

Metadades de dades obertes: estudi de cas a l'Uruguai

[Versión castellana]

SILVANA TEMESIO
Professora de l'Institut de Informació
Facultad de Información y Comunicación
Universidad de la República de Uruguay
silvana.temesio@fic.edu.uy

Resum

Objectiu: analitzar les metadades de dades obertes en voga des de la perspectiva de la seva adequació als objectius de recuperació i ús.

Metodologia: es planteja l'extensió de Data Catalog Vocabulary (DCAT) a una nova classe que modela el nivell de la dada i la seva semàntica. L'extensió s'analitza en un cas de dades obertes específic per a un procés de convenis de camineria rural a l'Uruguai.

Resultats: el desenvolupament s'implementa en la publicació d'un conjunt de dades obertes en el marc de descripció de recursos (RDF) al portal de dades obertes de l'Uruguai que n'exemplifica la utilitat.

Resumen

Objetivo: analizar los metadatos de datos abiertos en boga desde la perspectiva de su adecuación a los objetivos de recuperación y uso.

Metodología: se plantea la extensión de Data Catalog Vocabulary (DCAT) a una nueva clase que modele el nivel del dato y su semántica. La extensión se analiza en un caso de datos abiertos específico para un proceso de convenios de caminería rural en Uruguay.

Resultados: el desarrollo se implementa en la publicación de un conjunto de datos abiertos en el marco de descripción de recursos (RDF) en el portal de datos abiertos de Uruguay que ejemplifica la utilidad.

Abstract

This paper analyses metadata for the growing trend in open data, from the perspective of whether they meet the objectives of data recovery and use. We propose an extension of the Data Catalog Vocabulary (DCAT) to a new class that would model the level of the data and its semantics. The extension is analyzed in the case of specific open data for a process of agreements on rural road networks in Uruguay. The utility of the development is illustrated by its implementation in a set of open data in resource description framework (RDF), published on Uruguay's open data portal.

Rebut: 25/11/2014. Acceptat: 26/02/2015.

1 Introducció

Les aplicacions informàtiques en una organització governamental són les principals encarregades de gestionar les dades crues, que són dades font o primàries recopilades en els processos de l'organització i en les seves activitats. Aquestes dades són en general atòmiques, no estan agrupades. D'acord amb la diversitat de procediments que les involucren, són de dominis variats: textual, numèric, visual, geogràfic, etc. i es presenten en formats diferents. Si bé es pot usar el terme *crues* per a dades no validades provinents de dispositius de recollida de dades electròniques, en aquest document s'utilitza el terme en el sentit de dades primàries i atòmiques.

Aquestes dades crues de les institucions públiques normalment s'emmagatzemen en bases de dades, separades en esquemes i als quals diferents aplicacions accedeixen d'acord amb un sistema de permisos, rols i privilegis. El principi de les bases de dades relacionals de modelar les dades sense redundància opera de manera que cada categoria de dades té un esquema en què hi ha un responsable de l'actualització de la dada que proporciona drets de lectura a tots els qui la fan servir. D'aquesta manera s'utilitzen les dades dins de l'organització, un escriptor responsable de la dada o diversos responsables de diferents facetes de la dada i molts lectors consumidors. Una pràctica recomanable és que hi hagi consumidors de dades que hi accedeixin per mitjà d'un servei web, una prestació especial que ofereix les dades que s'utilitzaran en altres aplicacions informàtiques. D'aquesta manera es modela l'acoblament de forma senzilla.

Les aplicacions informàtiques que utilitzen bases de dades no són les úniques que gestionen dades crues, però sí que són les que ho fan de manera sistemàtica i global amb traçabilitat i seguretat. Són les que normalment registren les activitats que es fan dins de l'organització i, en aquest sentit, són una font d'informació auditable de les tasques i els procediments que porta a terme una institució per mitjà de les seves dades (Temesio, 2013).

El costum de treballar les dades aïlladament, sense connexió amb altres activitats relacionades, coincidint accions, sense alimentar un repositori global de dades —la base de dades— és un costum que, si bé encara persisteix en algunes tasques, cada vegada es desterra més com a modalitat de treball en els organismes públics amb l'adveniment de les bones pràctiques de l'administració electrònica.

Dins dels organismes públics, d'acord amb la llei d'accés a la informació pública —que té diverses expressions molt similars a Iberoamèrica; en aquest cas es refereix a la llei de l'Uruguai—, s'ha d'establir si la informació és pública o no.⁴ Si la informació és pública, es pot publicar com a dada oberta; en cas contrari, serà accessible per a les persones que designi el propietari de les dades —normalment qui ocupi un lloc d'alt nivell jeràrquic— amb la finalitat de fer tasques específiques.

