

Wikis semàntiques: Què són? Com són? Com treballen? Quan calen?

Jorge López Vacares

Enginyeria Tècnica en Informàtica de Gestió

Consultor: Joan Anton Pérez Braña

17 de Desembre de 2012



Aquesta obra està subjecta a una llicència de [Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FITXA DEL TREBALL FINAL

Títol del treball:	Wikis semàntiques: Què són? Com són? Com treballen? Quan calen?
Nom de l'autor:	Jorge López Vacares
Nom del consultor:	Joan Anton Pérez Braña
Data de lliurament (mm/aaaa):	12/2012
Àrea del Treball Final:	XML i Web Semàntica
Titulació:	Eng. Tècnica Informàtica de Gestió
Resum del Treball (màxim 250 paraules):	
<p>Des que al 1995 Ward Cunningham va crear la primera wiki (del hawaiaà "ràpid") amb l'objectiu de compartir les seves idees sobre programació, aquestes han ampliat el seu ventall de propòsits, com enciclopèdies, desenvolupament de programari, gestió del coneixement o de continguts, etc., i han esdevingut molt populars a la web i un dels pilars de la Web 2.0. Fins i tot en algunes àrees empresarials s'han començat a considerar les wikis com una eina que millora l'eficiència en facilitar la cerca d'informació.</p> <p>No obstant, les dades que emmagatzemen es troben en formats dissenyats per a ser consumides per l'ésser humà, el que dificulta que el coneixement pugui ser descobert i compartit per computadors.</p> <p>Amb les tecnologies de la Web Semàntica les wikis han evolucionat convertint-se en wikis semàntiques, on es fa possible marcar el contingut de les seves pàgines i les relacions entre elles amb informació semàntica, fent-la accessible per als computadors.</p> <p>En aquest treball s'analitza el funcionament d'una d'aquestes wikis semàntiques, en concret Semantic MediaWiki, que es basa en MediaWiki, una wiki tradicional mundialment coneguda per ser la base amb la que es va desenvolupar Wikipedia. Abans s'identifiquen les principals tecnologies de la Web Semàntica i es defineix el propòsit de les wikis semàntiques.</p>	
Abstract (in English, 250 words or less):	
Since Ward Cunningham created the first wiki (the hawaiian word for "fast") in 1995 for sharing their ideas about programming, they have expanded their range of purposes, such as encyclopedias, software development, knowledge	

or content management, etc., and have become very popular on the web and one of the pillars of the **Web 2.0**. Even in some areas of business wikis have begun to be regarded as a tool for improving the efficiency in finding information easier.

However, the data are stored in formats designed to be consumed by humans, which complicates the knowledge to be discovered and shared by computers.

With the Semantic Web technologies wikis have evolved and resulting in **semantic wikis**, which makes it possible to mark the content of its pages and the relationships between them with semantic information, making it accessible to computers.

This paper analyzes the performance of one of these semantic wikis, in particular Semantic MediaWiki, which is based on MediaWiki, a traditional wiki worldwide known as the base that was developed with Wikipedia. Before this, the paper identifies the key technologies of the Semantic Web and defines the purpose of semantic wikis.

Paraules clau (entre 4 i 8):

Web, Semantic, Wiki, Ontology, Vocabulary, MediaWiki.

Índex

1. Introducció.....	1
1.1 Context i justificació del Treball	1
1.2 Objectius del Treball.....	1
1.3 Enfocament i mètode seguit.....	1
1.4 Planificació del projecte.....	2
1.5 Breu sumari de productes obtinguts.....	4
1.6 Breu descripció dels altres capítols de la memòria	5
2. La Web Semàntica	6
2.1. XML.....	8
2.1.1. Well-formed documents	8
2.1.2. Document vàlid	9
2.1.3. Namespaces.....	10
2.2. RDF.....	10
2.2.1. RDF/XML	11
2.2.2. RDF Schema	12
2.3. OWL	12
2.4. Llenguatges de Consulta.....	13
2.4.1. XPath.....	13
2.4.2. XQuery.....	14
2.4.3. SPARQL	14
3. Wikis semàntiques	16
3.1. AceWiki	19
3.2. IkeWiki.....	20
3.3. Kaukolu	21
3.4. KnowWE.....	21
3.5. Makna.....	22
3.6. OntoWiki.....	22
3.7. Semantic MediaWiki	23
3.8. SweetWiki.....	24
3.9. TaOPis	25
3.10. Taula comparativa	25
4. Semantic MediaWiki.....	28
4.1. Instal·lació i configuració	28
4.2. Afegir semàntica.....	30
4.3. Semàntica basada en ontologies	32
4.4. Navegació semàntica	35
4.5. Cerca semàntica	38
4.6. Inferència.....	42
4.7. Exportació de les dades	44
5. Conclusions.....	49
6. Glossari	50
7. Bibliografia.....	52
8. Annexos	54
8.1. Diagrama de classes de la MusicOntology (MusicalManifestation).....	54

8.2. Diagrama de classes de la MusicOntology (MusicArtist, MusicGroup, SoloMusicArtist)	55
---	----

Llista de figures

Figura 2-1. Pila tecnològica de la Web Semàntica.....	8
Figura 2-2. Arbre del document XML del capítol 2.1.....	9
Figura 2-3. Graf RDF que descriu a Eric Miller del W3C.....	11
Figura 2-4. SPARQL endpoint de DBpedia.....	15
Figura 3-1. Graf RDF resultant de l'anotació de la pàgina de Barcelona.....	17
Figura 3-2. Interfície web d'AceWiki.....	20
Figura 3-3. Arquitectura d'IkeWiki.....	20
Figura 3-4. Finestra emergent amb anotacions de la pàgina.....	21
Figura 3-6. Arquitectura de Makna.....	22
Figura 3-7. Vista d'una instància en forma de mapa.....	23
Figura 3-8. Factbox a una pàgina sobre Berlín.....	24
Figura 3-9. Sistema d'auto-completat de SweetWiki.....	24
Figura 3-10. Visualització de la ontologia d'un projecte a TaOPis.....	25
Figura 4-1. Arquitectura de SMW... ..	28
Figura 4-2. Pàgina principal de la wiki semàntica sobre rock irlandès.....	29
Figura 4-3. Factbox a la pàgina del grup The Cranberries.....	31
Figura 4-4. Importació de la Music Ontology a SMW.....	33
Figura 4-5. Article sobre el grup U2.....	35
Figura 4-6. Pàgina de la propietat member.....	36
Figura 4-7. Pàgina Special:Browse amb les propietats d'U2.....	37
Figura 4-8. Pàgina Special:SearchByProperty per la propietat primary_instrument amb el valor "baix".....	38
Figura 4-9. Consulta semàntica amb Special:Ask.....	39
Figura 4-10. Membres d'U2 extrets amb la funció #ask.....	42
Figura 4-11. Gary Moore és una instància de SoloMusicArtist.....	43
Figura 4-12. Fitxer OWL/RDF de l'article sobre Bono.....	45
Figura 4-13. Graf simplificat sobre Bono.....	46
Figura 4-14. Exportació de les propietats de la wiki amb SMW_dumpRDF.php.....	47
Figura 4-15. Fragment de l'arxiu wiki_prop.rdf.....	48

1. Introducció

1.1 Context i justificació del Treball

Als darrers anys les wikis s'han convertit en populars entorns per a la col·laboració a la web i una important via de compartició de coneixement. Fins i tot en algunes àrees empresarials s'han començat a considerar com una eina que millora l'eficiència en facilitar la cerca d'informació sobre negocis, processos, etc. No obstant, les dades que emmagatzemen es troben en formats dissenyats per a ser consumides per l'ésser humà, el que dificulta que el coneixement pugui ser descobert i compartit per computadors, un dels principals objectius de la Web Semàntica.

Arran d'aquesta restricció les wikis evolucionen i neixen les wikis semàntiques. Aquestes incorporen ontologies subjacents que possibiliten l'addició d'anotacions semàntiques al contingut de les pàgines. Aquestes ontologies, ocultes per a les persones, habiliten a programes informàtics l'accés a la informació per a descobrir coneixement, compartir-lo i reutilitzar-lo. A més, les wikis semàntiques milloren la navegació i l'eficàcia de les cerques.

1.2 Objectius del Treball

- Estudiar els conceptes en que es basa la Web Semàntica.
- Conèixer l'estructura dels principals llenguatges de representació i consulta d'informació emprats a la Web Semàntica.
- Estudiar el món de les wikis semàntiques identificant els seus avantatges e inconvenients envers les wikis convencionals.
- Aprofundir en l'anàlisi d'una wiki semàntica concreta.

1.3 Enfocament i mètode seguit

Primerament caldrà familiaritzar-se amb el terme Web Semàntica i fer un mínim d'estudi de les tecnologies que l'envolten per tal d'identificar les més importants, i que serviran de base de coneixement per a l'estudi. Aquest, constarà bàsicament de tres fases.

En una primera fase s'introduiran lleugerament les tecnologies en que es basa la Web Semàntica.

En una segon s'aprofundirà en el món de les wikis semàntiques en general, d'on sorgeixen, quins són els seus objectius, com funcionen, etc. S'identificaran els seus avantatges i inconvenients envers les wikis tradicionals, i es

comentaran deu de les wikis semàntiques més conegudes, identificant i comparant les seves funcionalitats.

Per últim es crearà una wiki semàntica amb Semantic MediaWiki per tal de provar el seu funcionament en quant a la gestió de dades semàntiques, navegació i cerca, importació de vocabularis i exportació de dades.

1.4 Planificació del projecte

Per al correcte seguiment d'aquest TFC s'han identificat una sèrie d'entregues parcials que s'han de complir per tal d'aprovar l'assignatura:

Tasca	Data d'entrega
Pla de treball (PAC1)	01/10/2012
PAC2	05/11/2012
PAC3	17/12/2012
Lliurament final	08/01/2012

A continuació es detalla el que es pretén entregar en aquestes dates clau:

- **Pla de treball:** Aquest document.
- **PAC2:** Memòria desenvolupada fins al capítol 3 acabat. Programari instal·lat. Meta-dades a importar seleccionades.
- **PAC3:** Memòria desenvolupada completament.
- **Lliurament final:** Memòria finalitzada i revisada i vídeo presentació.

La setmana següent al lliurament final tindrà terme un debat virtual amb el tribunal que avaluarà el TFC, el qual podrà fer preguntes al respecte i que s'hauran de respondre en un plaç màxim de 24 hores.

Es preveu el següent calendari de treball. S'han identificat el groc les dates clau inamovibles i en verd els dies de vacances de la feina. Aquests dies de vacances, i també els dissabtes i els diumenges s'invertirà un mínim de 4 hores de feina, mentre que els dies laborables s'emprarà un mínim de 3 hores.

