PROYECTO NORMA TÉCNICA:
ESTRUCTURA DE METADATO PARA INTEROPERABILIDAD

Creative Commons

Reconocimiento-No comercial-Compartir bajo la misma licencia 3.0.

Usted es libre de:

Copiar, distribuir y reproducir públicamente la obra.

Hacer obras derivadas.

Bajo las siguientes condiciones:

Reconocimiento. Debe reconocer los créditos de la obra de la manera especificada por el autor o el licenciante (pero no de una manera que sugiera que tiene su apoyo o apoyan el uso que hace de su obra).

No comercial. no puede utilizar esta obra para fines comerciales.

Compartir bajo la misma licencia. Si altera o transforma esta obra, o genera una obra derivada, sólo puede distribuir la obra generada bajo una licencia idéntica a ésta.

Al reutilizar o distribuir la obra, tiene que dejar bien claro los términos de la licencia de esta obra. Alguna de estas condiciones puede no aplicarse si se obtiene el permiso del titular de los derechos de autor.

Nada en esta licencia menoscaba o restringe los derechos morales del autor.

Los derechos derivados de usos legítimos u otras limitaciones reconocidas por ley no se ven afectados por lo anterior.

Esto es un resumen fácilmente legible del texto legal de su versión original en idioma inglés (la licencia completa)

http://creativecommons.org/licenses/by-nc-sa/3.0/ve/

Contenido

Capítulo I: Generalidades....4
1.Objeto....4
2.Ámbito de Aplicación....4
3.Siglas y acrónimos....4
4.Definiciones....4
5.Formatos de metadatos....6
6.Nombre canónico....6
7.Mantenibilidad y escalabilidad....6
8.Representación física del metadato....6
Capítulo II: De la construcción de los esquemas XML....7
9.Reglas generales....7
10. Esquema XML....7
11. Identificación del esquema....7
12. Versión del esquema....9
13. Listas codificadas....10
14. Diagrama de clases....10
15. Diccionario de datos....10
16. Especificación del metadato....10
17. Estructura general del documento XMI....10
ANEXO 1....12
ANEXO 2....13

Capítulo I: Generalidades

1. Objeto

Establecer una estructura de metadato que permita describir, buscar, recuperar y preservar un dato u objeto digital facilitando la Interoperabilidad entre los sistemas de información del Poder Público del Estado Venezolano.

2. Ámbito de Aplicación

Este proyecto de norma técnica será aplicado al Poder Público a que se refiere el artículo 2 de la Ley Decreto con Rango, Valor y Fuerza de acceso e intercambio electrónico de datos, información y documentos entre los órganos y entes del Estado.

3. Definiciones

A efectos del presente proyecto de norma técnica se presentan las siguientes definiciones:

1. **Estándares abiertos:** Especificaciones técnicas, publicadas y controladas por alguna organización que se encarga de su desarrollo, aceptadas por la industria de las tecnologías de información, y que están a disposición de cualquier usuario para ser implementadas.
2. **Escalabilidad**: Propiedad deseable de un metadato que indica su habilidad para reaccionar y adaptarse sin perder calidad.
3. **Mantenibilidad**: Representa la capacidad de un metadato para ser modificado efectiva y eficientemente, debido a necesidades evolutivas, correctivas o perfectivas
4. **Metadatos:** Se refiere a toda aquella información descriptiva sobre el contexto, calidad, condición o características de un contenido digital, dato u objeto, que tiene la finalidad de facilitar su búsqueda, recuperación, preservación e interoperabilidad.
5. **Modelo de datos :** Representación gráfica o léxica de datos, especificando sus propiedades y sus interrelaciones.
6. **Objetos digitales:** También conocidos como documentos electrónicos, se refiere a todo soporte de información producida, recibida, enviada, almacenada, organizada, usada y procesada por medios electrónicos o migrada a estos, a través de un proceso de digitalización que requiere de una herramienta específica para ser legible o recuperable.
7. **Software libre:** Programa de computación en cuya licencia el autor o desarrollador garantiza al usuario el acceso al código fuente y lo autoriza a usar el programa con cualquier propósito, copiarlo, modificarlo y redistribuirlo con o sin modificaciones, preservando en todo caso el derecho moral al reconocimiento de autoría.