Publicar les dades obertes persegueix les finalitats de l'administració electrònica de donar accés a qui pugui i vulgui reutilitzar les dades, crear aplicacions que siguin d'interès i transparentar l'Administració (OECD, 2014).

Per aconseguir aquests fins, la dada ha d'arribar a qui la cerca, de manera que el descobriment i la recerca han de ser operacions clares, senzilles i comprensibles. En aquest sentit pot pensar-se en les dades obertes amb un abordatge des de les ciències de la informació. Una font de dades obertes, o conjunt de dades obertes, és un paquet d'informació que requereix explicitar-se i que es pot catalogar amb un estàndard de metadades general com el Dublin Core. D'aquesta manera, el conjunt de dades proveeix informació sobre si mateix, sobre el seu origen, sobre els aspectes administratius i el seu contingut. El conjunt de dades és un objecte digital que pot assimilar-se a un llibre i, per tant, les metadades aportades tenen una analogia amb la descripció bibliogràfica. Es podria assenyalar una altra analogia amb l'arxivologia veient la dada com una peça documental simple o composta i el seu registre com una descripció arxivística. Aquest símil, encara que forçat, serveix per donar suport a la idea que aquest tipus d'objecte digital, igual que en els exemples, compleix les necessitats de qualsevol recurs d'informació que els professionals de la informació gestionin: ser buscat i trobat d'acord amb criteris de descripció que permetin ràpidament identificar si compleix les necessitats de la cerca.

La qualitat en les dades obertes, tant en el contingut com en la descripció, contribueix a potenciar el principi de transparència, ja que facilita una participació ciutadana per mitjà de criteris de rellevància, usabilitat i qualitat. La usabilitat de les dades obertes s'expressa en la capacitat d'autoexplicar-se. Com més detalladament es descriu, més gran és la probabilitat de ser descobertes, analitzades i reusades, ja sigui en aplicacions que ofereixin un servei que no presti l'administració, en l'anàlisi d'investigacions periodístiques o com a entrada per a la decisió de portar a terme una intervenció tant des de l'àmbit de l'administració com de la societat civil o de les empreses.

2 Metadades en dades obertes

Es va dur a terme una investigació de les normes usades en la descripció de dades i experiències en la regió. Es van revisar els portals de dades obertes següents:

- L'Uruguai <<http://datos.gub.uy/> (<http://datos.gub.uy/>) >
- Xile <<http://datos.gob.cl/> (<http://datos.gob.cl/>) >
- Espanya <<http://datos.gob.es/> (<http://datos.gob.es/>) >, <<https://www.zaragoza.es/ciudad/risp/>>
- Colòmbia <<http://datosabiertocolombia.cloudapp.net/frm/buscador/frmBuscador.aspx> (<http://datosabiertocolombia.cloudapp.net/frm/buscador/frmBuscador.aspx>) >
- L'Argentina <<http://datospublicos.gov.ar/> (<http://datospublicos.gov.ar/>) >
- El Brasil <<http://dados.gov.br/> (<http://dados.gov.br/>) >
- El Perú <<http://lima.datosabiertos.pe/home/> (<http://lima.datosabiertos.pe/home/>) >

El CKAN (<http://ckan.org/>) és una aplicació de programari lliure molt usada en els portals de dades obertes a l'Amèrica Llatina. Permet l'ingrés de col·leccions de dades i les seves metadades per mitjà del permís i rol específic, després exposa aquests conjunts de dades i en permet la cerca a través de les metadades. Si la cerca dona resultats d'interès, llavors l'usuari té la possibilitat de descarregar el conjunt de dades. D'alguna manera, l'aplicació és com un repositori, però aquí el que s'introdueix són conjunts de dades en lloc de recursos digitals.

El CKAN treballa amb una sèrie de metadades, algunes d'obligatòries i altres d'opcionals: títol descriptiu, freqüència d'actualització, cobertura espacial, sistema de referència usat, cobertura temporal, descripció de les dades, etiquetes, llicència, organisme, visibilitat, i a més informació addicional, com el responsable de les dades i el responsable de l'actualització en el catàleg d'acord amb la freqüència assenyalada.

Hi ha una proposta de metadades específiques, el Data Catalog Vocabulary (DCAT), que té el seu origen en una iniciativa de normalització del World Wide Web Consortium (W3C) i que ha tingut en compte altres conjunts de metadades relacionades establint la necessitat de l'ús de taxonomies per als ítems que puguin normalitzar-se. El DCAT té un nivell de descripció quant a portal i un altre nivell pel que fa al conjunt de dades.

El DCAT, d'acord amb la seva especificació (DCAT, 2014), "es un vocabulari RDF dissenyat per facilitar la interoperabilitat entre catàlegs de dades publicats en la web".

El propòsit exposat és que en descriure conjunts de dades amb el DCAT es faciliti el descobriment i l'ús de les aplicacions en el consum de dades de múltiples catàlegs.

