



Escuela de
Ingeniería
Informática
Universidad de Oviedo

Sistemas de Información para la Web

Práctica 9 - Extracción de Información semántica

Esteban Montes Morales - UO246305

Segunda Fase: Cuestiones

¿Qué ontologías usa cada servicio para “tipar” las instancias detectadas en el texto?

- Open Calais
 - Utiliza su propia ontología
- FRED
 - www.ontologydesignpatterns.org
 - www.essepuntato.it
- DBPedia SpotLight
 - Dbpedia

¿Existe algún tipo de dichas ontologías que pudiera considerarse equivalente a otro tipo en schema.org? Señala todos los casos que te hayas encontrado con cada servicio, indicando también (si fuera posible) qué propiedades de un tipo mapearían sobre propiedades del otro.

Para la query:

The New York Times reported that John McCarthy died. He invented the programming language LISP.

OpenCalais: Company

Dbpedia: Newspaper

Digamos que sí comparten propiedades comunes, aunque no bajo el mismo nombre:

Opencalais	DBpedia
nationality	publishingCountry
Non-Executive Chairman	publisher

Reflexiona acerca de los motivos que pueden llevar a cada equipo de desarrolladores a producir una ontología propia. Investiga (someramente) sobre el problema de alineación de ontologías.

La principal razón de que se creen ontologías propias, es debido a que no existan ontologías que cubran las necesidades que se buscan. En cuanto al problema de alineación, viene dado a que al surgir ontologías propias, que se alejan de los estándares definidos, es difícil y tedioso relacionar la información entre ellas, sobre todo cuando desde una ontología “estándar” se intenta integrar datos provenientes de una ontología propia o “no estándar”.

De todas formas existen una gran variedad de técnicas de alineamiento para intentar dar solución a este problema.

Utiliza el servicio sameAs.org para localizar equivalencias para los tipos fundamentales de Open Calais, DBpedia Spotlight y FRED, así como los homólogos de Schema.org que detectaste. ¿Existe algún servicio para el que no se disponga de información en sameAs?

Hay que destacar, que por ejemplo Wikidata, en muchas ocasiones tenía varias propiedades como “sameAs”, “equivalentClass”, etc. que referenciaban a ontologías equivalentes de proveedores diferentes. Por lo que me imagino que también pase en schema.org, dbpedia y demás fuentes principales de ontologías y me imagino así que sea como el servicio web está funcionando, haciendo consultas y sacando los enlaces contenidos en dicho campo, así de forma recursiva hasta obtener todos los enlaces equivalentes.

Equivalent URIs for <http://schema.org/Person> –

1. <http://dbpedia.org/ontology/Person>
2. <http://dbpedia.org/ontology/PokerPlayer>
3. <http://proton.semanticweb.org/protonu#Profession>
4. <http://purl.org/vocabularies/princeton/wn30/synset-person-noun-1>
5. <http://schema.org/Person>
6. http://sw.opencyc.org/2009/04/07/concept/en/_http___dbpedia_org_ont...kerPlayer
7. http://sw.opencyc.org/concept/Mx8Ngh4rY_nOYBt-QdiS5seAT9e7DQ-naHR0c...VyUGxheWVy
8. http://sw.opencyc.org/2009/04/07/concept/Mx8Ngh4rY_nOYBt-QdiS5seAT9...VyUGxheWVy

9. <http://umbel.org/umbel#PersonTypes>
10. <http://umbel.org/umbel/rc/PersonWithOccupation>
11. <http://www.w3.org/2000/10/swap/pim/contact#Person>
12. <http://xmlns.com/foaf/0.1/Person>

Equivalent URIs for <http://schema.org/Organization> –

1. <http://dbpedia.org/ontology/Organisation>
2. <http://proton.semanticweb.org/proton#Organization>
3. <http://proton.semanticweb.org/protonu#SportOrganization>
4. <http://purl.org/vocabularies/princeton/wn30/synset-organization-noun-1>
5. <http://schema.org/Organization>
6. <http://umbel.org/umbel/rc/Organization>
7. <http://umbel.org/umbel/rc/Organizations>
8. <http://xmlns.com/foaf/0.1/Organization>

Equivalent URIs for <http://schema.org/Corporation> –

1. <http://proton.semanticweb.org/protonu#CommercialOrganization>
2. <http://purl.org/vocabularies/princeton/wn30/synset-corporation-noun-1>
3. <http://schema.org/Corporation>
4. <http://umbel.org/umbel/rc/CommercialOrganization>

Es curioso, porque no detecta “Coporation” y “Organization” como equivalentes, a pesar de su similitud, eso me hace deducir que está realizando bien la diferenciación entre una organización y una organización corporativa o empresarial.

Y también se me ocurrió un ejemplo que estuve buscando durante mucho tiempo en wikidata y dbpedia y no pude encontrar, y me sorprendió ver como sameAs.org me detectaba que efectivamente no tenía equivalentes o al menos no parecía tenerlos.

Equivalent URIs for <http://schema.org/Action> –

1. <http://schema.org/Action>