8. **Instituciones públicas:** Se refiere a los órganos y entes del Poder Público Nacional, Estadal y Municipal, las personas de derecho público nacionales, estadales y municipales y demás entes de carácter público, así como, a las universidades públicas nacionales autónomas y experimentales, centros y colectivos de investigación y desarrollo y cualquier otra institución de los sectores universitario, académico, científico o tecnológico de naturaleza pública.
9. **Interoperabilidad:** Capacidad de los órganos y entes del Estado de intercambiar por medios electrónicos datos, información y documentos de acceso público.
10. **Interoperabilidad sintáctica**: Se refiere a la capacidad de un sistema o componente de un sistema para proveer información portable entre aplicaciones, así como un mecanismo de control.
11. **Interoperabilidad semántica**: Conocimiento de un dominio necesario para que dos servicios informáticos se "comprendan" en término de sus capacidades e intenciones.
12. **Nombre canónico:** Se refiere al nombre dado a un metadato según un *canon* o regla preestablecida.

4. Siglas y acrónimos

El presente proyecto de norma técnica utilizará las siguientes siglas y acrónimos en idioma inglés, reconocidos en el vocabulario técnico universal:

1. **ABNF:** Augmented Backus–Naur Form (Formulario Aumentada Backus-Naur).
2. **EBNF:** Extended Backus–Naur Form (BNF extendido)
3. **UN/CEFACT:** Centro de las Naciones Unidas para la Facilitación del Comercio y el Comercio Electrónico.
4. **UML** Unified Modeling Language (Lenguaje Unificado de Modelado).
5. **UTF 8: 8-**bit Unicode Transformation Format (Formato de transformación Unicode)
6. **URL:** Uniform Resource Locator (Localizador de Recurso Uniforme).
7. **CDATA**: Character Data (Datos de carácter).
8. **XML:** Xtensible Markup Language (Lenguaje de Marcas Extensible).
9. **XML Schema Part 1: Structures:** Esquema XML Parte 1: Estructuras.
10. **XML Schema Part 2: Data Types:** Esquema XML Parte 2: Tipos de datos.
11. **W3C:** World Wide Web Consortium. (Consorcio World Wide Web).

Capítulo II: De la Presentación del Metadato

5. Formatos de metadatos

El formato de metadatos se refiere a las reglas usadas para estructurar los campos o elementos de los metadadtos. Es por ello que se definen tres niveles de abstracción sobre un descriptor de metadatos:

1. Un primer nivel donde se describa el contexto del metadato; en este contexto debe estar claro quien o quienes son los responsables de darle mantenimiento; en general corresponde con quien produzca los datos asociados. Se recomienda el lenguaje unificado de modelado UML. Se describen, mediante el uso de una jerarquía, las entidades u organizaciones responsables del metadato y sus relaciones
2. Un segundo nivel, en el cual se detallan la estructura y comportamiento de los datos que son descritos. Se debe proveer un diagrama de clases en UML describiendo la estructura y comportamiento asociado al metadato. Pueden incluirse cualquier diagrama adicional que describa mejor su comportamiento y su estructura.
3. Un tercer nivel, donde se especifique al detalle el metadato. Para ello se referencia los formatos sugeridos en el Anexo 1.

6. Nombre canónico

Se debe utilizar la jerarquía de clases resultante para obtener una nomenclatura que permita identificar los metadatos. Esto se logra concatenando, empezando desde la raíz, los diferentes nombres de las entidades hasta llegar al nombre del metadato.

7. Mantenibilidad y escalabilidad

Los metadatos deberán ser corregidos, mejorados o evolucionar mediante la modificación de los diagramas UML y reglas EBNF. Ambos métodos contempla y proveen mecanismo para ello. Por otro lado, los metadatos pueden extenderse mediante la concatenación de diagramas, así como también sus especificaciones.

8. Representación física del metadato

Los metadatos y todos sus elementos (*i.e.* diagramas UML, diccionario de datos y su especificación) deberán estar consolidados en un solo documento XML. El formato particular es XLM Metadata Interchange (XMI).