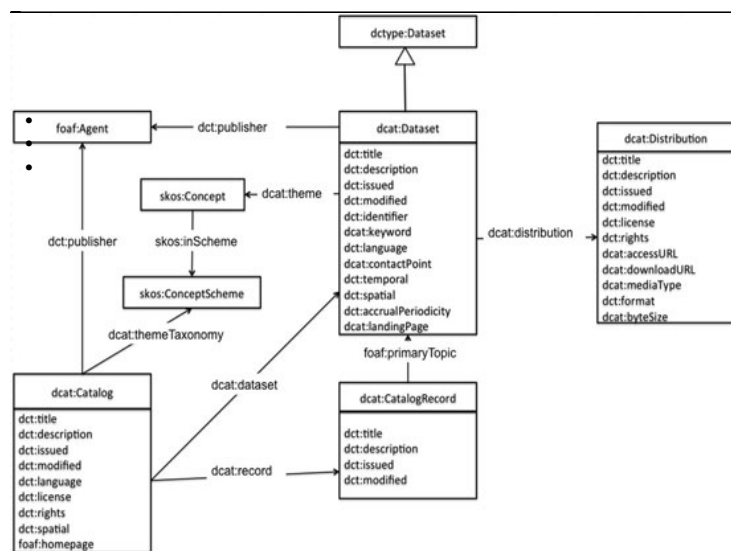


Figura 1. Diagrama UML de l'especificació DCAT (<http://www.w3.org/TR/vocab-dcat/>)

Tenim les classes següents:

- Catàleg ("dcat:CatalogRecord")
- Conjunt de dades ("dcat:Dataset")
- Distribució ("dcat:Distribution")

El conjunt de dades pot estar en diversos formats o estar disponible per mitjà d'una aplicació web (API). En aquest cas l'aplicació seria una instància de "dcat:Distribution".

El catàleg té múltiples conjunts de dades i cada un d'ells és una entrada que és descrita per la classe "dcat:CatalogRecord"; en canvi, "dcat:Dataset" descriu el conjunt de dades específicament.

Aquest conjunt de dades té un tema o diversos temes associats que es registren en un sistema taxonòmic, dins el qual corresponen a un concepte.

Així mateix, té un responsable de publicació que es correspon amb un valor en l'esquema FOAF (FOAF,

2000), amb la qual cosa queda normalitzada la institució o la persona que publica —una mena d'identificació d'autor o normalització del nom del productor, com a analogia en els àmbits bibliològic o arxivístic—, fent un acoblament amb un valor en el FOAF.

El DCAT fa servir termes d'altres estàndards i vocabularis com ara el Dublin Core (Dublin Core Metadata Initiative), l'SKOS (W3C SKOS, 2009) o el FOAF, amb la qual cosa aprofita i reutilitza els conceptes existents i es concentra en el disseny específic de catàlegs i conjunts de dades obertes.

L'ús de l'RDF (W3C RDF) no és prescriptiu, encara que si s'utilitza permet consultes semàntiques a través d'*endpoints* SPARQL (W3C SPARQL, 2008). Aquests *endpoints* són punts de recerca en un repositori de dades que en ser emmagatzemats en RDF tenen semàntica i poden ser interrogats amb un llenguatge de consulta —SPARQL— que permet buscar en un graf de triples RDF.

El CKAN permet l'exportació de registres en format DCAT i també permet recol·lectar registres d'altres catàlegs en aquest format. És possible que el DCAT pugui constituir-se en el format d'intercanvi de metadades de la mateixa manera que el Dublin Core és el conjunt de metadades de recol·lecció amb el protocol OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting).

Una altra iniciativa destacable en aquest àmbit és el VoID (W3C Describing Linked Datasets with the Void Vocabulary, 2011), que es pot usar en conjunció amb el DCAT. El VoID és un vocabulari RDF per expressar metadades dels conjunts RDF amb l'objectiu d'intervenir entre els qui publiquen les dades i els qui les usen, des dels processos de descobriment, catalogació i arxiu. Es basa en el Dublin Core.

Un aspecte de gran interès del VoID és que permet descriure els enllaços RDF entre diferents conjunts de dades, i permet vincular conjunts de dades diferents. Aquests enllaços RDF habiliten la navegació entre conjunts de dades diferents però lligats semànticament. Normalment fan servir el predicat "owl:sameAs", que vincula conceptes iguals amb noms diferents.

El VoID modela les matèries sobre les quals versen els conjunts de dades per mitjà de la propietat "dcterms:subject". Cada conjunt RDF utilitza un vocabulari o més d'un o ontologies OWL (W3C OWL, 2012).

