

GESTION DEL CONOCIMIENTO EN EL GOBIERNO ABIERTO: ONTOLOGIAS DE DOMINIO

Mg. Graciela Brusa
Cámara de Diputados
Gral. López 3050
Santa Fe

gracielabrusa@santafe.gov.ar

Dra. Ma. Laura Caliusco
CIDISI – UTN - FRSF
Lavaise 610
Santa Fe

mcaliusc@frsf.utn.edu.ar

Dr. Omar Chiotti
CIDISI – UTN - FRSF
Lavaise 610
Santa Fe

chiotti@ceride.gov.ar

RESUMEN

La evolución del Gobierno Electrónico hacia lo que hoy se conoce como Gobierno Abierto se sustenta fuertemente, entre otros elementos, en la evolución de las nuevas Tecnologías de la Información y las Comunicaciones (TICs). La comprensión de la información a la que se accede o se intercambia es la clave para su correcta utilización en la toma de decisiones, por ello, Gobierno Abierto no significa sólo mostrar los datos que el Estado posee sino facilitar su comprensión.

Este trabajo se enfoca en una solución tecnológica que permita contar con información disponible y abierta, no sólo para su utilización sino también para su acabada comprensión en un determinado contexto por parte del ciudadano. Esto exige un importante esfuerzo desde los diferentes niveles de la interoperabilidad, muy difícil de realizar por las propias dificultades de funcionamiento y magnitud de la administración gubernamental.

Se trabajo propone aquí un modelo para trabajar en el nivel más alto de la interoperabilidad, que es la interoperabilidad semántica, a través de la utilización de ontologías de dominio. Por otra parte, se aporta un método específico para resolver la complejidad del modelado conceptual de una ontología en escenarios complejos en los cuales las principales fuentes de información no son simplemente las bases de datos sino documentos legales y personas.

1. INTRODUCCION

El Gobierno Abierto puede pensarse como una evolución del gobierno electrónico donde existe menos normativa y donde las redes sociales y las nuevas formas de colaboración ayudan en el desarrollo de las relaciones entre el estado y los ciudadanos. Para una mejor comunicación entre el estado y los ciudadanos, el gobierno provee acceso a los servicios en línea para ciudadanos. Estos servicios son más eficientes porque permiten el acceso a la información pública desde cualquier ubicación, ya sea el hogar, oficina o lugares públicos, durante las 24 horas del día y utilizando diferentes medios y dispositivos de acceso. Este escenario requiere que los organismos públicos estén interconectados y el ciudadano pueda acceder a los servicios públicos por un único punto aún cuando sean provistos por diferentes departamentos, organismos o niveles del estado. Esto implica la existencia de un intercambio e integración efectiva entre las áreas involucradas. En muchos casos, la visión más común de integración es a través de los portales del gobierno, obviando que por

debajo de esta interfaz de integración, también conocida como “*front-office*” existe una distribución de tareas entre las áreas de gobierno, muy sofisticada en los niveles técnico y organizacional de la interoperabilidad. Este nivel de trabajo interno es muy complejo y se conoce como “*back-office*” (Lischewski y Scoll, 2006) y en el cual se ha avanzado mucho en los últimos años.

En este nivel de la interoperabilidad, es de destacar que en el Estado existe una compleja heterogeneidad organizacional. Por una parte, hay procesos estandarizados y completamente regulados por normativa. Por otra parte, no existen procesos estructurados de decisión, por lo que se requiere experiencia profesional y conocimiento especializado (Klischewski y Lenk, 2002). La estructura organizacional tiene muchos niveles de decisión para un mismo tema y a su vez, requiere una alta especialización tanto en su dimensión horizontal como vertical. Además, el modelo normativo no sólo regula los procesos de la administración sino también las comunicaciones dentro y fuera de su estructura.

Por otra parte, en los últimos años se ha producido una evolución en la web pasando de los documentos enlazados a un espacio donde se relacionan tanto documentos como datos, generando lo que se conoce como datos enlazados o “*linked data*” en inglés (Bizer y col, 2009). La integración e intercambio de información involucra entonces información estructurada que facilita la representación del conocimiento (Hoxha y Brahaj, 2011), constituyendo información que tiene alto valor agregado para la toma de decisiones de gobierno. Al hablar de conocimiento, surge inmediatamente el concepto de la semántica asociada a los datos en un determinado contexto para poder interpretar los mismos correctamente y tomar decisiones. Es justamente en este nivel de la interoperabilidad semántica donde la comunicación cobra verdadero sentido porque involucra la comprensión de los datos utilizados por dos partes que se comunican. Este nivel de comprensión es esencial en las comunicaciones del gobierno electrónico.

Como se muestra en la Figura 1, los principales componentes del gobierno electrónico son:

- **Cientes del gobierno electrónico.** Los clientes del gobierno electrónico son personas que utilizan las TICs para desarrollar sus propias tareas y mejorar los procesos que de ellos dependen. Son clientes externos e internos. Los clientes externos son los ciudadanos, empresas y otros gobiernos. Los clientes internos son los empleados públicos.
- **Actividades del Front-Office.** Las actividades del front-office representan la parte de un sistema de software que interactúa directamente con el usuario, por ejemplo, sitios de gobierno.
- **Actividades del Back-Office.** Las actividades del back-office están originadas en los sistemas transaccionales que dan soporte a las tareas administrativas y a la gestión de la información generada por un dominio (salud, educación, seguridad) que produce resultados para los clientes externos. Debido a los grandes volúmenes de información, si se desea lograr la integración e intercambio entre diferentes

sistemas, las actividades del back-office deben complementarse con una efectiva gestión del conocimiento.