Capítulo II: De la construcción de los esquemas XML

9. Reglas generales

Se deberá incluir la codificación del esquema XML descritas en este apartado, a los efectos de asegurar la máxima interoperabilidad. Tal como se describe a continuación:

1. No usar: xsd:notation; xsd:any; xsd:redefine; xsd:substitutionGroup; xsd:ID/xsd:IDREF.
2. Todos los elementos declarados deben tener un nombre, no se deberían usar un tipo "anónimo".
3. Cada objeto de metadatos definido en el esquema XML debe incluir la información contenida en el atributo "Definición" en la estructura "xsd:documentation", a modo de documentación.
4. Los nombres de elementos, atributos y tipos se corresponderán con el valor asignado a "Nombre Corto".
5. Un "xsd:complexType" debe ser usado para definir clases de objetos, documentos o elementos de datos, así como estructuras complejas que no puedan ser definidas usando "xsd:simpleType".
6. Los nombres de atributos deben usar la notación LCC.
7. Los nombres de elementos y tipos deben usar la notación UCC.
8. Se recomienda no usar acrónimos a menos que estos sean parte de un vocabulario controlado.
9. Se recomienda el uso del sufijo LC para definir una Lista Codificada. Como ejemplo, si los valores posibles de "Tipo Documento" se definen como una lista de valores posibles, entonces el nombre de la lista debe ser "TipoDocumentoLC".
10. Se recomienda el uso del sufijo "Tipo" en los nombres de dominios específicos.

10. Esquema XML

En caso de requerir mayor especificidad se debe adoptar como reglas generales de diseño de los esquemas XML las recomendaciones de la W3C, XML Schema Part 1: Structures y XML Schema Part 2: Data Types, y las reglas de diseño de esquemas XML adoptadas por la UN/CEFACT.

11. Identificación del esquema

Los esquemas XML deben adoptar el formato de codificación UTF-8. Se recomienda proporcionar en el esquema XML la información asociada con la identificación del esquema, la información de la organización y los derechos de propiedad de acuerdo al siguiente modelo:

```
<?xml version="1.0" encoding="UTF-8"?
<xsd:identification>
    nombre del metadato
```

```
    codigo (según la jerarquia)
</xsd:idenfitication>
<xsd:dependencies>
    establecer dependencias a otros metadatos, incluyendo el URL donde esta ubicado su XMI. Puede
utilizarse cualquier mecanismo válido de XML para la inclusion de documentos remotos.
</xsd:dependencies>
<xsd:documentation>
   Plataforma de Gobierno Electrónico CNTI
   Organización: CNTIVersión: nn.mm Fecha: aaaammdd
   Contexto/Modelo: NOMBRE CANÓNICO DEL METADATO
   Derechos de Propiedad
            Este esquema XML ha sido elaborado por CNTI. Usted es
            libre de usar, copiar, distribuir, comunicar y difundir
            públicamente este esquema así como hacer obras derivadas,
            siempre y cuando tenga en cuenta citar la obra de forma
            específica y no utilizar esta obra para fines comerciales.
            Toda obra derivada de esta deberá ser generada con estas
           mismas condiciones.
</xsd:documentation>
```

Nombre de Espacio (Namespace)

Se deberá adoptar para la codificación de los nombres de los *namespaces* las siguientes reglas:

a) Nombre del *namespace default*:

```
<xml schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
```

El prefijo “xsd” será usado para referirse a todos los elementos asociados a dicho *namespace*. Estructura de los nombres de *namespaces* alternativos:

b) Estructura de los nombres de *namespaces* alternativos, soportados bajo la Plataforma de Interoperabilidad.

Metadatos.xx.ve	Nivel de organizacion	Nivel tipo esquema	Nivel contexto	Versión

Donde :

1. El Nivel Organización identifica la Organización responsable por el mantenimiento del esquema XML.
2. El Nivel Tipo Esquema puede adoptar los siguientes valores: Datos; Lista Codificada y Documento.
3. El Nivel Contexto puede adoptar el nombre del contexto asociado al esquema, o adoptar el valor "Uso Comun" para definiciones comunes de tipos de datos.