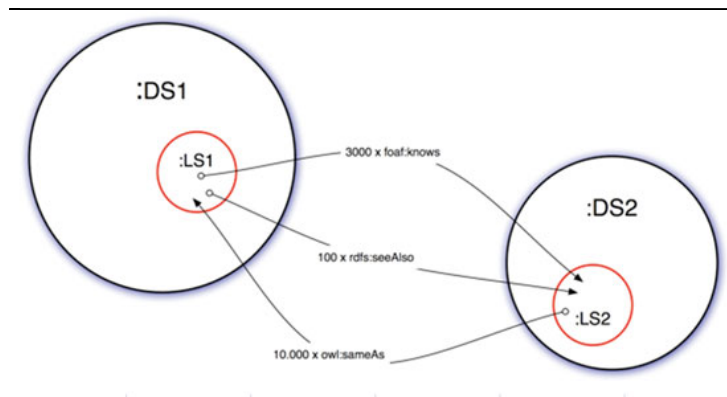


Figura 2. Vinculació en el VoID (<http://semanticweb.org/wiki/VoID>)

En la figura 2 es mostren dos conjunts de dades (DS1, DS2) que estan en RDF i es vinculen a través dels vincles (LS1, LS2). El vincle es fa a través d'una subclasse del VoID ("void:linkset"). En cada tripla (subjecte, predicat, objecte) de la vinculació, el subjecte és un recurs allotjat en un conjunt de dades (p. ex. DS1) i l'objecte, un recurs d'un altre conjunt (p. ex. DS2).

El VoID llavors permet establir les concordances entre conjunts de dades en esquemes diversos. Un exemple d'ús del DCAT i VoID és el cas de les dades obertes de Saragossa on així ho expliquen.²

L'ús d'ontologies en la descripció de dades obertes és una pràctica que s'implementa al portal de dades

obertes brasiler i constitueix una instància de la preocupació de normalitzar i consensuar els aspectes semàntics dels conjunts de dades que és molt important. En el cas específic de dades de pressupost federal brasiler, es va trobar una ontologia en OWL poblada amb dades pressupostàries i un *endpoint* SPARQL per a consultes.³

Al portal de dades obertes de Saragossa també es troba un *endpoint* SPARQL per a consultes que permet consultar sobre les dades enllaçades RDF.

L'ús d'ontologies permet no només relacionar conceptes i conjunts de dades, sinó establir un idioma comú consensuat que permeti la interoperabilitat semàntica en la vinculació.

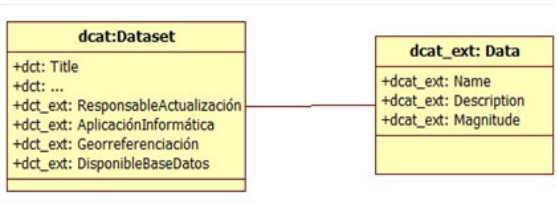
3 Usabilitat semàntica de les dades obertes

En els diversos casos analitzats hi ha aspectes que no es tenen en compte, com ara la semàntica de les dades, que es considera fonamental per a la interpretació i el correcte ús de les dades. Per tant, es planteja incorporar aquests aspectes en el conjunt de metadades que acompanyen les dades obertes.

L'enfocament es pot veure al portal de dades obertes de l'Uruguai.⁴ El Ministerio de Transportes y Obras Públicas (MTO) hi va publicar un conjunt de dades en què proposa una sèrie de metadades basades en Dublin Core, fent algunes extensions:⁵

- Qui és el responsable de l'actualització de les dades. Si bé en el DCAT hi ha la freqüència d'actualització ("AccrualPeriodicity") i pot inferir-se que l'organització que publica és la responsable de l'actualització, s'entén que l'explicitació ajuda a la localització del contacte específic, cosa que de vegades no és una tasca senzilla.
- Si les dades provenen d'una aplicació informàtica, consignar en aquest cas quina és l'aplicació. L'aplicació és una font sobre el procediment o activitat desenvolupats per produir la dada i dóna comprensió i context per poder interpretar adequadament el valor. El DCAT indica que l'API es pot definir com una instància de la distribució però sense arribar a la definició d'una classe específica. No obstant això, es considera que l'especificació expressa constitueix un element de traçabilitat i qualitat de la dada que mereix assenyalar-se específicament.
- Si les dades estan georeferenciades. Si la dada està georeferenciada és possible disposar-ne com a capa en un mapa. Quan el que es busca és aquest tipus de visualització, la indagació en cada conjunt de dades per determinar si disposa de les coordenades geogràfiques és una tasca que es pot simplificar. L'explicitació d'aquesta metadada permetria cercar ràpidament la condició, la qual cosa és de gran utilitat i es presumeix que tindrà un ús incremental en la mesura en què la territorialització de les dades s'incrementa.
- Si les dades estan disponibles en una base de dades. Aquest fet indica un nivell de tractament sistemàtic i la possible activitat coordinada de múltiples actors que treballen sense redundància i amb consistència. Si bé es considera si les dades es presenten mitjançant una aplicació estaran emmagatzemades en una base de dades, pot passar que el format de presentació no sigui a través d'una sortida de l'aplicació i en aquest cas no sabrà si les dades provenen o no d'una base de dades.
- Descripció granular de les dades en si i la seva estructuració. Seria una classe nova, ja que no s'aplicaria al conjunt de dades. Per als aspectes d'usabilitat i interoperabilitat de la dada aquest tercer nivell de descripció que correspon no ja al conjunt sinó a cada element de dades del conjunt permet entendre què significa cada dada i com està estructurada.