Teniendo en cuenta estos clientes externos e internos, existen dos tipos de conocimiento que debe ser gestionado desde el back-office: uno requerido por los clientes internos para entender los datos y hacer un uso eficiente de los mismos, integrarlos e intercambiarlos y otro, para proveer un acceso más rápido, preciso y seguro a la información del Estado por parte de cualquier ciudadano.

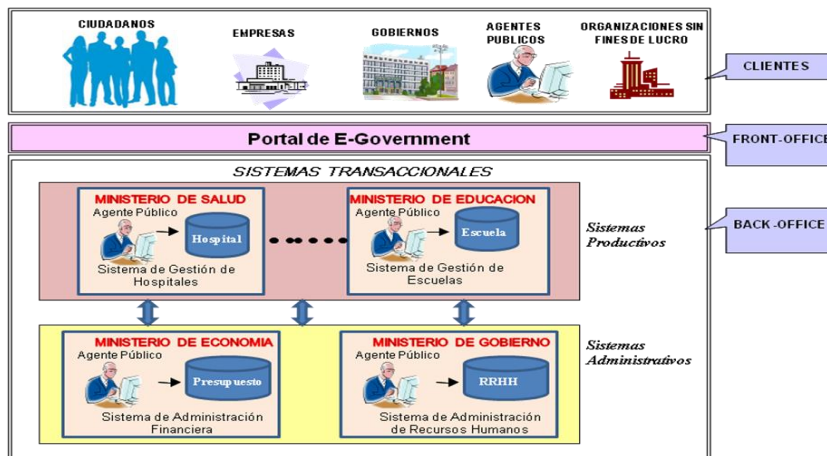


Figura 1 – Principales componentes del gobierno electrónico

Diversos trabajos sobre interoperabilidad semántica proponen un soporte en ontologías para describir servicios distribuidos a través de sitios de gobierno, focalizándose en la semántica específica de la información que se accede por los clientes externos (Álvarez Sabucedo y Anido Rifón, 2006), pero también existen trabajos que abordan el problema semántico desde la perspectiva del conocimiento que requieren los empleados públicos para mejorar el manejo de los procesos internos (Traunmüller y Wimmer, 2002). Ambos son importantes y deben considerarse desde una visión integral. El uso de ontologías de dominio, brinda una base para representar, adquirir y utilizar este conocimiento (Gómez-Pérez y col, 2004) y la visión integral, se propone a través de un modelo de gestión semántica que permita el intercambio e integración efectiva de la información.

2. RELEVANCIA DEL TRATAMIENTO DE LA INTEROPERABILIDAD SEMANTICA

El desarrollo de un modelo de gestión de la semántica asociada a la información pública es el punto de partida para fortalecer el concepto de transparencia en el Gobierno Abierto, especialmente si consideramos que la transparencia es “mostrar” (Calderón y col, 2010). Pero, ¿de qué sirve mostrar si lo que se muestra no es comprensible por la gran mayoría de personas que acceden a esa información?, ¿de qué sirven los datos si no se acompañan de la contextualización correspondiente? El Estado puede hacer accesible mucha información a través de sus portales y de hecho lo hace, pero a todos nos consta que el usuario común, el ciudadano,

muchas veces no comprende la terminología ni el contexto en el que esa información se despliega, por lo que no le resulta de utilidad y en muchas situaciones, se termina transformando en un arma de doble filo, pues desalienta al usuario, lo predispone mal frente a las TICs y pasando a ser un elemento que resta en lugar de sumar. Por ello, para realizar un proceso de apertura de datos realmente serio y útil, éstos deben ser veraces, completos y contextualizados (Concha y Naser, 2012).

Tal como manifiesta (Calderón y col, 2010) el Gobierno Abierto implica una evolución de nuestro sistema democrático de convivencia y valores basado en el establecimiento de mecanismos y espacios para la participación y colaboración de los ciudadanos, en el cual se aprovechan las nuevas tecnologías para garantizar estos nuevos canales de comunicación.

Por ejemplo, contar con información de salud o de un presupuesto de gobierno sin comprender acabadamente qué significa cada uno de los elementos que intervienen en su constitución ni comprender conceptos que se manejan en el contexto es, en la práctica, como si no se tuviera acceso a los mismos.