4. Versión, refiere a la versión del esquema en formato Vmmpnn.

El atributo xsd:elementFormDefault debe ser declarado como "qualified" y el atributo "xsd:attributeFormDefault" debe ser declarado como "unqualified".

```
elementFormDefault="qualified" attributeFormDefault="unqualified"
```

Un ejemplo de un namespaces:

```
http://metadatos.xx.xx.ve/cnti/documento/fideV01p01.xsd
```

12. Versión del esquema

"Vmmpnn" se refiere a la versión del esquema XML. El valor "mm" se refiere a cambios mayores no compatibles en el esquema XML, inicia en 01 y se incrementa bajo las siguientes condiciones:

1. Cambio de nombre de un elemento o atributo.
2. Cambio de un patrón de un tipo de datos.
3. Cambio de estructuras y relaciones.
4. Inclusión o exclusión de elementos o atributos obligatorios.
5. Cambios de cardinalidad.
6. Cualquier otro ajuste que requiera una revisión de los procedimientos que usan dicho esquema.

El valor "nn" se refiere a cambios menores que son compatibles con la última versión, inicia en 01 y se incrementa bajo las siguientes condiciones:

1. Inclusión o corrección de información documental y descriptiva.
2. Inclusión de un patrón en un tipo de datos.
3. Inclusión de nuevos elementos o atributos de carácter no obligatorio.
4. Inclusión de estructuras o relaciones opcionales.
5. Cualquier otro cambio que no requiera la revisión de los procedimientos que usan dicho esquema.

13. Listas codificadas

Se deberá adoptar para las listas codificadas las siguientes reglas:

1. Las listas codificadas internas deben ser definidas cuando no existen listas codificadas externas

definidas o cuando las listas codificadas existentes deben ser extendidas.

2. Las listas codificadas internas no deberían duplicar listas codificadas externas que pueden ser importadas.
3. Una lista codificada que pueda ser usada en otros esquemas, debería ser definida en un esquema XML propio o de "Uso Común" para posterior importación.
4. Cada lista codificada debe ser definida usando "xsd:simpleType".
5. Cada código en una lista codificada debe ser identificado como "xsd:enumeration" donde "xsd:value" es el valor del código actual.

14. Diagrama de clases

Los diagramas y modelos asociados al metadato pueden ser incluidos bajo la etiqueta <xsd:model>.
Se puede incluir todos los XMI relacionados con el mismo. Hay herramientas de software libre las cuales trabajan de manera nativa con el formato XMI.

15. Diccionario de datos

Los diccionarios de datos asociados al metadato deben ser incluidos como una entidad CDATA bajo la etiqueta <xsd:data>.

16. Especificación del metadato

Las especificaciones del metadato deberán ser incluidas bajo la etiqueta <xsd:especification>. Ahí pueden incluirse las descripciones en formatos como DublinCore y Marc21.

17. Estructura general del documento XMI

Esta estructura debe ser una guía para conformar el documento XML.

1. XML.
2. Identificación.
3. Documentación.
4. Dependencias.
5. Modelos.
6. Diccionario de datos.
7. Especificación.

ANEXO 1

1. Formatos de metadatos

Para comprender mejor el panorama de los formatos, a continuación presentamos la tabla 1. En esta tabla se muestran los diferentes formatos estudiados. Se describe su orientación (*i.e.* recurso físico, documento, recurso en red o contenido). La fecha de creación del formato, se indica la afiliación u origen(es) asociado(s) y el tipo (*i.e.* informal, semi-formal o formal).

Formato	Orientado a	Creación	Afiliación/Origen	Tipo
IAFA/Whois++ Templates	Sitios FTP/Archivos en Filesystem	1995	Internet Engineering Taskforce (IETF)	Semi-formal
Dublin Core	Documento	1995	Dublin Core Metadata Initiative (DCMI)	Semi-formal
Text Coding Iniciative (TEI)	Documento	1980	TEI Consortium	Formal
MARC	Documento	1960	Biblioteca del Congreso de Estados Unidos	Semi-formal
Gramáticas Formales	Contenido	1950	Noam Chomsky/Bacus Naur	Formal

Tabla 1 formatos.

ANEXO 2

Ejemplo: Especificación del Metadato Registro civil

Nombre canónico:

GOB::CNE::registro_civil::acta_nacimiento.

Modelo

Se utiliza la notación UML para especificar los diagramas. Utilizamos la convención *snake case* para escribir los nombres de clases, atributos y otros elementos de la notación.