Figura 3. Extensió proposada per al DCAT



El conjunt de dades és un recurs format per elements atòmics. Està constituït per atributs, i aquests atributs a la vegada han d'identificar-se perquè el conjunt sigui comprensible i usable en termes acceptables. Per tant, en el cas d'un recurs d'aquest tipus, es proposa una especificació explícita sobre cadascun dels atributs i també sobre la seva estructuració.

La dada té un nom generalment autoexplicatiu, però per treballar amb dades no n'hi ha prou amb un nom. S'ha de donar una descripció clara del que signifiquen els valors que conté perquè puguin desenvolupar-se aplicacions que usin la dada i que donin resultats correctes.

Si prenem com a exemple el nom del camp "estat de la ruta" i veiem el seu domini de valors (bo, regular o dolent), no ens indica res sobre quins indicadors es van tenir en compte per construir aquests valors. Darrere del valor hi ha una definició de treball que dirà que la ruta es considera en bon estat si té un valor en un rang determinat de l'índex d'estat superficial. Aquesta definició ("dcat_ext:Description") serà necessària per poder generar una aplicació que mostri els estats de les carreteres a Llatinoamèrica perquè s'haurà de tenir en compte que els valors que es posin del Brasil siguin anàlegs als de l'Uruguai, per exemple.

D'altra banda, el domini de valors ha de quedar clar quan no està indicat d'una altra manera; per exemple, un atribut "distància" es pot expressar en metres, quilòmetres, etc. Si el nom de l'atribut té per valor "suma en dòlars" no cal expressar la magnitud, però un atribut "suma" és ambigu.

L'altre aspecte de gran interès per a la reutilització de les dades és la seva estructuració, la manera com es vinculen uns atributs amb els altres. Normalment les dades que s'emmagatzemen en una base de dades relacional estan normalitzades i s'associen per mitjà de relacions. La millor manera de veure aquesta estructura és un diagrama d'entitat-relació.

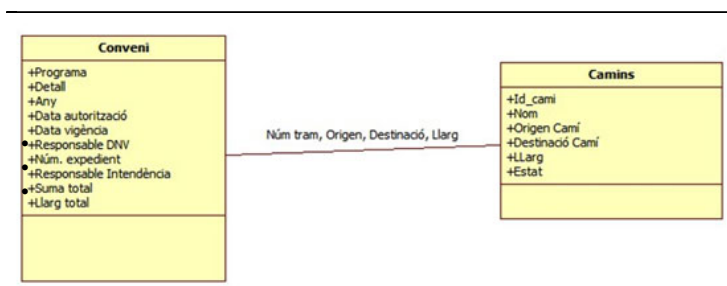


Figura 4. Diagrama d'entitat-relació conjunt de dades de conveni

El diagrama visibilitza que hi ha dos elements i un vincle:

- Els convenis amb les seves dades
- Els camins amb les seves dades
- Els camins assignats a cada conveni amb les porcions respectives (no tot el camí pot involucrar el conveni)

El format RDF visibilitza aquestes relacions i per això un conjunt de dades "rdfitzat" constitueix un major nivell de qualitat de les dades.

En el cas plantejat de les dades de convenis del Ministerio de Transportes y Obras Públicas, els atributs s'agrupen amb un criteri horitzontal en el conjunt de dades, però darrere d'aquesta agrupació hi ha un esquema de vinculació que ha de ser explícit per tal de ser comprensible. En presentar les dades en format RDF, s'aconsegueix una visió sobre l'estructuració interna de les dades. És així com s'aprecia la vinculació d'un conveni amb els elements del conveni en el format RDF, tal com es detalla en l'apèndix.

L'RDF modela el vincle en què cada conveni té associats determinats camins. També es pot veure mitjançant un diagrama de grafs.