Como primer paso, existe una propuesta de utilizar Datos Enlazados Abiertos ("*Open Linked Data*") para comprender cómo se relacionan los datos, a qué conceptos y reglas de negocio se asocian, cómo se estructuran y cuáles son las jerarquías que rigen estos datos para poder comprender el contexto en el que operan. El objetivo de Open Linked Data es mejorar la información publicada para eliminar barreras en su consumo por los usuarios finales. Tim Berners-Lee, propone cinco niveles para hacer una apertura progresiva de los datos, ordenados desde su menor a mayor apertura y en cada nivel asigna una estrella que indica un grado de compromiso que facilita la automatización de la información publicada. Así, los cinco niveles para a Open Linked Data son:

Nivel	Descripción
*	Publicar datos en la Web en cualquier formato pero con una licencia libre para que sean datos abiertos
**	Publicar dichos datos en formatos estructurados legibles por máquinas (ej.: un documento MS excel en lugar de una imagen escaneada)
***	Utilizar formatos no propietarios (ej.: CVS en lugar de MS excel)
****	Utilizar estándares del consorcio web (W3C) para la web semántica (ej.: RDF y SPARQL) para identificar las "cosas" y para que puedan así apuntar a la información desde el exterior
*****	Enlazar tus datos con otros datos (de terceros) para proveer un contexto.

Lo siguiente, es considerar la utilización de ontologías como la herramienta tecnológica más completa para representar la semántica de datos en un determinado contexto y hacer uso de las herramientas estándares que permiten cumplir con el cuarto nivel de compromiso para la información. La contribución de las ontologías al problema semántico es altamente significativa debido a que contempla un grupo de componentes para la construcción de un modelo que permite ejecutar una variedad de funciones de inferencia y gestión del

conocimiento. También, una ontología formal tiene la capacidad de ser interpretada por herramientas automatizadas.

Por último, un modelo de gestión de la semántica permitirá enlazar con otros datos. De este modo, se cumple con el doble objetivo de que el ciudadano pueda comprender acabadamente la información que se publique y también, que se pueda procesar la misma en los casos que se requiera, logrando así que la información se encuentre en el último nivel de la categorización de Berners Lee.

Este modelo de gestión facilita el fortalecimiento de la transparencia que propone el Gobierno Abierto porque brinda la contextualización de los datos que se publiquen en el front-office, facilitando su completa interpretación. Por otra parte y pensando desde el back-office, la posibilidad de contar con datos estructurados, enlazados y procesables por las máquinas, incluso en su dimensión semántica, optimizan los procesos de integración e intercambio dentro de la administración de gobierno, fortaleciendo la eficiencia, tanto en el cumplimiento y en los tiempos como en la calidad de cada uno de los procesos que luego se reflejan en los portales de gobierno a través de los trámites y servicios al ciudadano.

3. GESTION DE LA SEMANTICA

3.1. ONTOLOGIAS DE DOMINIO

Una ontología define los términos básicos y relaciones que incluye el vocabulario de un tópico, tanto como las reglas que combinan esos términos y relaciones definiendo las extensiones del vocabulario (Neches y col, 1991). Las ontologías se focalizan en resolver problemas semánticos desde las siguientes visiones: (1) la ayuda semántica que las personas pueden obtener en el gobierno electrónico a través del procesamiento automatizado (Wimmer, 2003) y (2) el soporte tecnológico para facilitar la ubicación y procesamiento automático de los recursos de información en la provisión de los servicios de gobierno electrónico (Klischewski, 2003).

Los Datos Enlazados, mencionados anteriormente, usan Resource Description Framework (RDF) para describir datos en la web, los cuales pueden ser cualquier cosa en el mundo. RDF y Web Ontology Language (OWL) (Gómez-Pérez y col, 2004) son lenguajes de implementación de ontologías y fijan las bases para su desarrollo y correspondiente gestión del conocimiento.

La literatura propone diferentes clasificaciones de ontologías. Estas clasificaciones están basadas en el contenido y la generalidad de la ontología (Gómez-Pérez y col, 2004). Se introduce aquí una nueva clasificación basada en la motivación de la ontología:

1. Una Ontología originada por el requerimiento surge motivada en necesidades o requerimientos del usuario y se desarrolla para hacer explícita la información que se utiliza para los clientes externos del gobierno. Estas

ontologías se implementan en los portales de gobierno asociadas a los servicios que desde allí se ofrecen (Apostolou y col, 2005) (Stojanovic y col, 2004).

2. Una ontología de dominio surge motivada por la necesidad de describir la semántica de los datos disponibles en un dominio y su conocimiento asociado. Esta clase de ontología está orientada a hacer explícito el conocimiento implícito en los sistemas transaccionales utilizados por los clientes internos. De este modo, una ontología de dominio permite a los empleados públicos, gestionar el conocimiento necesario para lograr un entendimiento común, contribuyendo a la eficiencia de los procesos internos de las organizaciones.

3.2. MODELO PROPUESTO: SISTEMA BASADO EN ONTOLOGIAS

Como se muestra en la figura 2, proponemos un modelo que interactúa con los sistemas transaccionales de la administración de gobierno. Los *Sistemas Transaccionales* son aquellos que dan soporte informático a la labor diaria del estado e involucra tanto los sistemas productivos, tales como los relativos a salud, educación, seguridad, medio ambiente, etc., como a los sistemas administrativos, como por ejemplo, administración financiera, gestión de compras del Estado, recursos humanos, etc.

El modelo propuesto tiene diferentes componentes: un Portal Web Inteligente, un repositorio semántico y un área de creación y mantenimiento de las ontologías.