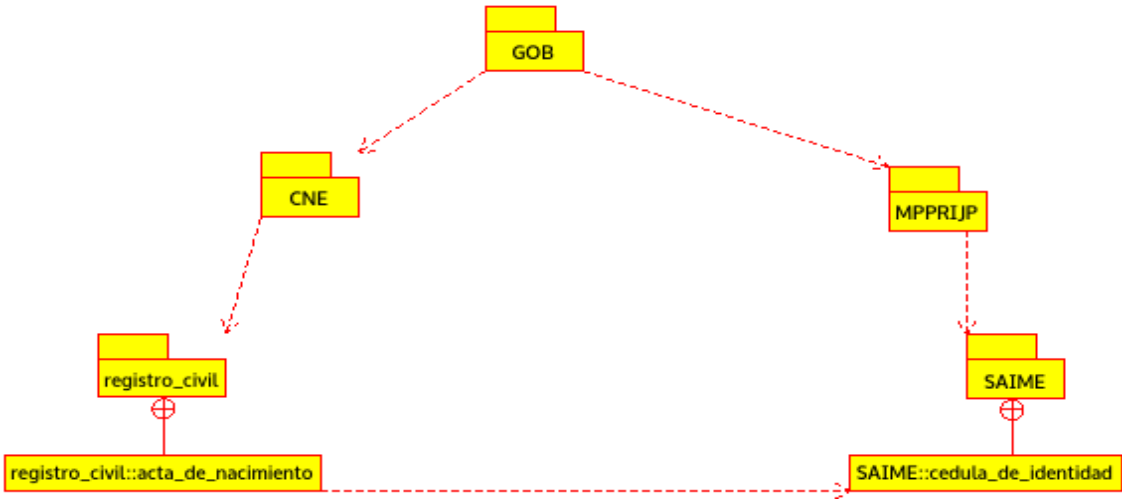

Figura 1: Taxonomía Organizacional

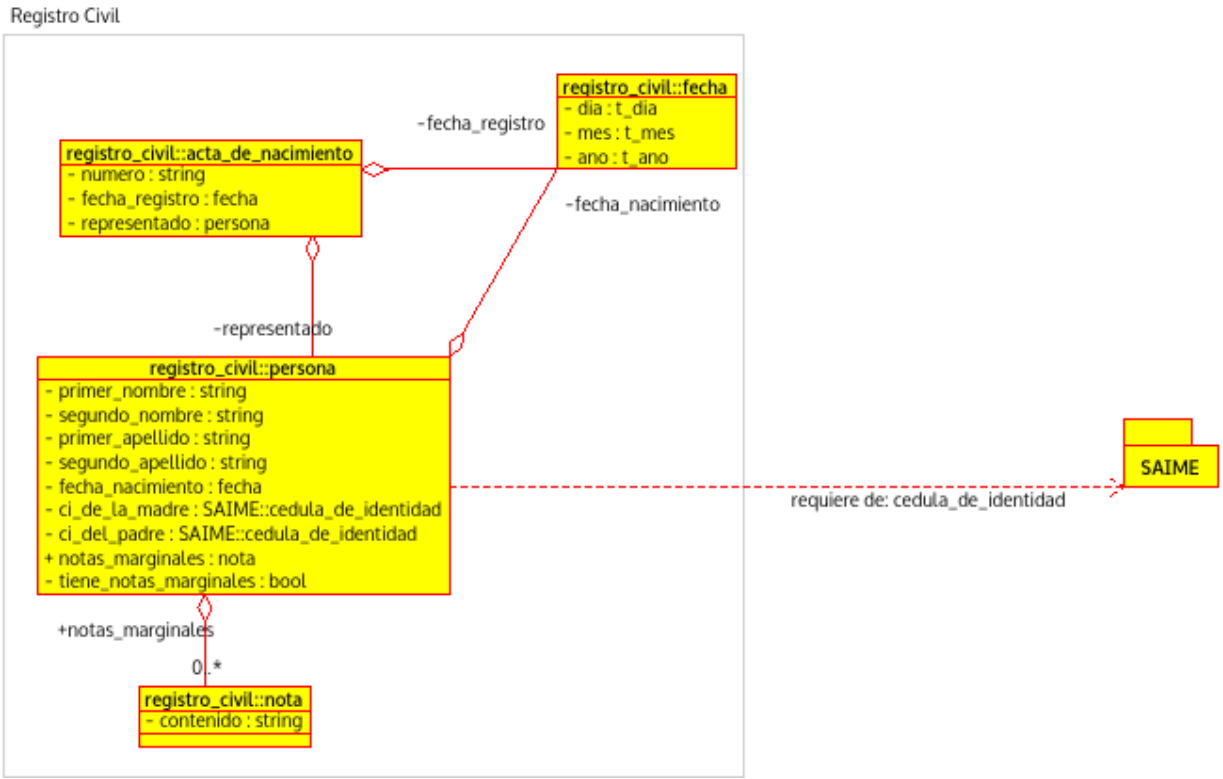

Figura 2: Diagrama representando la organización del metadato

Diccionario de datos

Para desarrollar el diccionario de datos, utilizamos una forma adaptada de la notación ABNF (Augmented Backus–Naur Form)[1]. Se escriben las reglas tal y como esta descrito en la especificación. Las condiciones locales a la regla se escriben en llaves a continuación de la misma.

acta_de_nacimiento = numero fecha_registro representado
numero ::= string
fecha_registro ::= t_día “/” t_mes “/” t_año
representado ::= primer_nombre
segundo_nombre
primer_apellido
segundo_apellido
fecha_nacimiento
ci_de_la_madre
[ci_del_padre]
notas_marginales
tiene_notas_marginales

primer_nombre ::= string
segundo_nomber ::= string
primer_apellido ::= string
segundo_apellido ::= string
fecha_nacimiento ::= t_día “/” t_mes “/” t_año

[1] https://tools.ietf.org/html/rfc5234

cedula_de_la_madre ::= GOB::MPPRIJP::SAIME::cedula_de_identidad
cedula_del_padre ::= GOB::MPPRIJP::SAIME::cedula_de_identidad
notas_marginales ::= nota+
nota ::= string
tiene_notas_marginales ::= verdadero / falso

string ::= 255caracter