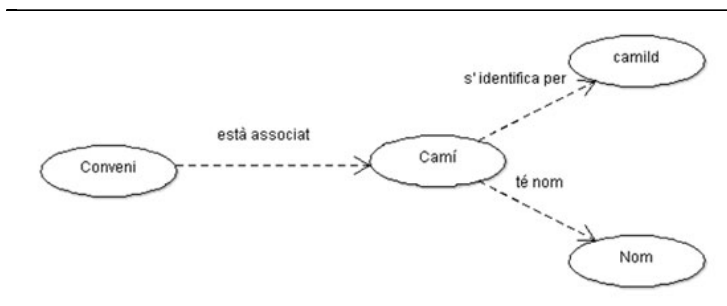


Figura 5. Graf de conveni

Si a més enllacem les dades amb conjunts de dades ja publicades, estem creant una autopista per navegar a la xarxa semàntica. En el cas d'aquest conjunt de dades això encara no s'ha fet, però, per exemple, si l'atribut que identifica el camí pogués enllaçar amb un conjunt de dades consensuat sobre la identificació dels camins, tots els que es referissin als camins podrien reutilitzar-se i d'aquesta manera s'aconseguiria la vinculació de dades amb diversos actors —el Ministerio de Transportes, la Intendencia Departamental— i amb facetes diferents, informació sobre convenis, informació departamental, informació

sobre obres dutes a terme en els camins, informació sobre avaluació de l'estat dels camins, etc. En tots els casos esmentats s'utilitzaria la identificació del camí d'aquesta font consensuada, diguem el "FOAF" del camí, i podríem navegar semànticament sobre aspectes totalment diversos, sabent que es tracta d'aquest camí i que tots l'anomenem igual.

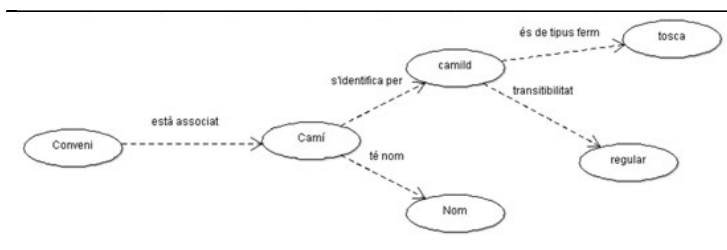


Figura 6. Graf de conveni estès

No obstant això, l'aspecte de comprensibilitat de la dada en si no es preveu per aquesta via, sinó a través de definir el significat de la dada, el diccionari de dades de la dada. Els dos aspectes són fonamentals: l'explicitació de la dada i la seva estructuració.

El diccionari de dades de les dades té dos aspectes:

La identificació del conjunt de valors sobre el qual mapa. Per exemple, la transitabilitat (bona, regular,

dolenta).

- La definició de treball de l'atribut que ha de permetre establir el valor, d'acord amb un criteri conegut que pugui ser entès i comparat; per exemple, la definició de *transitabilitat*.

Els portals de dades obertes funcionen com a dipòsits d'aquests objectes especials que són les dades i tenen, d'una banda, les mateixes necessitats d'un repositori digital de tipus genèric per donar solucions apropiades a la cerca d'informació. De l'altra, l'especificitat dels objectes emmagatzemats en aquests dipòsits requereix metadades granulars per donar un ús a les dades de forma consistent i amb qualitat. Podríem anomenar aquestes metadades de *tercer nivell* si considerem les metadades del portal de dades com un primer nivell i les metadades generals del conjunt de dades com un segon nivell; llavors les metadades dels atributs constituïrien un tercer nivell. Aquestes metadades de tercer nivell han

de tenir un criteri de normalització o establir la informació bàsica requerida. Són clau per a la reutilització de les dades de forma correcta, no operen en el nivell de recuperació d'informació de manera directa però són essencials per a la reutilització adequada, la qual cosa, en definitiva, és una operació seqüencial en la cerca i recuperació i és un objectiu primordial de la visibilització d'aquest tipus d'informació.

4 Conclusions

Les dades d'una organització són un element més en un entramat d'altres objectes d'informació. Perquè la informació d'una organització pugui acoblar-se i interoperar cal estandarditzar, sistematitzar i documentar clarament aquestes dades en la gestió interna de l'organització. Aquesta tasca també s'ha de dur a terme cap a fora, interactuant amb altres organismes governamentals i permetent el flux d'informació interinstitucional. Un nivell superior és l'ús de les dades de l'administració per part de la ciutadania, els periodistes o altres actors de la societat civil (*Ciudadanía/Cidadania 2.0*).

Les dades requereixen un nivell de precisió més gran que altres recursos d'informació, respecte al significat, el domini de valors i la forma com s'estructuren per poder-se utilitzar correctament i senzillament.

La integració de dades pot fer-se amb més qualitat i senzillesa si s'aporten les metadades proposades en el moment de la publicació. El productor de les dades coneix aquests aspectes, per la qual cosa s'estima que no pressuposa un gran esforç aportar-los en el moment de la publicació, i aquest esforç té un impacte molt significatiu en la qualitat i llegibilitat de les dades i, per tant, una major adequació en la reutilització.

