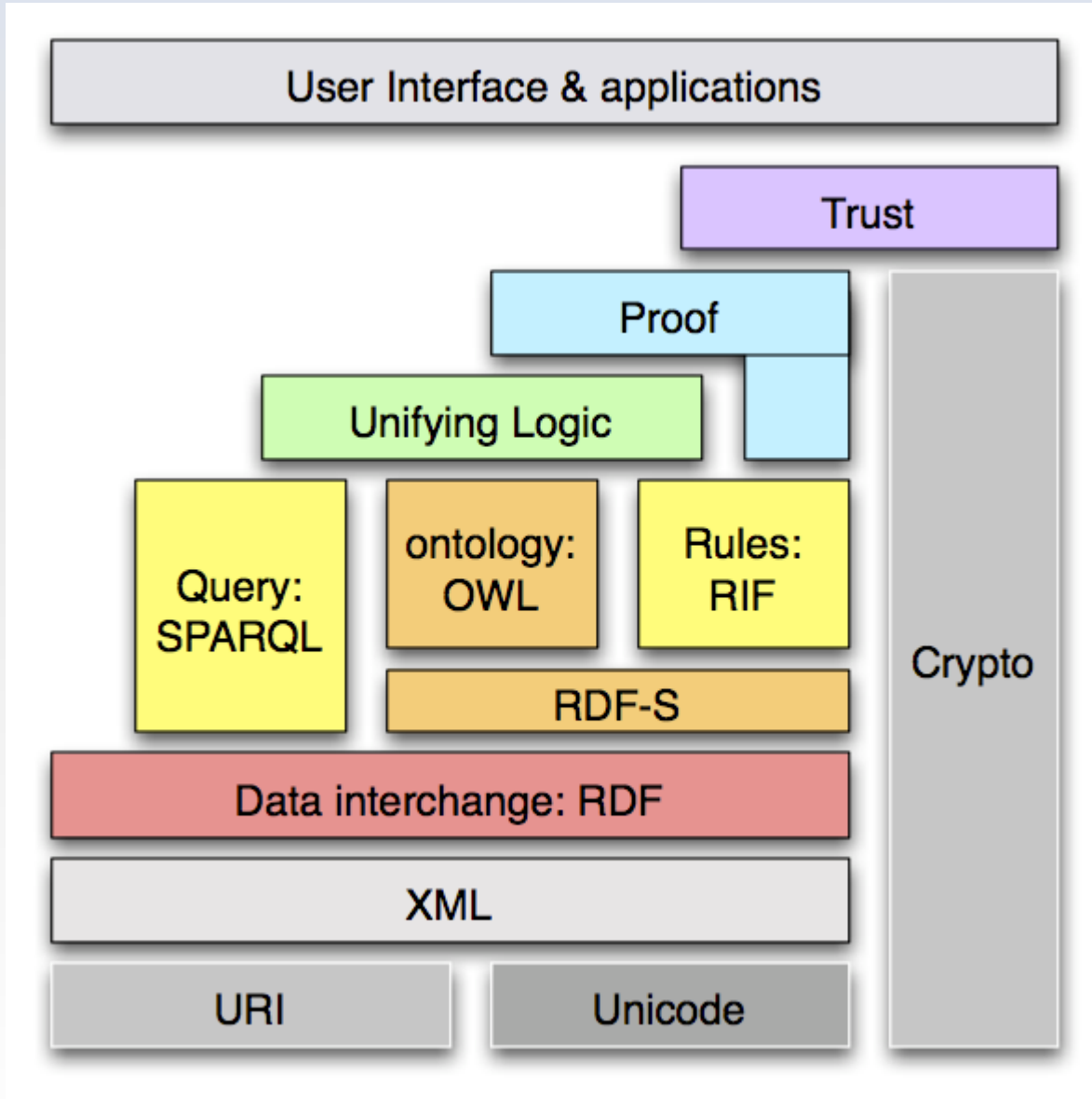
Foco en la estructura, no en el significado de recursos ni relaciones entre ellos

“El docente Juan Pérez tiene cédula de identidad 4105697”

```
<docente>  
  <nombre> Juan Pérez </nombre>  
  <ci> 4105697 </ci>  
</docente>
```

```
<docente nombre = “Juan Pérez”>  
  <ci> 4105697 </ci>  
</docente>
```


Arquitectura de la Web Semántica



Cada capa:

- entiende y se apoya en la capa inferior
- agrega semántica a la capa inferior

Unicode: estándar de codificación universal de caracteres

Uniform Resource Identifier (URI): estándar para identificación de recursos (cualquier cosa identificable)

XML, XML-Schema, Namespaces: descripción de documentos de forma estructurada, semántica débil. Sintaxis base para capas superiores.

RDF, RDF-Schema: descripción de recursos, relaciones entre recursos y clases de recursos. Integración de aplicaciones.

XML

<https://www.w3.org/TR/REC-xml/>

<https://www.w3.org/TR/xmlschema11-1/>

<https://www.liquid-technologies.com/online-xsd-validator>