A continuació es detalla el calendari:

Octubre							Novembre							7	Dia festiu
DL	DM	DX	DJ	DV	DS	DG	DL	DM	DX	DJ	DV	DS	DG		
1	2	3	4	5	6	7				1	2	3	4		
8	9	10	11	12	13	14	5	6	7	8	9	10	11		
15	16	17	18	19	20	21	12	13	14	15	16	17	18		Dates clau
22	23	24	25	26	27	28	19	20	21	22	23	24	25		
29	30	31					26	27	28	29	30				Vacances

Desembre							Gener							1	2
DL	DM	DX	DJ	DV	DS	DG	DL	DM	DX	DJ	DV	DS	DG		
					1	2		1	2	3	4	5	6		
3	4	5	6	7	8	9	7	8	9	10	11	12	13		
10	11	12	13	14	15	16	14	15	16	17	18	19	20		
17	18	19	20	21	22	23	21	22	23	24	25	26	27		
24	25	26	27	28	29	30	28	29	30	31					
31															

S'ha de tenir present que entre els dies 24 a 26 de Desembre es preveuen compromisos familiars, com s'especifica a l'apartat 7.3 d'aquest document. En resum, es podran dedicar al voltant de 337 hores de feina en el desenvolupament del TFC (no s'ha tingut en compte la setmana del debat virtual):

Mes	Hores laborals	Hores vacances/festius	Total
Octubre	66	36	102
Novembre	60	40	100
Desembre	39	72	111
Gener	12	12	24
Total	177	160	337

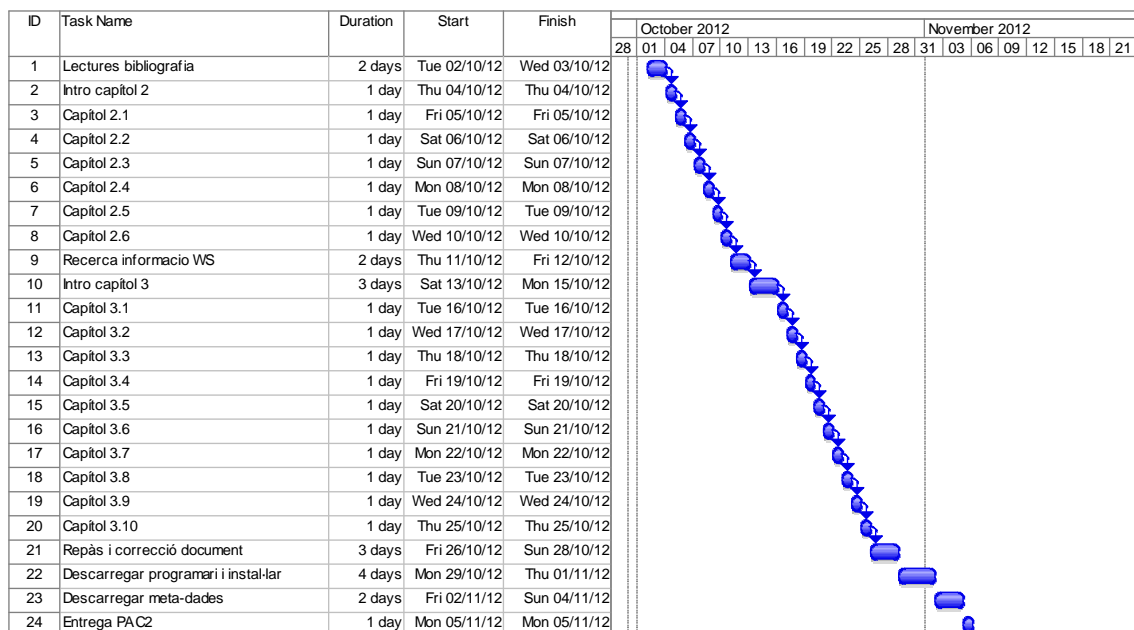
Com s'ha especificat anteriorment aquesta estimació d'hores correspondria a un mínim. S'ha de tenir en compte que, en cas que la feina s'endarrereixi en alguna fase i, per tal de no comprometre el treball final ni cap de les dates clau, podria ser necessari invertir més hores en alguns dies puntuals. Arribats a aquest punt, és factible treballar fins a 2 hores més els dies laborals i fins a 4 més durant els dies de vacances, dissabtes i festius.

Es detalla seguidament la planificació de cada una de les fases en que es divideix en TFC partint de les dates clau i es detalla cada una de les fases.

Planificació general:

ID	Task Name	Duration	Start	Finish												
					Oct '12	Nov '12	Dec '12	Jan '13								
					24	01	08	15	22	29	05	12	19	26	03	10
1	PAC2	25 days	Tue 02/10/12	Mon 05/11/12												
2	PAC3	30 days	Tue 06/11/12	Mon 17/12/12												
3	Lliurament final	16 days	Tue 18/12/12	Tue 08/01/13												
4	Debat virtual	7 days	Wed 09/01/13	Thu 17/01/13												

Planificació PAC2:



Planificació PAC3:



S'ha optat per no ajustar la planificació d'aquesta fase al temps total establert des d'un principi, ja que és bastant probable que el capítol creixi a mesura que avanci el treball i augmentin els coneixements en la matèria.

És recomanable també deixar marge per a possibles modificacions en l'índex i, per tant, de la metodologia de treball, atenent a les indicacions que el consultor de l'assignatura pugui fer en corregir la PAC2.

No s'ha planificat el Lliurament final ja que només correspon a una tasca, ni tampoc el debat amb el tribunal del TFC en no haver cap tasca a desenvolupar.

1.5 Breu sumari de productes obtinguts

Al final d'aquest projecte es desenvolupa una petita wiki amb Semantic MediaWiki. Es veu breument la seva instal·lació, configuració, interfície i funcionalitats bàsiques.

Així mateix, es fa servir una ontologia externa i es creen alguns articles per analitzar en profunditat com s'insereix informació semàntica i com es pot extreure per a consulta o per a compartir amb aplicacions externes.

1.6 Breu descripció dels altres capítols de la memòria

El **Capítol 2** descriu les tecnologies en que es basa el funcionament de la Web Semàntica. Partint del llenguatge més bàsic per a la compartició de dades a Internet com és l'XML, fins als llenguatges de consulta més complexos com l'SPARQL.

El **Capítol 3** situa les wikis semàntiques dins del context de la Web Semàntica, identifica les seves característiques més importants i les compara amb les wikis tradicionals. A més s'introdueixen algunes de les wikis semàntiques més importants, entre elles Semantic MediaWiki, una de les més emprades arreu.

Semantic MediaWiki s'analitza en profunditat al **Capítol 4**. En aquest, s'importa una ontologia externa per crear una petita wiki sobre el món de la música i estudiar com gestiona les dades semàntiques.

Les conclusions d'aquest estudi es comenten al **Capítol 5**, mentre que als **Capítols 6 i 7** es pot trobar el glossari i la bibliografia respectivament. L'annex, al **Capítol 8** presenta part de l'esquema de classes de la Music Ontologia, la ontologia emprada a la wiki desenvolupada.

2. La Web Semàntica

La major part del contingut de la Web avui en dia està dissenyada per ser consumida per l'ésser humà. Inclús si aquest contingut s'ha generat automàticament partint d'una base de dades es presenta generalment sense la seva informació estructural original a la base de dades. Un ús típic que l'ésser humà fa de la Web avui en dia és buscar i consumir informació, entrar en contacte amb altres éssers humans, revisar catàlegs de botigues online, etc.

Per trobar la informació que es necessita, sovint es fa servir un motor de cerca com Google, Yahoo, o Bing, però la utilització d'aquestes eines presenten els següents problemes:

- La majoria de vegades la cerca retorna un gran nombre de direccions a pàgines web que poden ser rellevants o no.
- Altres vegades pot succeir que la cerca no retorni cap resultat o que d'entre els pocs resultats obtinguts no es trobin pàgines rellevants. Això ens porta al tercer problema:
- Els resultats depenen altament del vocabulari. Sovint hem de fer servir paraules semànticament similars per trobar allò que estem buscant, i això és així perquè els documents poden fer servir una terminologia diferent de la que fem servir a la cerca.
- Els resultats són pàgines web. En cas de necessitar informació que es troba distribuïda en més d'un document hem de fer diverses consultes per trobar els documents rellevants i extreure'n la informació manualment de cadascun d'ells.

La intervenció doncs de l'esser humà es fa necessària per interpretar i combinar els resultats trobats en una cerca perquè aquests no són processables pels computadors. Per a que això fóra possible s'haurien de fer servir sofisticades tècniques basades en Intel·ligència Artificial, però tot i els avenços en aquest camp en els últims temps, sembla encara un objectiu massa ambiciós.

Una altra solució és la que proposa el World Wide Web Consortium¹ (W3C) que, impulsat pel creador de la Web Tim Berners-Lee, ha creat i manté actualitzada una col·lecció d'estàndards i tecnologies que busquen representar el contingut de la Web en un format llegible per a l'ésser humà però a la vegada processable per als computadors. El W3C es refereix a aquest marc com la Web Semàntica.

A grans trets, les tecnologies en les que es basa la Web Semàntica es poden resumir en tres grans grups:

¹ <http://www.w3.org/>

Metadades

El terme metadades es refereix a informació sobre les dades, presentada en un format que pugui ser entès tant per humans com per a ordinadors. Avui en dia el llenguatge de metadades predominant a les pàgines web és l'HTML. Aquest llenguatge aporta informació de com s'estructura un document (tipus de lletra, taules, colors, etc.) i descriu com s'ha de presentar, p.ex. a un navegador web.

La Web semàntica proposa canviar l'HTML per XML, un llenguatge que a més de la informació de format del document incorpori informació sobre el seu contingut. XML és descrit amb més detall a l'apartat 2.1.

Ontologies

D'acord amb la W3C: "les ontologies o vocabularis defineixen conceptes i relacions per a descriure i representar un àrea de coneixement". A aquesta àrea de coneixement se la coneix com a domini.

Típicament, una ontologia consta d'un nombre finit de conceptes (termes) i les relacions entre ells. Els termes són classes d'objectes d'un domini, mentre que les relacions representen habitualment jerarquies entre aquestes classes, p.ex., un gos és un tipus de mamífer.

Les ontologies o vocabularis en la web semàntica ajuden a la integració de dades quan, p.ex., poden existir ambigüitats amb els termes utilitzats en diferents datasets, o quan un mica de coneixement extra pot esdevenir el descobriment de noves relacions. Considerem, per exemple, una aplicació d'ontologies en el camp de la salut, que els metges fan servir per representar coneixement sobre símptomes, malalties i tractaments. Per altra banda, els farmacèutics la fan servir per representar informació sobre medicaments, dosis i al·lèrgies. La combinació d'aquests dos coneixements amb dades sobre pacients permet una àmplia gamma d'aplicacions intel·ligents, com eines de suport de decisions per buscar possibles tractaments, sistemes de monitoratge de l'eficàcia d'un medicament i els efectes secundaris, etc.

Les ontologies o vocabularis també són útils per organitzar coneixement en llibreries, museus, diaris, portals de governs, companyies, aplicacions per a xarxes socials, i altres comunitats que gestionen grans col·leccions de llibres, articles, notícies, entrades de blogs, etc.

Una altra utilitat és la de millorar la precisió de les cerques a Internet. Els motors de cerca ontològics poden buscar pàgines o fragments de pàgines que facin referència a un concepte precís d'una ontologia.

Inferència

En la Web Semàntica, la inferència es caracteritza pel descobriment de noves relacions entre termes dins l'àrea de coneixement donat. Per exemple, entre les

múltiples relacions d'una ontologia pot trobar-se la relació “un gos és un mamífer”, i per altra banda trobar-se també l’afirmació “Pluto és un gos”. D’aquí es dedueix, és a dir, s’infereix, que “Pluto és un mamífer”.

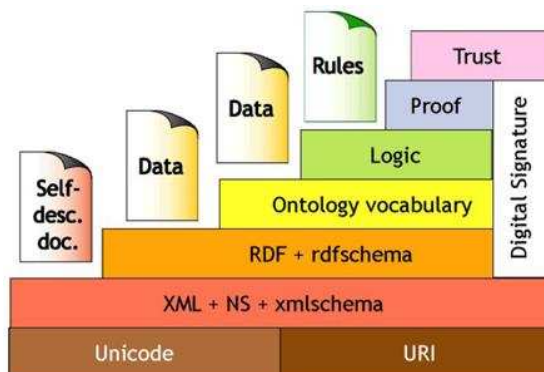


Figura 2-1. Pila tecnològica de la Web Semàntica.

El desenvolupament de la Web Semàntica procedeix en fases, cadascuna de les quals construeix una capa amb estàndards damunt l’anterior construint una pila tecnològica (**Figura 2-1**).

2.1. XML

L’Extensible Markup Language (XML²), és una simplificació d’un estàndard anterior molt més complex anomenat SGML. Tot i el seu nom no és un llenguatge en si mateix, sinó un conjunt de regles recomanades de manera oficial pel W3C per representar qualsevol tipus d’informació d’una manera estructurada, des de llibres fins a fórmules matemàtiques. Avui en dia és un dels formats més àmpliament utilitzats per intercanviar informació entre computadors i entre aquests i les persones.