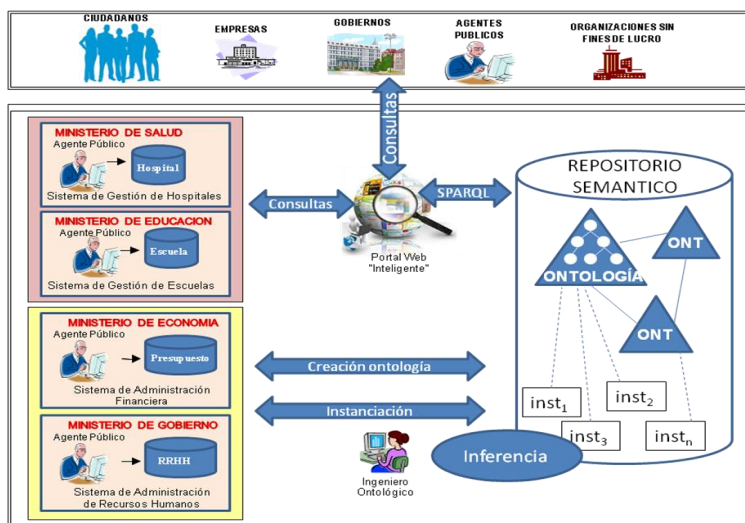


Figura 2 – Modelo para organizar la semántica en el back-office del gobierno electrónico

El Portal Web Inteligente tiene las características de que permite tanto a los clientes externos (ciudadanos, empresas, gobiernos, organizaciones sin fines de lucro) como a los clientes internos del Gobierno (empleados públicos y sistemas administrativos transaccionales) interactuar con el repositorio semántico de forma transparente y amigable.

En el área de creación y mantenimiento de las ontologías es donde trabajan los ingenieros ontológicos. Su trabajo consiste en crear las ontologías de dominio que correspondan, realizar el proceso de instanciación de las mismas, tomando los datos de los sistemas transaccionales que correspondan y realizar inferencia. Esta forma de trabajo garantiza que los clientes nunca acceden a los datos que gestionan los sistemas transaccionales, ya que éstos alimentan al repositorio semántico y es éste quien provee la información a los clientes. El repositorio semántico es un cliente externo de los sistemas transaccionales, aportando un nivel de seguridad a los datos que maneja el Estado, ya que la instanciación de las ontologías se realiza con las políticas de privacidad y acceso que el Estado determine. La instanciación puede realizarse de dos formas: manualmente, a través de una interfaz que permita insertar instancias dentro de la ontología o, de forma automática, poblando la ontología a través de métodos de extracción de recursos heterogéneos (Barrasa y col, 2004) (Stojanovic y col, 2002).

3.3. PROCESO PARA LA CONSTRUCCION DE ONTOLOGIAS DE DOMINIO

La construcción de ontologías es una tarea compleja en el ámbito gubernamental así como también para usuarios no expertos en desarrollo de ontologías. Para su construcción, se conocen dos grandes grupos de metodologías. El primer grupo es el de las metodologías basadas en la experiencia, tales como la Metodología de Grüninger y Fox (Grüninger y Fox, 1995) definida a partir del Proyecto TOVE y la Metodología de Uschold y King (Uschold, 1996) basada en la experiencia del desarrollo de la Ontología Enterprise. El segundo grupo, es el que propone un conjunto de actividades para el desarrollo de ontologías basado en un ciclo de vida de refinamiento de un prototipo, como es el caso de METHONTOLOGY, la cual se apoya en el estándar para desarrollo de software de la IEEE.

Al considerar el desarrollo de ontologías para el gobierno electrónico, se debe contemplar la complejidad y magnitud de las organizaciones gubernamentales y por esta razón se consideró apropiado trabajar con una metodología de prototipo evolutivo. Aunque METHONTOLOGY es la más indicada, las herramientas que propone, especialmente en las etapas de especificación y conceptualización, resultan insuficientes para representar la complejidad del Estado. Por esta razón, tomando dicha metodología como base, se agregaron herramientas para representar y modelar el conocimiento. Estas herramientas fueron tomadas de otras metodologías, tal como la Metodología de Desarrollo 101 (Noy y McGuinness, 2001). Por último, con el objetivo de completar los requerimientos de la ontología, se definieron algunas representaciones intermedias basadas en técnicas ya maduras de ingeniería de software, obteniendo un proceso que permite modelar dominios complejos, como es el caso particular de las administraciones gubernamentales.

El proceso que se expone a continuación, fue probado en el dominio de presupuesto del Gobierno de la

Provincia de Santa Fe, Argentina, ofreciendo resultados satisfactorios (Brusa y col, 2006). Aunque su aplicación debe ser extendida dentro de este dominio, se lograron importantes resultados y lecciones aprendidas.

El objetivo de un proceso de desarrollo de ontologías es establecer una estructura expresada en un lenguaje definido formalmente basado en técnicas de Inteligencia Artificial. El proceso se compone de tres subprocesos principales: Especificación, Conceptualización e Implementación, los cuales se describen a continuación.

3.3.1. Subproceso: Especificación

La meta de este subproceso es adquirir el conocimiento informal sobre el dominio y estimar el esfuerzo necesario para desarrollar la ontología de dominio. Las salidas son un Documento de Especificación de Requerimientos y un Documento de Estimación de Costos. Este subproceso tiene cuatro actividades (Figura 3).