L'àmbit de l'administració electrònica permet una recomanació institucional en aquest sentit, perquè és interessant poder navegar per la informació tant dins de la institució com entre institucions.

En el cas implementat de les dades dels convenis, s'estableix una identificació del camí rural per mitjà d'una URI (cadena de caràcters curta) que permet que altres actors, com ara les intendències departamentals, aportin una altra informació i vagin construint un graf navegable que modela l'articulació de la informació entre diferents institucions.

Aquest primer nivell de connexió i estandardització permet aventurar la construcció d'ontologies institucionals que puguin anar-se vinculant. Quan una intendència departamental usi la identificació del camí, pot establir que es tracta de la mateixa identificació que utilitza el Ministerio de Transportes ("owl:sameAs").

Un altre aspecte que s'ha de tenir en compte en els conjunts de dades és la necessitat de tenir una política de preservació digital, i en aquest sentit les metadades llancen informació sobre els formats, però també modelen un encapsulament descriptiu de la semàntica mateixa.

La informació de l'administració electrònica forma part d'un sistema més ampli que creix, s'influència i canvia constantment. Aquest vast ecosistema pot navegar a través dels fluxos i atracar a les illes o dipòsits que es troben al llarg del trajecte, però cal construir una brúixola precisa que marqui rumbos segurs. Aquesta brúixola és una combinació d'elements: metadades que descriuigui correctament, terminologia que permeti comprendre i acoblar, tesaurus, ontologies, grafs RDF i dades enllaçades.

Els professionals de la informació tenen la visió i la formació per participar en equips de treball multidisciplinaris que modelin corrents segurs en l'àmbit de l'administració electrònica.

Bibliografia

- Ciudadanía/Cidadania 2.0. <<http://www.ciudadania20.org/> (<http://www.ciudadania20.org/>) >. [Consulta: 25/11/2014].
- Dublin Core Metadata Initiative (DCMI). <<http://dublincore.org/> (<http://dublincore.org/>) >. [Consulta: 25/11/2014].
- FOAF: 2000–2014+, Friend of a friend. <<http://www.foaf-project.org/> (<http://www.foaf-project.org/>) >. [Consulta: 25/11/2014].
- OECD (2014). "Recommendation of the Council on Digital Government Strategies". <<http://www.oecd.org/gov/public-innovation/recommendation-on-digital-government-strategies.htm> (<http://www.oecd.org/gov/public-innovation/recommendation-on-digital-government-strategies.htm>) >. [Consulta: 25/11/2014].
- Open Archives Initiative Protocol for Metadata Harvesting. <<http://www.openarchives.org/pmh/> (<http://www.openarchives.org/pmh/>) >. [Consulta: 25/11/2014].
- Temesio, Silvana (2013). "Interoperabilidad de la información en el gobierno electrónico". *Acervo: Revista do arquivo nacional*, vol. 26, no. 2. <<http://www.revistaacervo.an.gov.br/seer/index.php/info/article/view/615> (<http://www.revistaacervo.an.gov.br/seer/index.php/info/article/view/615>) >. [Consulta: 25/11/2014].
- W3C, Data Catalog Vocabulary (DCAT) (2014). <<http://www.w3.org/TR/vocab-dcat/> (<http://www.w3.org/TR/vocab-dcat/>) >. [Consulta: 25/11/2014].
- W3C, Describing Linked Datasets with the VoID Vocabulary (2011). <<http://www.w3.org/TR/void/> (<http://www.w3.org/TR/void/>) >. [Consulta: 25/11/2014].
- W3C, Guía breve linked data. <<http://www.w3c.es/Divulgacion/GuiasBreves/LinkedData> (<http://www.w3c.es/Divulgacion/GuiasBreves/LinkedData>) >. [Consulta: 25/11/2014].
- W3C, OWL: web ontology language. <www.w3.org/2004/OWL (<http://www.w3.org/2004/OWL>) >. [Consulta: 25/11/2014].
- W3C, Resource Description Framework (RDF). <<http://www.w3.org/RDF/> (<http://www.w3.org/RDF/>) >. [Consulta: 25/11/2014].
- W3C (2008). Sparql query language for RDF. <<http://www.w3.org/TR/rdf-sparql-query/> (<http://www.w3.org/TR/rdf-sparql-query/>) >. [Consulta: 25/11/2014].
- W3C (2009). SKOS Simple Knowledge Organization System. <<http://www.w3.org/2004/02/skos/> (<http://www.w3.org/2004/02/skos/>) >. [Consulta: 25/11/2014].
- W3C (2012). OWL 2: Ontology Web Language. <<http://www.w3.org/TR/2012/REC-owl2-overview-20121211/> (<http://www.w3.org/TR/2012/REC-owl2-overview-20121211/>) >. [Consulta: 25/11/2014].