Un document XML consisteix en un pròleg i un o més elements que l’estructuren en parts. Els elements es delimiten per mitjà de dues etiquetes a l’estil HTML, una d’obertura i una altra de tancament, i pot contenir dades, altres elements, o estar buit:

```
<?xml version="1.0"? encoding="UTF-16">
<album>
  <title>The Joshua tree</title>
  <author>U2</author>
  <released>1987</released>
  <price currency="Euro"></price>
</album>
```

2.1.1. Well-formed documents

² <http://www.w3.org/TR/xml/>

Quan un document XML és sintàcticament correcte es diu que està ben format (well-formed). Hi ha unes regles bàsiques per a considerar que un document XML està ben format:

- Només hi ha un element més extern, anomenat arrel.
- Els elements han d'estar correctament niats.
- Cada element conté una etiqueta d'obertura i una de tancament.
- Els elements buits estan tancats.
- Els atributs dels elements fan servir noms únics i els seus valors s'escriuen entre cometes dobles.

Els documents well-formed es poden representar gràficament en forma d'arbre, com es pot veure a la **Figura 2-2**.

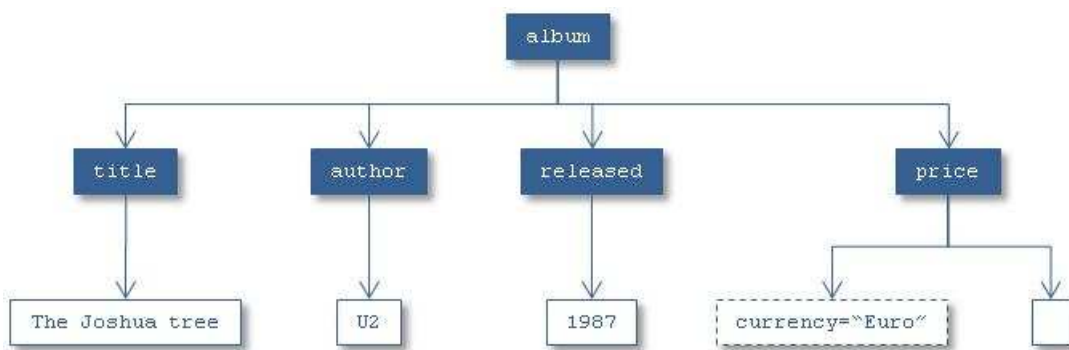


Figura 2-2. Arbre del document XML del capítol 2.1.

2.1.2. Document vàlid

A més de sintàcticament correcte, un document XML ho ha de ser també estructuralment. Cal definir prèviament quins noms d'elements i d'atributs es faran servir, quins valors poden prendre aquests atributs, quins elements poden anar dins de quins elements, etc. Quan un document està ben format, fa servir una informació d'estructura i està d'acord a aquesta estructura, es diu que és un document vàlid.

Existeixen dues maneres de definir l'estructura d'un document: mitjançant DTD (extern o intern) i XML Schema, essent aquest últim més complet i potent. La seva sintaxis es basa en el propi llenguatge XML i permet la creació de nous esquemes a partir d'esquemes ja existents, el que proveeix una aproximació a la programació Orientada a Objectes. El següent esquema defineix l'estructura del document XML de l'apartat 2.1.

```
<schema "http://www.w3.org/2000/10/XMLSchema">
  <element name="album" type="albumType"/>

  <complexType name="albumType">
    <sequence>
      <element name="title" type="string"/>
      <element name="author" type="string"/>
      <element name="released" type="Year"/>
    </sequence>
  </complexType>
</schema>
```

```

    <element name="price" type="decimal"/>
    <attribute name="currency" type="string"
        use="required"/>
</sequence>
</complexType>

```

2.1.3. Namespaces

Un document XML pot fer servir més d'un DTD o esquema alhora amb l'objectiu de reutilitzar noms d'elements i atributs ja definits. Aquest fet pot esdevenir en conflictes si un mateix nom es fa servir amb significats diferents. Per exemple, si un document XML es construeix en base a dos esquemes, i tots dos defineixen un element *e* però amb diferents significats, una eina que intentés processar el document seria incapaç, en trobar l'element *e*, d'identificar a quin esquema fa referència.

La solució passa per fer servir un prefix davant del nom d'element o d'atribut que identifiqui en cada moment a quin esquema pertany. Els prefixos es defineixen dins un element XML i, un cop definits, es poden emprar dins el mateix element o en qualsevol dels seus fills.

L'exemple que mostra a continuació correspon al document XML de l'apt. 2.1. en que s'especifica que es fa servir l'esquema de l'apt. 2.1.2 amb el prefix *mus*:

```

<?xml version="1.0"? encoding="UTF-16">
<mus:album xmlns:mus="http://www.exemple.cat/schema/">>
  <mus:title>The Joshua tree</mus:title>
  <mus:author>U2</mus:author>
  <mus:released>1987</mus:released>
  <mus:price></mus:price>
</mus:album>

```

2.2. RDF

El llenguatge XML proveeix un marc per estructurar i intercanviar dades, però els documents manquen de semàntica. El Resource Description Framework (RDF³) és un model de dades per representar informació sobre els recursos que hi ha a la web (serveis, pàgines, documents, direccions de correu electrònic, etc.) i les seves característiques o propietats.

La descripció d'una propietat d'un recurs és un triplet *subjecte-predicat-objecte*. Un subjecte és simplement un recurs, i es defineix mitjançant una cadena de caràcters anomenada URI (Uniform Resource Identifier). Un predicat és la propietat que es vol descriure del recurs, mentre que el valor del predicat pot ser una cadena de caràcters, un valor numèric o un altre recurs. Els predicats també s'identifiquen fent servir URIs.

³ <http://www.w3.org/TR/rdf-primer/>

Existeix una manera de representació d'un triplet fàcilment comprensible per a l'ésser humà en forma de graf, on els nodes simbolitzen subjectes i objectes, i els predicats es dibuixen en forma de fletxa amb origen al subjecte.

A la **Figura 2-3** es mostra un graf amb triplets que descriuen al Dr. Eric Miller del W3C.



Figura 2-3. Graf RDF que descriu a Eric Miller del W3C.

Font: <http://www.w3.org/TR/rdf-primer/>.

“Copyright © 2004 W3C® (MIT, ERCIM, Keio), All Rights Reserved.”.

2.2.1. RDF/XML

RDF/XML⁴ proveeix una sintaxis basada en XML que permet escriure documents que representin un graf, fent així que la informació pugui ser accessible i processable per computadors. A continuació es mostra el document RDF/XML per al graf de l'Eric Miller:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person
    rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```

⁴ <http://www.w3.org/TR/rdf-syntax-grammar/>

2.2.2. RDF Schema

El llenguatge RDF Schema⁵ permet especificar vocabularis que es puguin fer servir en triplets per indicar tipus de recursos i les seves propietats d'una manera similar a com en un llenguatge OO es defineixen classes i atributs amb UML. Les instàncies en aquest cas són els propis recursos.

Un vocabulari, doncs, pot ser esquematitzat com una jerarquia de classes. Per fer-lo llegible per als ordinadors, però, també es pot plasmar en un document RDF/XML.

2.3. OWL

Les ontologies són necessàries per aplicacions que volen cercar i fusionar informació en diverses comunitats. Tot i que XML, DTD i XML Schema són suficients per intercanviar dades entre parts coneixedores de cert vocabulari, la seva falta de semàntica evita que els computadors puguin fer aquesta tasca de manera fiable quan es dóna un vocabulari nou: un mateix terme podria tenir significats diferents en diferents contextos, i diferents termes podrien tenir el mateix significat.

RDF i RDF Schema comencen a apropar-se a aquest problema al permetre associar una simple semàntica amb identificadors. Amb RDF Schema es poden definir classes amb múltiples subclasses i superclasses, i es poden definir propietats amb subpropietats. En aquest sentit, RDF Schema és un llenguatge d'ontologies simple. No obstant, per tal d'aconseguir interoperabilitat entre nombrosos esquemes desenvolupats i gestionats de forma autònoma es necessiten semàntiques més riques. Per exemple, RDF Schema no pot especificar que les classes Persona i Cotxe són disjunts, o que un quartet de corda està format exactament per quatre músics.

El Grup de Treball d'Ontologies per a la Web del W3C (Web-Ontology Working Group) va identificar una sèrie de casos d'ús⁶ per a la Web Semàntica que requerien molta més expressivitat de la que oferien RDF i RDF Schema. També altres grups d'investigadors en Amèrica i Europa van arribar a aquesta conclusió, i van liderar un projecte conjunt que va ser el punt de partida de l'OWL⁷, actualment en la seva versió 2, que pretén ser el llenguatge ontològic estàndard i àmpliament acceptat per a la Web Semàntica.

Encara que hi ha altres sintaxis (Manchester⁸, Functional⁹, Turtle¹⁰, OWL/XML¹¹), un document OWL habitualment està escrit en RDF/XML, fent ús de l'element `owl:`

⁵ <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>

⁶ <http://www.w3.org/TR/2004/REC-webont-req-20040210/>

⁷ <http://www.w3.org/TR/owl2-overview/>

⁸ <http://www.w3.org/TR/2009/NOTE-owl2-manchester-syntax-20091027/>

⁹ <http://www.w3.org/TR/2009/REC-owl2-syntax-20091027/>

¹⁰ <http://www.w3.org/TeamSubmission/turtle/>

```

<owl:Class rdf:about="associateProfessor">
  <owl:disjointWith rdf:resource="#professor"/>
  <owl:disjointWith rdf:resource="#assistantProfessor"/>
</owl:Class>

```

S'han identificat tres subllenguatges d'OWL, cadascun d'ells orientat a diferents necessitats:

- **OWL Full:** ofereix màxima expressivitat sense restriccions, fet que limita la seva capacitat computacional.
- **OWL DL:** aplica certes restriccions a OWL Full, el que permet donar un bon suport al raonament automàtic.
- **OWL Lite:** aplica encara més restriccions fent-lo més simple, convertint-lo en un llenguatge més fàcil d'entendre i d'implementar, però a costa de limitar la seva expressivitat.

2.4. Llenguatges de Consulta

Com a qualsevol sistema d'informació, en el context de la Web Semàntica calen llenguatges de consulta que permetin accedir i recuperar les dades. Seguidament es llisten alguns dels llenguatges més importants.

2.4.1. XPath

L'XML Path Language (XPath¹²) permet navegar per l'arbre d'un document XML, modelant aquest com un arbre de nodes. Existeixen varis tipus de nodes, entre els quals es troben nodes element, nodes atribut i nodes text. La construcció sintàctica bàsica en XPath és una expressió que, un cop avaluada produeix un objecte d'un dels següents tipus de dades bàsics: booleà, número, cadena o conjunt de nodes.

Un tipus d'expressió important és el camí de localització, que selecciona un conjunt de nodes partint del node arrel (camí absolut) o de qualsevol altre (camí relatiu), i s'especifica mitjançant una seqüència d'un o més passos de localització separats pel caràcter '/'. Per exemple, la expressió:

```
/bookstore/book[price>35.00]
```

retorna un conjunt de nodes `book` que són fills del node `bookstore` i tenen un element `price` amb un valor major que 35.

¹¹ <http://www.w3.org/TR/2009/REC-owl2-xml-serialization-20091027/>

¹² <http://www.w3.org/TR/xpath/>

2.4.2. XQuery

L'XML Query Language (XQuery¹³), és una recomanació del W3C per recuperar i interpretar informació emmagatzemada en XML. XQuery és a XML el que SQL és a les taules d'una base de dades relacional. Ha sigut dissenyat de tal manera que les seves consultes són concises i fàcilment enteses, i és prou flexible per consultar un ampli espectre de fonts d'informació XML, incloent bases de dades i documents.