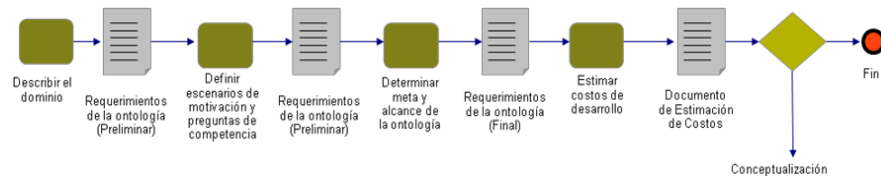


Figura 3. Actividades del Subproceso de Especificación

3.3.1.1. Describir el Dominio

En esta actividad se sugieren diferentes encuentros con expertos para adquirir el conocimiento sobre el dominio e identificar a los expertos del dominio como también los roles de cada uno de los involucrados.

3.3.1.2. Definir Escenarios de Motivación y Preguntas de Competencia

Con el objetivo de modelar el mundo real, Grüninger y Fox(1995) propone obtener requerimientos analizando diferentes escenarios y definir un modelo lógico informal de conocimiento. Para ello, sugiere definir escenarios de motivación y preguntas de competencia.

Los escenarios de motivación describen un conjunto de requerimientos que la ontología debería satisfacer después de haber sido implementada formalmente. También provee un conjunto de soluciones intuitivamente posibles para los problemas del escenario. Estas soluciones dan una primera idea informal de la semántica pretendida de los objetos y relaciones que más adelante se incluirán en la ontología. Este paso contempla plantillas para describir escenarios (Brusa y col, 2006) debido a que pueden existir una gran cantidad de contextos en un mismo dominio del sector público.

Por último, se definen las preguntas de competencia. Estas son preguntas en un nivel conceptual, escritas informalmente en lenguaje natural que la ontología deberá poder contestar (Grüninger y Fox, 1995). Las preguntas de competencia pueden agruparse en función de las entidades principales del dominio y se puede generar una jerarquía de las mismas.

3.3.1.3. Determinar la Meta y Alcance de la Ontología

La meta de la ontología es la definición de la forma en la cual la ontología será utilizada. Dependiendo de esta meta, se definirán sus elementos. El alcance limita la ontología, especificando qué debe ser incluido y qué no. Esto significa que un concepto no debería ser incluido si ninguna pregunta de competencia lo utiliza. Esta es una actividad muy importante en el subproceso de especificación ya que permite minimizar la cantidad de datos y conceptos que serán analizados, especialmente en el caso de la extensión y complejidad del sector público

3.3.1.4. Estimar los costos de desarrollo

La estimación de costos es la actividad de predecir la cantidad de esfuerzo requerido para construir una ontología. Esta tarea depende de las particularidades del proyecto como también del producto, personal y aspectos del proceso de desarrollo de la ontología.

En el sector público, esta actividad es muy importante debido al hecho de que la continuidad del proyecto puede depender de la aprobación de este documento por las autoridades correspondientes y a veces, su aprobación involucra también la elaboración de normativa.

3.3.2. Subproceso: Conceptualización

El objetivo de este subproceso es definir un Modelo Conceptual de Dominio organizando el conocimiento relevante adquirido en el subproceso previo y utilizando un lenguaje de modelado conceptual. Esta es la tarea más importante de la metodología propuesta y tiene que estar acordada con los expertos del dominio. Por ello, la utilización de representaciones gráficas es esencial para facilitar la comunicación entre los ingenieros ontológicos y los expertos. Así es que se pueden utilizar las técnicas de ingeniería de software que son familiares para los expertos del dominio, tal como el Lenguaje Unificado de Modelado (*UML*) (UML, 2006). Si bien UML no es el estándar más apropiado para una representación semántica, sus representaciones gráficas si son adecuadas para facilitar la comunicación entre las personas involucradas ("*stakeholders*") (Falbo, 2004).

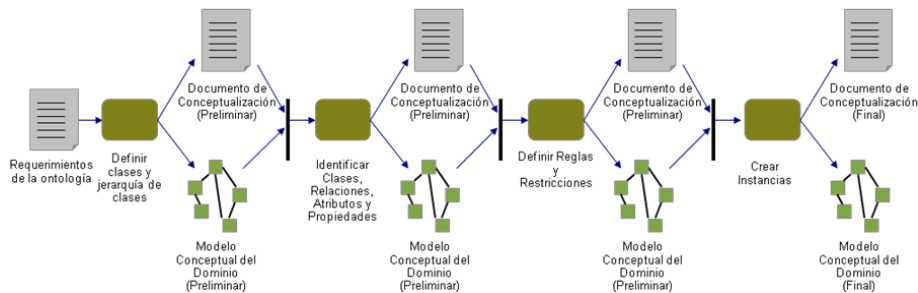


Figura 4. Actividades del Subproceso de Conceptualización

La Figura 4 muestra la representación esquemática del subproceso de conceptualización. La entrada es el

documento de especificación de requerimientos de la ontología definido en el subproceso previo. La salida es un modelo conceptual de dominio basado en diagramas de clase UML hasta el agregado de instancias.