Notes

¹ La llei de l'Uruguai es pot consultar a: <<http://www.impo.com.uy/informacionpublica> (<<http://www.impo.com.uy/informacionpublica>) >.

² <<https://www.zaragoza.es/ciudad/risp/vocabulario-dcat.htm#ftn2>>.

³ Per visualitzar el model ontològic de les categories de despesa del pressupost federal es pot consultar: <<http://vocab.e.gov.br/2013/09/loa> (<<http://vocab.e.gov.br/2013/09/loa>) >.

L'endpoint SPARQL per fer consultes està disponible a: <<http://dados.gov.br/dataset/orcamento-federal/resource/6971c6c1-109e-4462-9463-7a4fee2cd702> (<<http://dados.gov.br/dataset/orcamento-federal/resource/6971c6c1-109e-4462-9463-7a4fee2cd702>) >.

⁴ El portal de dades obertes de l'Uruguai es pot consultar a: <<http://datos.gub.uy/> (<<http://datos.gub.uy/>) >.

⁵ El conjunt de dades es pot consultar a: <<https://catalogodatos.gub.uy/dataset/convenios>>.

Apèndix. Parcial de l'arxiu RDF: *Convenios caminería rural - MTOP*

<<https://catalogodatos.gub.uy/dataset/convenios/resource/32d942b3-6977-4b0f-ae65-a78d66095a2d>>

```
<?xml version="1.0" encoding="UTF-8"?>
< rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" (http://www.w3.org/1999/02/22-rdf-syntax-ns) ">
< rdf:Description rdf:about="http://dgv.mtop.gub.uy/0 (http://dgv.mtop.gub.uy/0) ">
< rdf:type rdf:resource="http://dgv.mtop.gub.uy/convenios (http://dgv.mtop.gub.uy/convenios) "/>
< convenioid xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) " rdf:datatype="http://www.w3.org/2001/XMLSchema#int (http://www.w3.org/2001/XMLSchema#int) ">4.0</convenioid>
< conveniodepto xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">COLONIA </conveniodepto>
< ConvenioPrograma xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">370</ConvenioPrograma>
< ConvenioDetalle xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "></ConvenioDetalle>
< ConvenioAÀ±o xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#int (http://www.w3.org/2001/XMLSchema#int) ">2014.0</ConvenioAÀ±o>
< ConvenioAutorizacion xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime (http://www.w3.org/2001/XMLSchema#dateTime) ">Wed Apr 30
00:00:00 GMT-03:00 2014</ConvenioAutorizacion>
< ConvenioVigencia xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime (http://www.w3.org/2001/XMLSchema#dateTime) ">Wed Dec 31
00:00:00 GMT-03:00 2014</ConvenioVigencia>
< ConvenioResponsable xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">GUSTAVO MIERES</ConvenioResponsable>
< ConvenioExpediente xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#int (http://www.w3.org/2001/XMLSchema#int) "></ConvenioExpediente>
< ConvenioResponsableIntendencia xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">ING. HECTOR
ANZALAS</ConvenioResponsableIntendencia>
< ConvenioMontoTotal xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#double (http://www.w3.org/2001/XMLSchema#double)
">1.8876565E7</ConvenioMontoTotal>
< ConvenioLargoTotal xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#double (http://www.w3.org/2001/XMLSchema#double)
">682.3</ConvenioLargoTotal>
< rdf:type rdf:resource="http://dgv.mtop.gub.uy/Caminos (http://dgv.mtop.gub.uy/Caminos) "/>
< CaminoId xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) " rdf:datatype="http://www.w3.org/2001/XMLSchema#int (http://www.w3.org/2001/XMLSchema#int) ">427.0</CaminoId>
< CaminoNombre xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">Beltramo</CaminoNombre>
< CaminoNombreIntendencia xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "></CaminoNombreIntendencia>
< CaminoNroTramo xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">463</CaminoNroTramo>
< CaminoOrigen xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">Ruta 21, 230k000</CaminoOrigen>
< CaminoDestino xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) ">Al Este, Est. Sr. Beltramo</CaminoDestino>
< CaminoLargoKm xmlns="http://dgv.mtop.gub.uy/ (http://dgv.mtop.gub.uy/) "
rdf:datatype="http://www.w3.org/2001/XMLSchema#double (http://www.w3.org/2001/XMLSchema#double)
">3.0</CaminoLargoKm>
< /rdf:Description>
```

Citació recomanada

Nemesio, Sylvana (2015). «Metadades de dades obertes : estudi de cas a l'Uruguai». *BiD: textos universitaris de biblioteconomia i documentació*, juny, núm. 34. <<http://bid.ub.edu/34/temesio.htm>> [Consulta: 18-06-2015].