Existeixen més de 100 funcions construïdes en XQuery que es poden cridar i que cobreixen un ampli rang de funcionalitats, com manipular cadenes, realitzar càlculs matemàtics o combinar seqüències d'elements entre moltes altres. A més és possible definir funcions pròpies, ja sigui dins el propi document de consulta o en llibreries externes. La següent expressió retorna, ordenat alfabèticament, un llistat de títols dels llibres amb preu més gran que 30.

```
for $x in doc("books.xml")/bookstore/book
where $x/price>30
order by $x/title
return $x/title
```

2.4.3. SPARQL

SPARQL¹⁴ (pronunciat "sparkle") és un acrònim recursiu d'SPARQL Protocol And RDF Query Language. Com es desprèn de la part 'RQL' del nom, està dissenyat per consultar dades guardades en format RDF, encara que també és possible consultar dades relacionals o XML. En paraules del director del W3C, Tim Berners-Lee, "Intentar fer servir la Web Semàntica sense SPARQL és com intentar utilitzar bases de dades relacionals sense SQL".

La veritable utilitat d'SPARQL s'aprecia en consultar datasets públics. La majoria d'aquests datasets ofereixen un *SPARQL endpoint* que permet llençar consultes directament des del navegador web, sense necessitat d'instal·lar cap programari localment. Un dels dataset públics més populars és DBpedia amb el seu endpoint SNORQL¹⁵, que alberga les dades que es poden trobar als quadres grisos de la dreta de les pàgines de Wikipedia.

¹³ <http://www.w3.org/TR/xquery/>

¹⁴ <http://www.w3.org/TR/rdf-sparql-query/>

¹⁵ <http://dbpedia.org/snorql/>

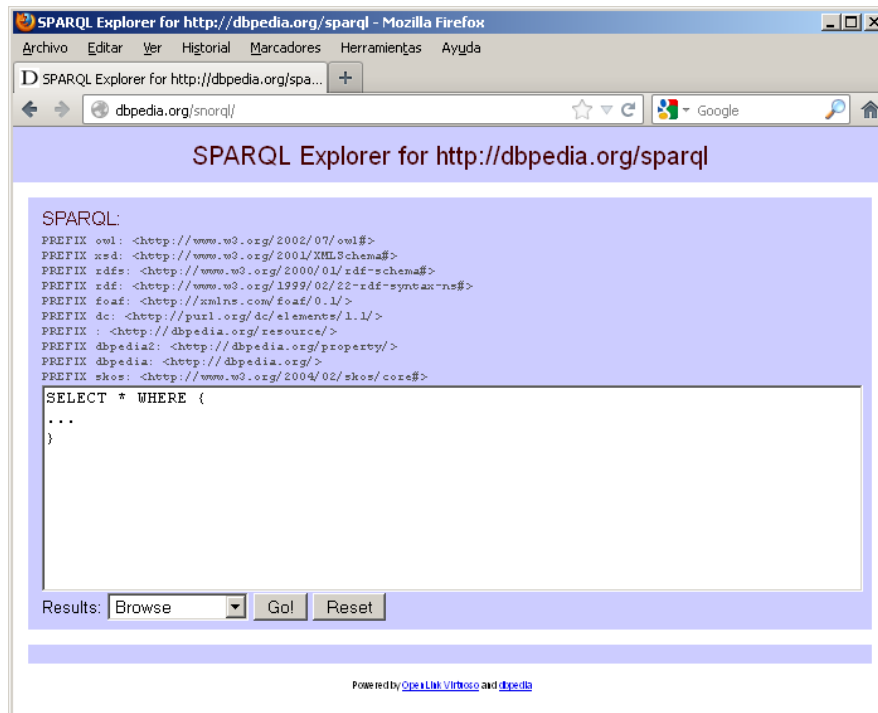


Figura 2-4. SPARQL endpoint de DBpedia.

Com a exemple del tipus de consulta que es pot construir amb SPARQL, la següent instrucció executada sobre SNORQL recupera un llistat d'àlbums i els seus intèrprets que han sigut produïts per Timbaland:

```
SELECT ?artistName ?albumName
WHERE
{
  ?album d:producer :Timbaland .
  ?album d:musicalArtist ?artist .
  ?album rdfs:label ?albumName .
  ?artist rdfs:label ?artistName .
}
```

3. Wikis semàntiques

A grans trets, una wiki és una aplicació web que gestiona una col·lecció de pàgines web, on qualsevol pot crear noves pàgines i editar-ne d'existents fent servir un simplificat llenguatge de marques. Una aplicació anomenada motor wiki s'encarrega de fer-la funcionar.

La primera wiki va ser creada al 1995 per Ward Cunningham¹⁶, i als darrers anys s'han convertit en populars entorns per a la col·laboració a la web i una important via de compartició de coneixement. Fins i tot en algunes àrees empresarials s'ha començat a veure ja com, de manera interna, les wikis poden ajudar a ser més eficients en facilitar la cerca d'informació sobre negocis, processos, etc. Els seus avantatges més importants es llisten a continuació:

- **Simplicitat a l'hora de crear una nova pàgina:** per crear una nova pàgina només cal introduir l'adreça al navegador. Automàticament es visualitza una plantilla que permet crear la pàgina.
- **Simplicitat a l'hora d'editar una pàgina existent:** tota pàgina disposa d'un enllaç que permet editar-la sense haver de conèixer HTML ni cap altre llenguatge, ja que aquest es genera automàticament.
- **Simplicitat per enllaçar amb altres pàgines:** no cal conèixer la direcció sinó tan sols indicar la pàgina al motor wiki.
- **Històric de canvis:** totes les modificacions queden registrades i es pot tornar a versions anteriors en cas d'error.

No obstant, les wikis també tenen importants limitacions que cal valorar:

- **Descobriment del coneixement:** la informació està estructurada per ser llegida per humans i no per ser processada per un ordinador. Per exemple, no hi ha cap manera de “preguntar” a Wikipedia per un llistat de ciutats amb més de 100 anys d'història, que tinguin una alcaldessa i que hagin albergat uns jocs olímpics. Tot i que les wikis poden fer cerques de text, es poden presentar problemes de llenguatge (ambigüitat, idioma, etc) que fan que el resultat s'hagi de revisar i filtrar posteriorment.
- **Reutilització del coneixement:** degut a la problemàtica exposada al punt anterior, la reutilització del coneixement de manera automàtica és del tot impossible.
- **Consistència en el coneixement:** arrel de que tothom pot afegir i editar contingut en una wiki, un mateix concepte pot aparèixer en múltiples pàgines i en múltiples termes, dificultant la cerca d'informació.

¹⁶ http://es.wikipedia.org/wiki/Ward_Cunningham

Cal, per tal de fer la informació accessible i processable per als ordinadors, dotar-la de semàntica, el que fa evolucionar les wikis a wikis semàntiques. De fet, no són realment les wikis les que evolucionen, sinó els motors. Així, una wiki semàntica es pot definir com una wiki que funciona amb un motor wiki semàntic. Aquest nous motors permeten als usuaris afegir marcat semàntic a les pàgines que després es pot fer servir per realitzar cerques més eficaces i intercanvi d'informació.

En tota wiki, un document està format per dos elements: enllaços i text. Als primers tipus de wikis semàntiques, dotar de semàntica al document consistia en anotar aquests dos elements amb propietats i valors per tal de descriure el seu significat, el que resulta en *enllaços classificats* i en *atributs*. Els enllaços classificats incorporen una relació binària entre pàgines de la wiki, mentre que els atributs representen una relació entre la pàgina i una dada.

Per exemple, l'anotació 'Located in:Catalonia' representaria un enllaç classificat que enllaçaria la pàgina de Barcelona amb la de la Catalonia. Per altra banda, la propietat 'has population:1621537' representa un atribut de la ciutat, on 1621537 és el valor de la propietat i no una altra pàgina. Cada una d'aquestes anotacions es pot representar com un triplet RDF, on el subjecte és la pròpia pàgina que conté l'anotació:

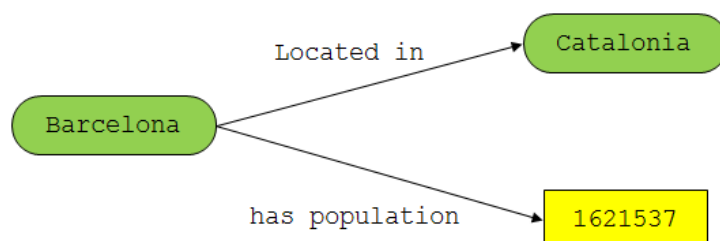


Figura 3-1. Graf RDF resultant de l'anotació de la pàgina de Barcelona.

Aquest mètode aporta una manera de construir ontologies simples, però aquestes tenen l'inconvenient d'aportar una baixa capacitat de raonament. Per exemple, no es poden descriure relacions d'herència. La necessitat d'haver d'expressar ontologies més riques va impulsar la creació d'un nou tipus de wikis semàntiques. Aquestes, incorporen internament una petita ontologia que permet dotar d'una informació semàntica més completa als documents. A més, és possible importar noves ontologies desenvolupades externament.

Un dels primers motors wiki en treballar en aquest sentit va ser Semantic MediaWiki, que incorpora SWiVT¹⁷. El Semantic Wiki Vocabulary and Terminology (SWiVT), està escrit en OWL DL, i incorpora termes com `Wikipage` o `wikiPageModificationDate`.

A les wikis semàntiques es poden identificar certes fortaleeses i debilitats que cal comentar. Comencem per les primeres:

¹⁷ <http://semantic-mediawiki.org/swivt/>

- **Facilitat d'ús:** per als usuaris d'una wiki, el pas a una wiki semàntica no requereix cap aprenentatge sobre ontologies, sinó que únicament ha d'aprendre a fer anotacions.
- **Navegació:** la informació semàntica d'una pàgina es mostra un quadre resum amb una llista de propietats i atributs. Aquesta llista es pot fer servir per navegar pel contingut semàntic de la wiki.
- **Facilitat per recuperar la informació:** el fet de disposar la informació en triplets RDF fa que es pugui fer servir un llenguatge de consulta com SPARQL per fer cerques d'una manera eficaç. A més, habitualment les wikis semàntiques disposen de pàgines especials que permeten fer cerques sense fer ús de cap llenguatge, simplement seleccionant la propietat que es vol consultar.
- **Extensions:** es poden integrar extensions per oferir noves funcionalitats.
- **Reutilització:** la correspondència de les dades d'una wiki semàntica amb components d'una ontologia facilita l'exportació i compartició del coneixement.

En quant a les debilitats, es poden sintetitzar en:

- **Ambigüïtat de les dades:** encara existeix ambigüïtat en les dades. L'ús de sinònims, homònims, etc, pot crear confusió en els computadors, per exemple, en no saber distingir si Java és un llenguatge de programació o una illa.
- **Seguretat i privacitat:** el fet que qualsevol pugui afegir i modificar la informació fa que aquesta sigui propensa a tenir errors, ser inexacta o fins i tot falsa.
- **Raonament limitat:** com a conseqüència dels dos punts anteriors, el raonament automàtic és encara poc efectiu.

L'addició de semàntica a les wikis ha fet que aquestes s'incorporin com una eina d'utilitat en múltiples àrees a causa de la seva millorada navegabilitat i recuperació i compartició de la informació. Entre les àrees d'aplicació que ja fan servir les wikis semàntiques, podem trobar:

- **Enginyeria del coneixement:** aquesta disciplina consisteix en la integració del coneixement en sistemes informàtics per tal de resoldre problemes complexos.
- **Enciclopèdies, diccionaris, col·leccions de documents,** com pot ser Wikipedia, un diccionari de biologia o informació d'una conferència.

- **Gestió de continguts i portals semàntics:** certs organismes gestionen els seus portals fent ús d'una wiki semàntica.
- **Gestió d'informació personal:** pràctica i estudi de les activitats que una persona (o un grup) realitza per adquirir, organitzar, mantenir i recuperar informació pel seu ús diari, com documents, pàgines web, missatges de correu electrònic, etc.
- **Aprenentatge:** compartició d'informació, activitats o pràctiques dins d'una comunitat.
- **Disseny de processos industrials:** determinar el flux de materials, necessitats de maquinari, etc, per un procés industrial.
- **eGovernment:** sovint s'han de gestionar una gran varietat de dades provinents de diferents administracions, on les ontologies poden jugar un paper determinant.
- **Intranet corporativa:** gestió de projectes i els seus recursos associats.
- **Coordinació d'equips:** els membres d'un projecte poden coordinar les seves activitats i anotar el seu progrés. Aquesta dada pot ser recuperada pel coordinador de l'equip per fer-ne un seguiment.
- **Desenvolupament de programari:** la documentació de multitud de projectes de codi obert es troba en una wiki.
- **Sistemes de suport a decisions:** sistemes basats en computador que acompanyen la resolució d'un problema ajudant en la presa de decisions. Es tracta d'un concepte lligat al Bussines Intelligence i la Minería de Dades.