3.3.2.1. Definir Clases y Jerarquías de Clases

La primera tarea aquí es identificar una lista de los términos más relevantes, utilizando la estrategia “*middle-out*” (Uschold, 1996), a partir del documento de especificación de requerimientos de la ontología. Inicialmente, es apropiado tomar una lista de términos sin considerar el solapamiento entre conceptos, propiedades o relaciones entre ellos y tampoco si los conceptos son clases o atributos (Noy y McGuinness, 2001).

A continuación, se representan los términos de dicha lista como una jerarquía de clases y serán jerarquizados o especializados, según corresponda. Proponemos representar la jerarquía de clases utilizando un diagrama de clases UML, definiendo una clase llamada “*Thing*” como la clase superior de la jerarquía. Si el diagrama se torna muy complejo y se dificulta su manejo, se podría separar en módulos usando paquetes. Esta separación puede ser muy útil más adelante, para decidir si se necesita definir una ontología de tareas para cada paquete. Una ontología de tarea define conceptos relacionados a una tarea específica (Gómez-Pérez y col, 2004). Finalmente, se identifican clases disjuntas, descomposiciones exhaustivas y particiones (Horridge y col, 2004) en el diagrama de clases UML.

3.3.2.2. Identificar Relaciones, Atributos y Propiedades

Para continuar, una tabla que refleje las relaciones bidireccionales puede ser elaborada asignando nombres con un criterio uniforme, identificando el dominio y el rango, cardinalidad y relaciones inversas.

Esta tabla de información se representa en un diagrama de clases UML y puede ser utilizado para comunicar el conocimiento adquirido a los expertos. Además, podría ser útil para comparar el modelo conceptual con un gráfico de la ontología generado por un software de edición de ontologías. Estudios empíricos prueban que las notaciones gráficas son más fáciles de comprender que otros formalismos, tal como los predicados lógicos.

3.3.2.3. Definir Reglas y Restricciones

Una restricción, en su uso general, es un tipo de especificación que establece una frontera finita (y en general, absoluta) definida para un tipo de proceso o función. En las ontologías, una restricción se refiere a una limitación impuesta porque la forma en que se han estructurado los conceptos. Por ejemplo, la cardinalidad o valores permitidos representan restricciones. Las restricciones que se pueden capturar a partir de la notación gráfica no deberían ser escritas explícitamente (Falbo, 2004). Por ello, las restricciones deberían definirse mientras se definen las relaciones entre clases, atributos y propiedades.

Por otra parte, una regla es una norma ampliamente aceptada, un concepto, verdad, definición o calificación en el dominio del discurso. En el dominio del sector público, las reglas están generalmente definidas de forma muy

clara en la normativa del Estado. Entonces, para definir las reglas, un ingeniero ontológico no sólo debe analizar las plantillas de descripción del escenario asociadas con los términos en consideración, sino también la normativa asociada a dichos escenarios. Las reglas deben ser explícitamente definidas utilizando un lenguaje formal, tal como la lógica de primer orden, apropiado para guiar al ingeniero ontológico en la definición de las restricciones de la ontología.

3.3.2.4. Crear Instancias

La última actividad del subproceso de conceptualización es crear instancias individuales en las clases. La definición de una instancia individual para una clase en el diagrama UML requiere (1) elegir la clase, (2) crear una instancia individual para esa clase y (3) completar los valores de los atributos. Para esta actividad, es útil analizar las plantillas de descripción del escenario y preguntas de competencia. Particularmente, una ontología de dominio/datos se poblará con instancias que se pueden extraer automáticamente o semi-automáticamente desde los sistemas transaccionales.

3.3.3. Subproceso: Implementación

La meta en este proceso es construir una ontología en un lenguaje procesable por las máquinas. Para ello, hay tres actividades principales: (1) Crear una ontología procesable, (2) Verificar la ontología y (3) Validar la ontología (Figura 5).

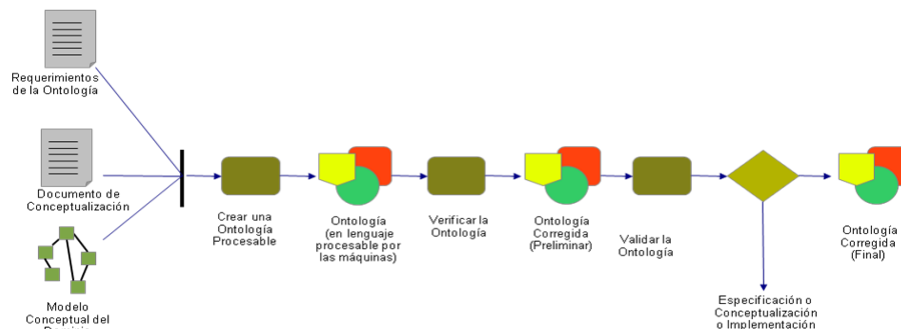


Figura 5. Actividades del subproceso de Implementación

3.3.3.1. Crear una Ontología Procesable

Crear una ontología procesable significa representar una ontología en un lenguaje procesable por las máquinas. Hay diferentes lenguajes pero los más relevantes son RDF y OWL. El primero fue desarrollado como un lenguaje basado en redes semánticas para describir recursos Web. El segundo, fue creado como un nuevo lenguaje de marcación para la Web Semántica.