caracter ::= minuscula / mayuscula / digito / simbolo

t_día ::= digito digito {01 <= t_dia <= 31}
t_mes ::= digito digito {01 <= t_mes <= 12}
t_año ::= digito digito digito digito {t_año >= 1900}

mayuscula ::= “A” / “B“ / “C“ / “D“ / “E“ / “F“ / “G“ / “H“ / “I“ / “J“ / “K“ / “L“ / “M“ / “N“ / “Ñ“ / “O“ / “P“ / “Q“ / “R“ / “S“ / “T“ / “U“ / “V“ / “W“ / “X“ / “Y“ / “Z“ / “À“ / “Á“ / “Â“ / “Ã“ / “Ä“ / “Å” / “È” / “É” / “Ê” / “Ë” / “Ì” / “Í” / “Î” / “Ï” / “Ò” / “Ó” / “Ô” / “Õ” / “Ö” / “Ù” / “Ú” / “Û” / “Ü”

minuscula ::= “a” / “b“ / “c“ / “d“ / “e“ / “f“ / “g“ / “h“ / “i“ / “j“ / “k“ / “l“ / “m“ / “n“ / “ñ“ / “o“ / “p“ / “q“ / “r“ / “s“ / “t“ / “u“ / “v“ / “w“ / “x“ / “y“ / “z“ / “à“ / “á“ / “â“ / “ã“ / “ä“ / “å” / “è” / “é” / “ê” / “ë” / “ì” / “í” / “î” / “ï” / “ò” / “ó” / “ô” / “õ” / “ö” / “ù” / “ú” / “û” / “ü”

digito ::= “0” / “1” / ”2” / ”3” / ”4” / ”5” / ”6” / ”7” / ”8” / ”9” ;

simbolo ::= ”[” / ”]” / ”{” / ”}” / ”(” / ”)” / ”<” / ”>” / ”” / ”” / ”=” / ”/” / ”.” / ”,” / ”;” / ” ”

Especificaciones

El metadato puede ser especificado en una diversa cantidad de formatos (e.g. Dublin Core, XML schema, Marc21, etc). De ser mas de uno deben ir numerados.

1. Especificación Dublin Core
2. XML Schema
3. Marc21
4. …