En els següents apartats es fa un recull d'alguns projectes de motors wiki semàntics, descrivint les seves principals característiques, àrea d'aplicació, usuaris als quals van dirigits, etc. Existeix un llistat més complet i actualitzat a Semanticweb.org¹⁸.

3.1. AceWiki

Desenvolupada per la universitat de Zurich, AceWiki¹⁹ fa ús d'un llenguatge natural controlat anomenat Attempto Controlled English (ACE) per representar la informació. ACE és un subconjunt de l'anglès amb una gramàtica restringida i una semàntica formal, i el seu ús té dos avantatges respecte altres wikis semàntiques: primera, millora la usabilitat i l'aprenentatge, i segona, s'aconsegueix una millor expressivitat que altres llenguatges semàntics formals.

¹⁸ http://semanticweb.org/wiki/Semantic_wiki_projects

¹⁹ <http://attempto.ifi.uzh.ch/acewiki>



Figura 3-2. Interfície web d'AceWiki.

3.2. IkeWiki

Originalment, IkeWiki²⁰ va néixer com una eina per desenvolupar ontologies en un entorn col·laboratiu i per a tasques de gestió del coneixement. Està orientada a usuaris experts i disposa d'avançades funcionalitats semàntiques com les relacionades amb el raonament o l'enginyeria d'ontologies. Els desenvolupadors d'IkeWiki van donar molta importància a la facilitat d'ús i van desenvolupar una interfície basada en navegador web, molt semblant la de Wikipedia i amb un editor WYSIWYG.

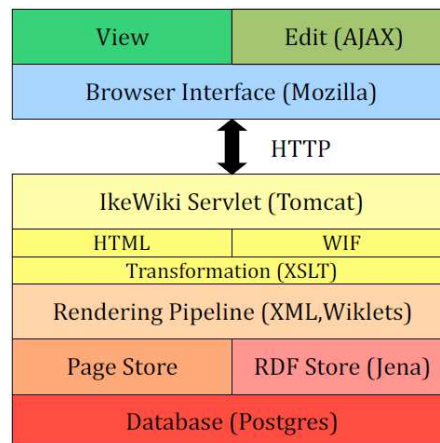


Figura 3-3. Arquitectura d'IkeWiki.

IkeWiki emmagatzema la informació en una base de dades Postgres però diferenciant dos magatzems: un per al contingut de les pàgines i un altre, basat en Jena²¹, per emmagatzemar triplets RDF amb les anotacions semàntiques. A l'hora de recuperar un recurs, el contingut de la pàgina i les anotacions es

²⁰ <http://semanticweb.org/wiki/IkeWiki>

²¹ <http://jena.apache.org/>

fusionen i es representen en XML amb la finalitat de ser exportades o bé transformades en HTML per visualitzar-les al navegador web.

Actualment el projecte IkeWiki està aturat, ja que va quedar substituït al març de 2008 pel projecte Kiwi²².

3.3. Kaukolu

Kaukolu²³ es fa servir com a banc de proves per enfocaments alternatius de mapejat de continguts wiki amb una base de coneixement. A diferència de la majoria de wikis semàntiques, on es fan servir triplets RDF per identificar els enllaços i el text, Kaukolu implementa una capa d'indirecció entre el model de pàgina wiki i el model RDF. Els recursos RDF es poden crear i adjuntar a seccions de text formant anotacions. L'usuari pot navegar fent servir tant l'entorn wiki estàndard com la capa d'anotacions segons les seves necessitats.

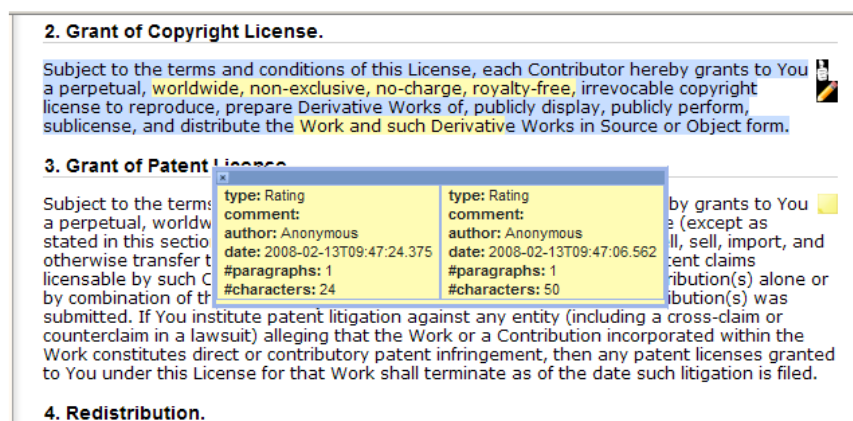


Figura 3-4. Finestra emergent amb anotacions de la pàgina.

3.4. KnowWE

KnowWe²⁴ és una wiki semàntica basada en el motor wiki de codi obert JSPWiki²⁵. Va ser creada per la universitat de Würzburg amb el propòsit de ser una eina orientada en la creació de sistemes de suport de decisions i de resolució de problemes relacionats amb ontologies i el coneixement.

KnowWE implementa un concepte (refinat) de les wikis semàntiques: a més del coneixement ontològic bàsic -com la definició de les classes, propietats taxonòmiques i definides per l'usuari- és capaç de representar coneixement per a la resolució de problemes que s'aplica a determinades classes de l'ontologia. Cada pàgina wiki incorpora, no només el text amb anotacions semàntiques, sinó també coneixement explícit per a la resolució de problemes.

²² <http://www.kiwi-project.eu/>

²³ <http://kaukoluwiki.opendfki.de/>

²⁴ <http://www.is.informatik.uni-wuerzburg.de/en/research/applications/knownwe/>

²⁵ <http://www.jspwiki.org/>

```

<Kopic id="SwimmingRules">
  <Rules-section>
    IF   (Training Goals = reducing stress) OR (Training Goals = endurance)
    THEN Swimming = SUGGESTED

    IF   Medical restrictions = allergy
    THEN Swimming = EXCLUDED
  </Rules-section>
</Kopic>

```

Figura 3-5. Coneixement explícit sobre natació a KnowWe.

3.5. Makna

Aquest motor es va crear com una eina per a l'enginyeria de coneixement distribuït. Al igual que KnowWE, està basat en el motor wiki JSPWiki, ampliant-lo amb diversos components per a la manipulació de dades semàntiques. En concret permet afegir, eliminar o modificar sentències RDF a través del framework Jena.

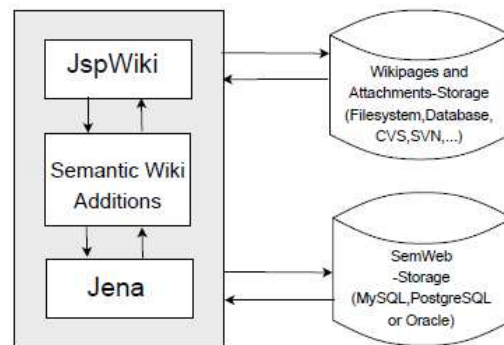


Figura 3-6. Arquitectura de Makna.

Una de les àrees on ha tingut aplicació ha sigut en el Workflow Management o gestió del flux de treball, tant en entorns industrials com de negocis o de desenvolupament de programari. Amb la inclusió de conceptes de flux de treball a un model semàntic es poden recollir els esdeveniments que sorgeixen dels diferents processos (inici de tasca, fi de tasca, etc) i es poden associar a accions personalitzades i aplicar inferència.

3.6. OntoWiki

OntoWiki²⁶ és una eina que pretén donar suport a l'enginyeria del coneixement en un entorn web. Proporciona sofisticats mitjans per a la navegació, visualització i creació de bases de coneixements basades en RDF. Les pàgines són configurables per l'usuari i contenen anotacions en forma de sentències RDF. D'aquesta manera OntoWiki permet crear diferents punts de vista de les dades per al consum humà (com ara pàgines wiki, representacions tabulars o mapes) i per al consum computacional (documents RDF, SPARQL).

²⁶ <http://ontowiki.net/Projects/OntoWiki>

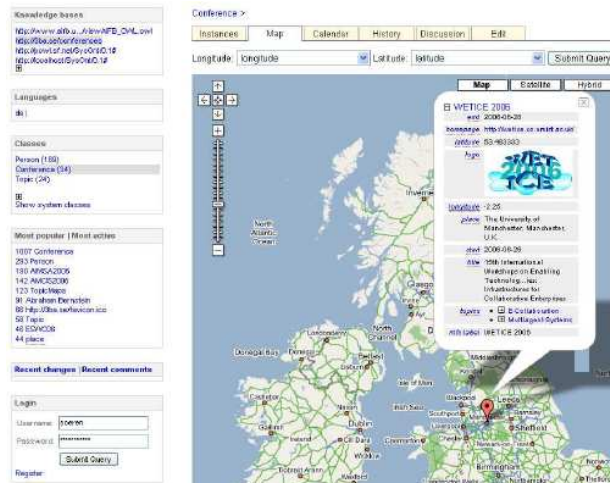


Figura 3-7. Vista d'una instància en forma de mapa.

OntoWiki proveeix diversos mètodes per navegar a través de les bases de coneixement, com la navegació taxonòmica i jeràrquica, la basada en facetes, i la cerca per text. Els mètodes es poden combinar entre ells, de manera que, per exemple, un usuari pot començar amb una cerca de text i refinar els resultats restringint-los a una determinada jerarquia.

3.7. Semantic MediaWiki

Semantic MediaWiki²⁷ (SMW) és una extensió de MediaWiki²⁸ – un dels motors wiki més coneguts arreu, especialment per ser sobre el qual s'ha construït Wikipedia – que afegeix anotacions semàntiques millorant la navegació, cerca, organització avaluació i compartició del contingut d'una wiki. El seu propòsit inicial va ser el de crear una Wikipedia semàntica, i amb el temps va esdevenir molt popular entre usuaris no experts.

El seu funcionament es basa en anotar els enllaços d'una pàgina (enllaços classificats) i el seu text (atributs), i d'aquesta manera crear sentències RDF on cada pàgina es converteix en un recurs ontològic. SMW mapeja aquestes anotacions en una descripció formal fent servir el llenguatge OWL DL. Les anotacions es troben disponibles per a l'usuari en forma d'un quadre resum (anomenat factbox) al peu de la pàgina visualitzada. Aquest quadre, a més, permet que es pugui navegar d'una forma semàntica pel contingut de la wiki i exportar les anotacions en un document OWL/XML.

Aquest mode de funcionament per mitjà d'anotacions no requereix de cap ontologia definida, i els usuaris poden anar afegint anotacions a mida que ho requereixen sense restriccions, el que fa que SMW no suporti funcionalitats avançades de raonament.

²⁷ <http://semantic-mediawiki.org/>

²⁸ <http://www.mediawiki.org/wiki/MediaWiki/es>

Facts about "Berlin" ⓘ		RDF feed
Area	891.85 km² (344.343 sqmi) +	
Average rainy days	106.3 +	
Capital of	Germany +	
Has coordinates	52° 30' 60" N, 13° 24' 0" E +	
Located in	Germany +	
Population	3,391,409 +	

Figura 3-8. Factbox a una pàgina sobre Berlín.

3.8. SweetWiki

SweetWiki²⁹ és una eina de treball col·laboratiu basada en un navegador web que combina l'etiquetatge social o folcsonomia amb ontologies formals. Els usuaris poden anotar pàgines amb etiquetes arbitràries, que al seu torn poden estar associades amb conceptes de les ontologies subjacents.