El componente principal componente de la ontología es el conjunto de reglas. Considerando que el lenguaje

OWL es el estándar para implementar ontologías y no siempre es suficiente para realizar alguna deducción, se requiere combinar OWL con otro formalismo de representación, como son las reglas. Una forma de abordar este tema es utilizar el Lenguaje de Reglas de la Web Semántica (*SRWL- Semantic Web Rule Language*) el cual pretende ser el lenguaje estándar de reglas de la Web Semántica. Provee la capacidad de expresar reglas tipo cláusulas de Horn. SWRL permite a los usuarios escribir reglas que pueden ser expresadas en términos de conceptos OWL y que pueden razonar sobre individuos OWL (O'Connor y col, 2005).

Para extraer información desde ontologías OWL se requiere un lenguaje de consultas. El lenguaje más potente es SQWRL, el cual está basado en el lenguaje de reglas SWRL y utiliza su fuerte sustento semántico y su estructura formal. SQWRL provee un pequeño pero potente arreglo de operadores que permite a los usuarios construir consultas sobre ontologías OWL.

A continuación, se debe utilizar una herramienta de desarrollo de ontologías. En los últimos años, se ha desarrollado una nueva generación de entornos de ingeniería ontológica con el objetivo de integrar la tecnología de ontologías en los actuales sistemas de información. Entre estas herramientas se encuentra Protégé 4.0 (Horridge y col, 2004) y TopBraid Composer (Knublauch, 2006).

3.3.3.2. Verificar la Ontología

Para probar que la ontología es correcta, se debe verificar que sea consistente, concisa y completa.

- Consistente. Una definición dada es consistente si no se pueden inferir frases contradictorias a partir de otras definiciones y axiomas.
- Completa. No se puede probar que una ontología esté completa o sus definiciones sean completas pero sí se puede probar cuán incompleta es una definición individual y así deducir cuán incompleta está la ontología. También, se puede hablar de lo incompleto de una ontología si al menos se omite una definición atendiendo a un modelo de referencia establecido.
- Concisa. Una ontología es concisa si no almacena ninguna definición innecesaria o no utilizada, si no existen redundancias explícitas en las definiciones y si las redundancia no pueden ser inferidas utilizando otras definiciones y axiomas.

Se pueden identificar dos tipos de medidas para verificación: medidas estructurales y medidas formales. Las primeras se requieren para representar gráficamente la ontología. Las segundas implican utilizar un razonador.

3.3.3.3. Validar la Ontología

La validación de una ontología se refiere a si las definiciones de la ontología realmente modelan el mundo real para el cual fue creada la ontología (Gómez-Pérez y col, 2004). Entonces, validar la ontología significa verificar si la ontología cumple con lo establecido en los requerimientos en el documento de requerimientos de la

ontología.

El proceso de validación siempre debería realizarse contra un modelo de referencia. Este modelo podría estar dado por la especificación de requerimientos y las preguntas de competencia. La evaluación podría ser ejecutada de forma automática, si las preguntas de competencia se representan formalmente o de forma semi-automática, utilizando una heurística específica o el juicio humano.

Para formalizar las preguntas de competencia, se pueden utilizar los lenguajes de consulta RDQL (Seaborne, 2004) y OWL-QL (Fikes y col, 2003), Si es necesario realizar inferencias, es más apropiado utilizar OWL-QL.

En la validación de la ontología no sólo debe verificarse que las preguntas de competencia se respondan correctamente sino también se debe determinar si las preguntas de competencia plantean en realidad las preguntas correctas de acuerdo al propósito de la ontología (Falbo, 2004). Aquí, es esencial la comunicación con los expertos del dominio.

4. Conclusión y sugerencias

Este trabajo expone una forma útil de gestionar el conocimiento, elemento esencial en un Gobierno Abierto, presentando un modelo para facilitar la representación y difusión dentro del back-office. Es una herramienta simple que puede implementarse fácilmente, ya que como usa su propia base de instancias para la ontología, es menos resistida por parte de los agentes públicos. En segundo lugar, se expone un proceso de desarrollo de ontologías que muestra una forma práctica para completar el subproceso de conceptualización en el Estado permitiendo manejar su complejidad y poder ser utilizado por usuarios no expertos. Por último, se expone una nueva categorización de ontologías originada en la motivación de su construcción, lo cual ayuda al determinar su propósito y alcance.

Finalmente, se sugiere considerar las ideas presentadas en aquellos trabajos que se enfoquen en la filosofía del Gobierno Abierto, ya que permite trabajar a través de dos líneas interrelacionadas: (1) hacer explícito el conocimiento implícito y (2) sistematizarlo y difundirlo (Falivene y Silva, 2003). Esto es importante para hacer más transparentes, participativo y colaborativos los procesos del Estado.

5. Referencias

- Álvarez Sabucedo, L., Anido Rifón. L. (2006) "Semantic support for eGovernment transaction, Developing services for a better public service". eGovINTEROP Conference.
- Apostolou, D., Stojanovic, L., Pariente Lobo, T., Thoenssen, B. (2005)"Towards a Semantically-Driven Software Engineering Environment for eGovernment.".In: E-Government: Towards Electronic Democracy. LNCS 3416, Springer Verlag, pp.157-168.
- Barrasa, J., Corcho, O., Gómez-Pérez, A. (2004) "R2O, an extensible and semantically based database-to-

- ontology mapping language", 2nd Workshop on Semantic Web and Databases.
- Bizer, C., Heath, T., Berners-Lee, T. (2009), "Linked Data - The Story So Far", International Journal on Semantic Web and Information Systems (IJSWIS), 5(3), 1-22. doi:10.4018/jswis.2009081901.
- Brusa, G., Caliusco, Ma. L., Chiotti, O. (2006)"Building Ontology in Public Administration: A Case Study". 1st Intl Workshop on Applications and Business Aspects of the Semantic Web, pp.16-30.
- Calderón, C., Lorenzo, S. (2010). "Open Government: Gobierno Abierto". Alcalá la Real, España: Algón, Editores, ISBN 978-84-937218-5-5.
- Concha, G. y Naser, A. (Eds.) (2012) "El desafío hacia el gobierno abierto en la hora de la igualdad", Santiago de Chile: CEPAL/ECLAC. <http://www.cepal.org/ddpe/publicaciones/xml/9/46119/W465.pdf>.
- Falbo, R.A.. (2004), "Experiences in Using a Method for Building Domain Ontologies". 16th Intl Conf. on SEKE, Workshop on Ontology In Action pp.474-477.
- Falivene G., Silva G., La formación de directivos para la Gestión del Conocimiento en las organizaciones públicas, Segundo Congreso Argentino de Administración Pública, Sociedad, Gobierno y Administración Pública -" Reconstruyendo la estatalidad: Transición, instituciones y gobernabilidad ", 2003, Córdoba.
- Fikes, R., Hayes, P., Horrocks, I. (2003), "OWL-QL-A Language for Deductive Query Answering on the Semantic Web", KL Laboratory, Stanford University.
- Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004), "Ontological engineering".(2nd ed.). London: Springer Verlag.
- Grüniger, M., Fox, M. S. (1995), "Methodology for the Design and Evaluation of Ontologies".IJCAI Workshop on Basic Ontological in Knowledge Sharing.
- Harrison, T., Guerrero, S., Burke, B., Cook, M., Cresswell, A., Helbig, N., Hrdinova, J., Pardo, T. (2011)."Open Government and E-Government: Democratic Challenges from a Public Value Perspective". 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times (dg.o'11), ACM, New York, USA, pp.245-253.
- Horrige, M., Knublauch, H., Rector, A., Stevens,R., Wroe,C. (2004)"A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0", The University Of Manchester, Manchester, UK.
- Hoxha, J.; Brahaj, A. (2011) "Open Government Data on the Web: A Semantic Approach, Emerging Intelligent Data and Web Technologies" (EIDWT), International Conference on Digital Object Identifier:10.1109/EIDWT.2011.24, pp.107-113.
- Klischewski, R., Lenk, K. (2002), "Understanding and Modelling Flexibility in Administrative Processes", EGOV pp.129-136.

- Klischewski, R. (2003)"Semantic Web for e-Government". In: Traunmuller, R.(ed.), EGOV, LNCS No. 2739, Springer Verlag, pp.288-295.
- Klischewski, R, Scholl, HJ. (2006), "Information Quality as a Common Ground for Key Players in e-Government Integration and Interoperability", HICSS.
- Knublauch, H.(2006) "RDF a support in TopBraid Composer", Available: <http://lists.w3.org/Archives/Public/public-rdf-in-xhtml-tf/2006Sep/0019.html>.
- Neches, R, Fikes, R.E, Rice, J, Wilder, S, Senator, T, Swartout, W. (1991)"Enabling technology for knowledge sharing", AI Magazine 12(3) pp.36-56.
- Noy, N., McGuinness, D. (2001)"Ontology Development 101: A Guide to Creating Your First Ontology". Technical Report SMI-2001-0880, Stanford Medical Informatics, Stanford University, Stanford, USA.
- O'Connor, M., Knublauch, H., Tu, S., Musen, M. (2005)"Writing Rules for the Semantic Web Using SWRL and Jess". 8th International Protege Conference, Protege with Rules Workshop.
- Seaborne, A. (2004), "RDQL - A Query Language for RDF, W3C Submission". <http://www.w3.org/Submission/2004/>
- Stojanovic, L., Stojanovic, N., Volz, R. (2002), "Migrating data-intensive Web Sites into the Semantic Web". In: Symposium on Applied Computing.
- Stojanovic, L., Abecker, A., Stojanovic, N., Studer, R.(2004)"An Approach for the Change Management in the E-Government Domain", 2nd International Conference on Knowledge Economy and Development of Science and Technology.
- Traunmüller, R., Wimmer, M. (2002)"KM for Public Administration: Focusing on KMS Feature Requirement, Practical Aspects of Knowledge Management".LNAI 2569, Springer-Verlag, pp.314-325.
- UML (2006)"Unified Modeling Language" <http://www.uml.org/>
- Uschold, M. (1996)"Building Ontologies: Towards a Unified Methodology". In: 16th Annual Conference of the British Computer Society Specialists Group on Expert Systems.
- Wimmer, M.(2003)"Knowledge Management in e-Government". In: Wimmer, M. (ed.).4th IFIP International Working Conf, KMGov. LNCS No. 2645, Springer Verlag, pp.26-28.