Disposa d'un editor WYSIWYG que habilita, des del mateix navegador, editar el contingut d'una pàgina, així com realitzar-ne anotacions. Mentre l'usuari està introduint una anotació, un mecanisme d'autocompletat proposa paraules clau existents llançant consultes SPARQL al servidor web semàntic per tal de trobar conceptes compatibles amb les etiquetes, i mostra el número de les pàgines que incorporen aquest concepte.

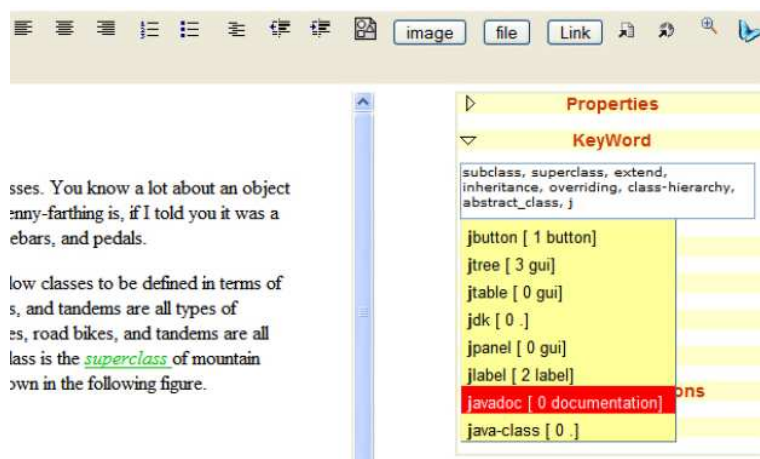


Figura 3-9. Sistema d'auto-completat de SweetWiki.

SweetWiki també proposa un enfocament mixt per tal d'organitzar les etiquetes, vinculant-les dins d'una folcsonomia descrita fent servir llenguatges de la web semàntica, on les etiquetes s'organitzen en una jerarquia i es relacionen amb les altres amb relacions del tipus `subClassOf`, `seeAlso`, etc.

²⁹ <http://www-sop.inria.fr/teams/edelweiss/wiki/wakka.php?wiki=SweetWiki>

3.9. TaOPis

TaOPis³⁰ és una plataforma de codi obert per a l'allotjament de projectes per a comunitats que integra un motor wiki semàntic programat en el llenguatge ontològic FLORA-2. Aquest llenguatge és al seu torn un dialecte del llenguatge F-logic.

La wiki semàntica de TaOPis es basa en un marc orientat a objectes, on qualsevol tema en un fòrum, qualsevol entrada d'un blog, un usuari, etc, és considerat com un objecte genèric que pot ser especialitzat fent ús d'etiquetes amb valors d'atributs. Existeixen unes etiquetes especials, com `class` i `subclass`, que permeten definir una jerarquia de classes semàntica, i on els enllaços entre objectes es consideren relacions.

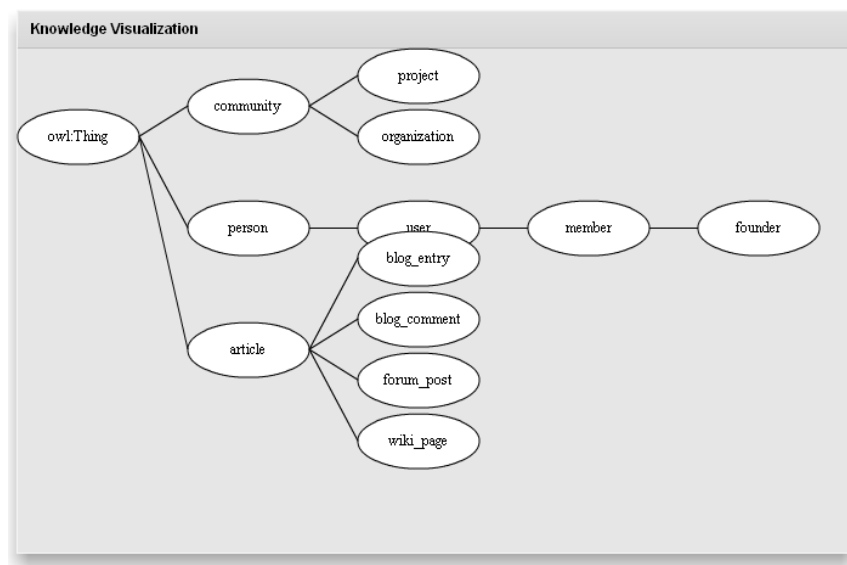


Figura 3-10. Visualització de la ontologia d'un projecte a TaOPis.

3.10. Taula comparativa

En aquest apartat es mostren dues taules resum amb l'objectiu de comparar les wikis analitzades als apartats anteriors.

En una primera taula (**Taula 3-1**) es mostren característiques de caràcter més general de cada motor, com el tipus de llicència, estat del projecte o últim llançament:

³⁰ <http://autopoiesis.foi.hr/>

	Llenguatge de programació	Llicència	Estat	Últim llançament	Mètode d'edició
AceWiki	Java	LGPL	Alpha	0.5.3 (19/01/2012)	Editor predictiu
IkeWiki	Java	GPL	Substituit	2.0 (05/03/2008)	Wiki-syntax
Kaukolu	Java			737 (30/10/2007)	Formulari
KnowWE	Java	LGPL	Estable	5.0 (30/05/2012)	Wiki-syntax
Makna	Java				Wiki-syntax
OntoWiki	PHP	GPL	Estable	0.9.5 (14/06/2010)	WYSIWYG
Semantic MW	PHP	GPL	Estable	1.7.1 (03/05/2012)	Wiki-syntax Formulari
SweetWiki	Java		Inactiu	1.7 (29/02/2008)	Wiki-syntax WYSIWYG Etiquetat
TaOPis	FLORA-2 Python PostgreSQL PHP	GPL	Prototipus		Wiki-syntax Etiquetat

Taula 3-1. Aspectes generals de les wikis semàntiques.

A la segona (**Taula 3-2**) es llisten aspectes més tècnics, i que es descriuen a continuació:

- **Adaptació de la presentació de continguts:** la forma en que es presenta el contingut de les pàgines pot variar en base a les anotacions.
- **Cerca semàntica:** cerca basada en la informació semàntica i no en el contingut de les pàgines.
- **Suport al raonament:** la wiki pot treure conclusions derivades del coneixement, que al seu torn deriva del contingut semàntic.
- **Importació/exportació d'ontologies:** capacitat d'exportar o importar ontologies per ser reutilitzades.
- **Canvi del model subjacent:** els usuaris poden canviar el model subjacent de coneixement que conté la wiki.

Un punt negre significa que la wiki incorpora aquesta característica, mentre que si el quadre roman blanc significa que, o bé no la incorpora, o bé es desconeix en el moment de redactar aquest document:

	Adaptació de la presentació de continguts	Cerca semàntica	Suport al raonament	Importació d'ontologies	Exportació d'ontologies	Canvi del model subjacent
AceWiki	●		●			
IkeWiki	●		●	●	●	●
Kaukolu				●	●	
KnowWE	●		●	●	●	●
Makna				●	●	
OntoWiki	●	●			●	●
Semantic MW	●	●	●	●	●	
SweetWiki						●
TaOPis		●	●		●	●

Taula 3-2. Aspectes tècnics de les wikis semàntiques.

4. Semantic MediaWiki

4.1. Instal·lació i configuració

Com s'ha comentat a l'apartat 3.7, Semantic MediaWiki (d'ara endavant SMW) és una extensió del motor MediaWiki, i per tant és requisit imprescindible tenir aquest correctament instal·lat i configurat prèviament. Addicionalment, MediaWiki necessita altres components per a funcionar, que en el moment de confeccionar aquest document són els següents:

- El **servidor web** Apache o IIS.
- El llenguatge d'scripts per al desenvolupament web **PHP**.
- Un **servidor de base de dades**, preferiblement MySQL.

Cal consultar les pàgines web de SMW³¹ i de MediaWiki³² per trobar aquesta informació actualitzada, ja que contínuament s'afegeixen servidors web i de base de dades compatibles amb MediaWiki.

La **Figura 4-1** mostra una visió de l'arquitectura emprada per SMW. Sota aquesta arquitectura, SMW captura i gestiona certs events. A més, en ser una extensió, SMW no sobreescriu cap part de MediaWiki, sinó que n'agrega funcionalitat. Per aquest motiu, SMW pot ser afegida a qualsevol wiki sense cap cost.

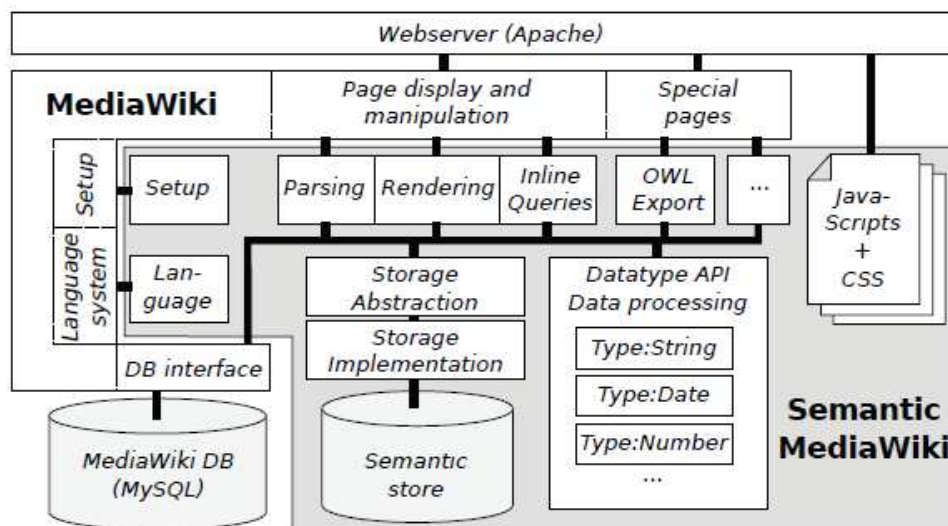


Figura 4-1. Arquitectura de SMW.

SMW i la resta de components poden funcionar sota els principals sistemes operatius. Això és, Linux, Windows, Mac OS, Solaris, NetWare i FreeBSD, si

³¹ <http://semantic-mediawiki.org/wiki/Help:Installation#Requirements>

³² http://www.mediawiki.org/wiki/Manual:Installation_requirements

més no és en equips sota Linux on la seva presència està més estesa i on s'obtenen millors resultats.

Per a la realització de les proves s'ha fet servir una màquina virtual VirtualBox amb sistema operatiu Linux Ubuntu. Sobre aquesta base s'ha construït una infraestructura coneguda com a LAMP (Linux, Apache, MySQL i PHP). Existeix força literatura que explica pas a pas com construir un sistema LAMP, p.ex. a la documentació d'Ubuntu³³.

Un cop instal·lada la infraestructura LAMP, Mediawiki i SMW, s'ha construït una petita wiki semàntica sobre música rock a Irlanda (grups, discos i músics), i que serà objecte d'estudi al llarg d'aquest capítol. A la figura següent se'n mostra la pàgina principal de la wiki i s'identifiquen els aspectes més importants de la interfície de MediaWiki.

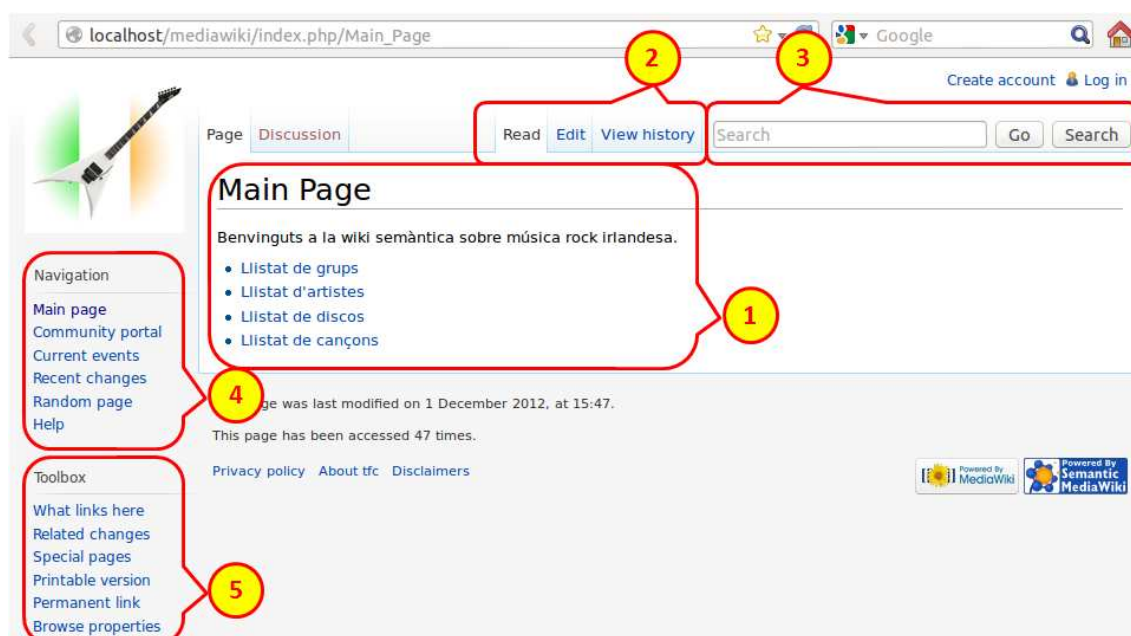


Figura 4-2. Pàgina principal de la wiki semàntica sobre rock irlandès.

A l'àrea central i més gran **(1)** es on es mostren els diferents articles, en aquest cas la pàgina principal. A la part de dalt hi ha una sèrie de pestanyes **(2)** que permeten veure l'article en mode lectura o d'edició, així com l'historial de canvis. Al costat **(3)** un camp de cerca de text que té la funcionalitat afegida d'habilitar l'accés directe a un article de la wiki amb el botó Go.

A l'esquerra, el menú Navigation **(4)** presenta alguns accessos directes a la pàgina principal, l'ajuda o l'històric de canvis de la wiki. I per últim el menú Toolbox **(5)** que conté funcionalitats importants. Destacar l'enllaç Special Pages, que accedeix a un recull de pàgines especials per realitzar informes, llistats, gestió d'usuaris, exportar dades, etc. Una altra funcionalitat important és la que ofereix l'enllaç Browse properties.

33

https://help.ubuntu.com/community/ApacheMySQLPHP#To_install_the_default_LAMP_stack_in_Ubuntu_10.04_and_above

En quant a la configuració, SMW permet un alt grau de configuració per mitjà d'un arxiu anomenat `LocalSettings.php` i que es troba dins la carpeta on s'instal·la MediaWiki. Aquest arxiu està ple de variables que es poden editar per adaptar el funcionament de MediaWiki i SMW a les nostres necessitats.

4.2. Afegir semàntica

Les wikis tradicionals com Wikipedia consten de pàgines amb text i amb enllaços a altres pàgines. A Mediawiki, per indicar que un text és un enllaç s'escriu entre claudàtors dobles. Així, si per exemple, a la pàgina que descriu el grup The Cranberries volem afegir un enllaç a la ciutat on es va formar, ho indiquem de la següent forma:

```
[[Limerick]]
```

Amb SWM és possible incloure informació semàntica a una pàgina afegint anotacions davant dels enllaços. Aquestes anotacions formen triplets RDF on el subjecte és la pròpia pàgina, l'objecte és la pàgina enllaçada, i el paper del predicat el fa l'anotació. En incloure una anotació el que s'està fent en realitat és especificar una propietat que estem descrivint, que en el nostre cas podria ser ben bé així:

```
[[format a::Limerick]]
```

A SMW no existeix una col·lecció de propietats definides, sinó que l'usuari té total llibertat a l'hora de crear-les. Així, en comptes de "format a" es podria fer servir "lloc de formació" o qualsevol altre nom per a la propietat. Un cop escollit el nom, però, és molt recomanable emprar sempre el mateix a tota la wiki.

Una altra forma d'afegir informació semàntica consisteix en anotar fragments de text. D'aquesta manera, si decidim que l'any de formació d'un grup de música és una propietat que cal indicar, ho podem fer així:

```
[[data formacio::1989]]
```

Aquesta anotació té l'inconvenient de que transforma el text 1989 en un enllaç a una pàgina que, tot i que es podria crear, no sembla tenir sentit. La solució passa per indicar-li a SMW que la propietat indica un número, és a dir, cal especificar el tipus de dada de la propietat, que ha de ser un dels tipus de dada predefinits a SMW. En el moment de la realització d'aquest document els tipus de dades disponibles són els següents:

- Annotation URI
- Boolean
- Code
- Date
- Email
- Geographic Coordinate

- Number
- Page
- Quantity
- Record
- String
- Telephone number
- Temperature
- Text
- URL

La utilització dels tipus de dades, a més d'evitar que el valor d'una propietat es converteixi en un enllaç, és important pels següents motius:

- Indica a les eines que llegeixen la propietat com han de manipular les dades, per exemple, a l'hora d'ordenar els resultats d'una cerca, ja que no és el mateix ordenar alfabèticament que numèricament.
- Relaciona diferents valors amb un mateix significat. Per exemple, els valors "1532", "1,532" i "1.532e3" corresponen al mateix número.
- Alguns tipus de dades tenen comportaments especials, com el tipus *Quantity*, que permet indicar les unitats en que es poden representar els valors numèrics (Kg, mm, etc.) i que esdevé bàsic per a fer conversions entre diferents unitats.

Per tal d'indicar el tipus de dada de la propietat *data formacio* cal, primer de tot, crear una pàgina que s'anomenarà *Property:data formacio*, per després afegir-hi aquest text:

```
[[Has type::number]]
```

Que fa ús d'una propietat especial predefinida a SMW anomenada *Has type*. Existeix un bon número de propietats especials³⁴ que tenen una utilitat interna i que no s'avaluen com la resta de propietats.

És possible definir propietats sense indicar-ne el seu tipus. De fet, per a la propietat *format* a no ho vàrem fer. En aquest cas el tipus per defecte és *Page*, és a dir, una pàgina de la wiki, que és precisament el que es pretenia. Tot i això és una bona pràctica definir sempre el tipus de dada d'una propietat per evitar confusions i per prevenir problemes futurs en el cas de que, per exemple, un administrador decideixi canviar el tipus de dada per defecte.



Figura 4-3. Factbox a la pàgina del grup The Cranberries.

³⁴ http://semantic-mediawiki.org/wiki/Help:Special_properties

Un cop afegides les anotacions i definits els tipus de dades de les propietats, a peu de pàgina de l'article sobre The Cranberries es farà visible el factbox (veure apartat 4.1.1 per més informació) que resumeix les propietats que s'han especificat, així com el valor d'aquestes (**Figura 4-3**). Cal comentar, però, que és possible que la visualització del factbox no vingui activada per defecte a SMW, i podria ser necessari modificar l'arxiu `LocalSettings.php` incloent-hi la següent línia:

```
$smwgShowFactbox = SMW_FACTBOX_NONEMPTY;
```

4.3. Semàntica basada en ontologies

Tot i que podem anar “inventant-nos” propietats per descriure allò del que estem parlant en els diferents articles de la wiki, la veritable potència de les anotacions es fa patent quan s'utilitzen termes que representen classes i propietats d'ontologies ja definides. De res serveix que la wiki s'ompli d'anotacions si aquestes no poden ser compreses des de l'exterior.

Sempre que sigui possible és convenient la reutilització d'ontologies ja definides, fet que millorarà notablement la compatibilitat de la nostra wiki amb altres eines. A semanticweb.org es pot trobar un llistat³⁵ de les ontologies més emprades a l'entorn de la web semàntica, però n'existeixen milers. Podem utilitzar un buscador ontològic com Swoogle³⁶ per cercar les ontologies que millor s'adeqüin a la nostra wiki.

Per al cas d'estudi que ens ocupa, i ja que el contingut de la wiki fa referència a conceptes musicals, s'ha optat per la reutilització de la ontologia Music Ontology³⁷. Aquesta ontologia defineix una gran quantitat de termes que descriuen tot allò que envolta el món de la música, com artistes, discs, instruments, events musicals, discogràfiques, etc.

Per tal d'utilitzar a SMW una ontologia existent, aquesta primer s'ha d'importar. La importació de la Music Ontologia a SMW requereix els següents passos:

Primerament, cal crear una pàgina especial amb el nom `Mediawiki:smw_import_namespace`, on “namespace” és una abreviació de l'espai de noms de la ontologia. Així, la pàgina creada per importar el vocabulari musical es dirà:

```
Mediawiki:smw_import_mo
```

Un cop creada la pàgina s'ha d'indicar dins el seu cos quin vocabulari estem important:

³⁵ <http://semanticweb.org/wiki/Ontology>

³⁶ <http://swoogle.umbc.edu/>

³⁷ <http://musicontology.com/>

`http://purl.org/ontology/mo|[http://musicontology.com/ The Music Ontology]`

Aquesta línia la podem desglossar en dos parts:

- **`http://purl.org/ontology/mo`** és exactament l'espai de noms que fa servir l'ontologia.
- Després del caràcter `|`, **`[http://musicontology.com/ The Music Ontology]`** és un enllaç a la pàgina web de l'ontologia on un usuari es pot adreçar per tal de recaptar més informació. El text "The Music Ontology" simplement presenta l'enllaç d'una forma més fàcilment llegible per a l'ésser humà.

Val a dir que no tots els termes definits a la ontologia (classes i propietats) s'importen automàticament un cop indicat l'espai de noms, sinó que s'han d'especificar de manera explícita quins termes s'importaran i, un cop importats, de quina manera seran utilitzats a la wiki. Així, cada classe de la ontologia es converteix en una **Categoria** de SMW amb el text `Nom_de_Classe|Category`, mentre que una propietat equival a una **propietat** de SMW, que serà d'un dels tipus de dades predefinits amb la sentència `Nom_de_Propietat|Type:tipus_de_dada`.

Un cop analitzada la ontologia, s'ha decidit fer servir les categories (classes) `Record`, `Track`, `MusicGroup`, `SoloMusicArtist` i `MusicArtist`, i les propietats `track`, `track_number`, `member`, `member_of`, `homepage`, `produced`, `producer` i `primary_instrument`.

A la següent figura es pot veure detalladament com queda la pàgina Mediawiki:smw_import_mo un cop editada:

MediaWiki:Smw import mo

[http://purl.org/ontology/mo/%7C%5Bhttp://musicontology.com/The Music Ontology](http://purl.org/ontology/mo/%7C%5Bhttp://musicontology.com/The%20Music%20Ontology%5D)

```
Record|Category
Track|Category
MusicGroup|Category
MusicArtist|Category
SoloMusicArtist|Category
track|Type:Page
track_number|Type:Number
member|Type:Page
member_of|Page
homepage|Type:URL
produced|Type:Page
producer|Type:Page
primary_instrument|Type:String
```

Figura 4-4. Importació de la Music Ontology a SMW.

Per últim cal que creem una pàgina nova per a cada terme importat (propietats i categories) i fer saber a SMW que aquest es correspon exactament amb un de la ontologia. Per a fer-ho disposem de la propietat especial `imported from`.

Així, per exemple, a la pàgina `Property:track_number` podem afegir el següent:

```
Propietat importada de [[imported from::mo:track_number]].  
Indica el número d'una cançó dins un disc.
```

Com vàrem comentar a l'apartat 4.2, en definir una propietat s'ha d'indicar el seu tipus, que ha de ser un dels tipus de dades predefinits a SMW. En el cas de les propietats importades d'un vocabulari extern aquesta indicació resulta redundant i innecessària, ja que el tipus de dada s'ha indicat en importar el terme. Recordem:

```
track_number|Type:Number
```

Un cop creades les pàgines per a tots els termes ja estem en condicions d'afegir anotacions semàntiques a la wiki. És important comentar que, a més, per a cada pàgina que sigui instància d'una classe de la ontologia s'ha d'afegir l'etiqueta `[[Category:nom_de_categoria]]` per tal que SMW la reconegui com a tal.

Igualment cal fer un incís en que, si bé SMW permet l'ús de qualsevol propietat amb qualsevol categoria, sembla obvi entendre que aquestes es facin servir d'acord a la seva especificació a la ontologia, més concretament, d'acord al seu domini. Altrament es perdria compatibilitat amb les diferents eines coneixedores de la ontologia. Per tant, més enllà del fet que en el món real no té sentit aplicar la propietat `track_number` a un grup musical, conceptualment aquesta no pertany al domini de la classe `MusicGroup`.

SMW permet treballar amb més d'un vocabulari extern. Per tal de testar aquest fet s'ha decidit incorporar també la ontologia FOAF³⁸ (Friend of a Friend), per tant s'ha creat la pàgina `Mediawiki:smw_import_foaf` i s'ha importat el terme `name`.

A la **Figura 4-5**, a mode d'exemple, es pot veure un dels articles creats a la wiki. Es tracta de l'article sobre el grup de música U2. A l'esquerra es presenta la pàgina en mode lectura, mentre que a la dreta es troba en mode edició. S'hi pot veure la utilització d'algunes propietats importades de la Music Ontology, així com la propietat `name` de FOAF. Noteu com, a la última línia, es fa saber a SMW que l'article és una instància de la classe `MusicGroup`.

³⁸ <http://xmlns.com/foaf/spec/>

U2

U2 és un grup de rock de Dublín.

Pàgina Oficial [\[edit\]](#)

- <http://www.u2.com/> [\[edit\]](#)

Discografia [\[edit\]](#)

1. [Boy](#) (1980)
2. [October](#) (1981)
3. [War](#) (1983)
4. [The Unforgettable Fire](#) (1984)
5. [The Joshua Tree](#) (1987)
6. [Rattle and Hum](#) (1988)
7. [Achtung Baby](#) (1991)
8. [Zooropa](#) (1993)
9. [Pop](#) (1997)
10. [All That You Can't Leave Behind](#) (2000)
11. [How to Dismantle an Atomic Bomb](#) (2004)
12. [No Line on the Horizon](#) (2009)



Membres [\[edit\]](#)

Membres actuals

- Bono
- The Edge
- Adam Clayton
- Larry Mullen

Members antics

- Dick Evans

Category: MusicGroup

Facts about U2 ⓘ		RDF feed ⓘ	
Homepage	http://www.u2.com/ [edit] [info]		
Member	Bono [edit] [info] , The Edge [edit] [info] , Adam Clayton [edit] [info] and Larry Mullen [edit] [info]		
Name	U2 [edit] [info]		
Produced	Boy [edit] [info] , October [edit] [info] , War [edit] [info] , The Unforgettable Fire [edit] [info] , The Joshua Tree [edit] [info] , Rattle and Hum [edit] [info] , Achtung Baby [edit] [info] , Zooropa [edit] [info] , Pop [edit] [info] , All That You Can't Leave Behind [edit] [info] , How to Dismantle an Atomic Bomb [edit] [info] and No Line on the Horizon [edit] [info]		

Editing U2

Warning: You are not logged in.

Your IP address will be recorded in this page's edit history.



```
'''[[name::U2]]''' és un grup de rock de Dublín.  
  
[[image:U2.jpg|frame|right|200x200px]]  
  
==Pàgina Oficial==  
  
* [[homepage::http://www.u2.com/]]  
  
==Discografia==  
  
#[[produced::Boy]] (1980)  
#[[produced::October]] (1981)  
#[[produced::War]] (1983)  
#[[produced::The Unforgettable Fire]] (1984)  
#[[produced::The Joshua Tree]] (1987)  
#[[produced::Rattle and Hum]] (1988)  
#[[produced::Achtung Baby]] (1991)  
#[[produced::Zooropa]] (1993)  
#[[produced::Pop]] (1997)  
#[[produced::All That You Can't Leave Behind]] (2000)  
#[[produced::How to Dismantle an Atomic Bomb]] (2004)  
#[[produced::No Line on the Horizon]] (2009)  
  
==Membres==  
  
;Membres actuals  
* [[member::Bono]]  
* [[member::The Edge]]  
* [[member::Adam Clayton]]  
* [[member::Larry Mullen]]  
  
;Members antics  
* [[Dick Evans]]  
  
[[Category:MusicGroup]]
```

Figura 4-5. Article sobre el grup U2 a l'esquerra.
A la dreta les anotacions semàntiques.

4.4. Navegació semàntica

SMW ofereix a l'usuari diversos mètodes de navegació fent servir la informació semàntica en comptes del contingut textual de les seves pàgines. Tot seguit se'n fa un recull dels més importants.

4.4.1. Factbox

Com s'ha vist al principi d'aquest capítol, en incloure anotacions semàntiques a un article apareix a la part inferior d'aquest un factbox (veure **Figura 4-3** i **Figura 4-5**). Un factbox no és més que un quadre que resumeix tota la informació semàntica que conté la pàgina. A més, ofereix algunes utilitats interessants.

La informació semàntica dins el factbox es presenta en dues columnes: a l'esquerra apareixen les propietats, mentre que a la dreta hi tenim els valors que prenen. Cada nom de propietat és també un enllaç que porta a la pàgina que descriu aquesta propietat. Per exemple, i prenent com a base el factbox de l'article sobre el grup U2, si es fa clic sobre la propietat `member` s'accedeix a la pàgina `Property:Member`. Observeu que aquesta ara presenta informació sobre el seu ús dins la wiki (**Figura 4-6**).

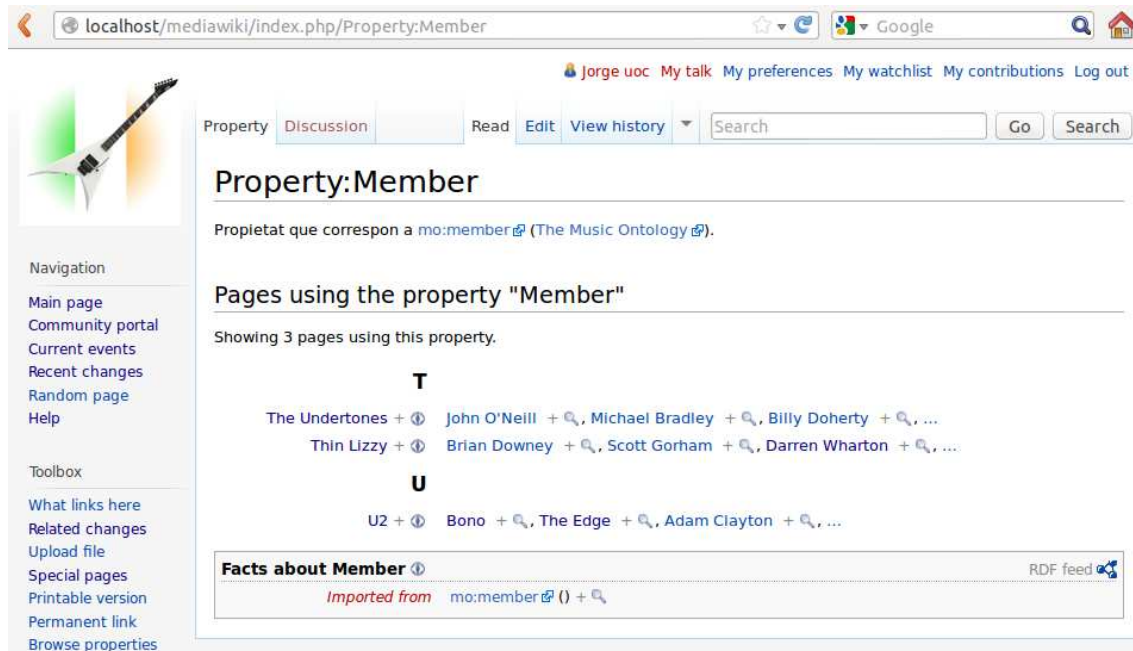


Figura 4-6. Pàgina de la propietat Member.

La informació sobre l'ús de la propietat torna a ser presentada en dues columnes: a l'esquerra les pàgines que la incorporen i a la dreta els valors que pren en cada pàgina. Entre totes dues columnes la icona ⓘ dóna accés a una pàgina especial anomenada `Special:Browse` i de la que es parla a l'apartat 4.4.2 amb més deteniment. Noteu que la pàgina `Property:member` també presenta el seu factbox.

Tornant al factbox de l'article sobre U2, i concretament als valors que prenen les propietats, en funció del seu tipus aquestes poden ser enllaços a altres pàgines de la wiki (propietat de tipus `Page`), enllaços externs (URL o URI), o simplement un valor (`String`, `Number`, etc). Especialment d'utilitat és la icona + ⓘ a la dreta de cada valor d'una propietat. Aquesta icona mostra el resultat d'una cerca que recull totes les pàgines de la wiki que contenen aquesta propietat amb exactament aquest valor. De fet, la icona és en realitat un enllaç a una altra pàgina especial anomenada `Special:SearchByProperty`, (veure apartat 4.4.3) que permet fer cerques d'aquest tipus.

A la part superior del factbox es troben possiblement dues de les seves característiques més importants: la icona ⓘ que enllaça amb la pàgina

especial `Special:Browse` i la icona **RDF feed** que porta a la pàgina especial `Special:ExportRDF` (veure apartat 4.7).

4.4.2. Special:Browse

Es tracta d'una pàgina especial que presenta totes les propietats i els seus valors d'una entitat. A diferència del factbox no només recull les propietats que es troben a la pròpia pàgina (propietats sortints), sinó també les que des d'altres pàgines enllacen amb aquesta (propietats entrants).

Algunes formes d'accedir a la pàgina `Special:Browse` són fent clic a la icona ⓘ que apareix a la part superior del factbox, també des de l'enllaç "Browse properties" del menú "Toolbox", o escrivint "`Special:Browse`" directament a la barra de direccions del navegador. Amb aquest últim mètode es mostra un petit formulari on s'ha d'entrar manualment el nom de l'entitat que es vol explorar.

The screenshot shows a web browser window with the URL `localhost/mediawiki/index.php/Special:Browse/U2`. The page title is "Special page" and the main heading is "Browse wiki". The entity being browsed is "U2".

Homepage	http://www.u2.com/
Member	Bono ⓘ, The Edge ⓘ, Adam Clayton ⓘ, Larry Mullen ⓘ
Name	U2 ⓘ
Produced	Boy ⓘ, October ⓘ, War ⓘ, The Unforgettable Fire ⓘ, The Joshua Tree ⓘ, Rattle and Hum ⓘ, Achtung Baby ⓘ, Zooropa ⓘ, Pop ⓘ, All That You Can't Leave Behind ⓘ, How to Dismantle an Atomic Bomb ⓘ, No Line on the Horizon ⓘ
Categories	MusicGroup ⓘ
Modification date	27 November 2012 18:55:31 ⓘ

hide properties that link here

Adam Clayton ⓘ, Bono ⓘ, Larry Mullen ⓘ, The Edge ⓘ	Member of
Achtung Baby ⓘ, All That You Can't Leave Behind ⓘ, Boy ⓘ, How to Dismantle an Atomic Bomb ⓘ, No Line on the Horizon ⓘ, October ⓘ, Pop ⓘ, Rattle and Hum ⓘ, The Joshua Tree ⓘ, The Unforgettable Fire ⓘ, War ⓘ, Zooropa ⓘ	Producer

Enter the name of the page to start browsing from.



Figura 4-7. Pàgina `Special:Browse` amb les propietats d'U2.

Des d'aquesta pàgina especial es pot navegar a través de la xarxa de dades que es creen entre les entitats de la wiki fent clic a la icona ⓘ d'aquelles propietats que siguin de tipus pàgina. Per a la resta de propietats la icona dona accés a la pàgina `Special:SearchByProperty`.

Les propietats entrants suposen una clara millora respecte les wikis tradicionals en mostrar una informació que d'altra manera seria impossible de recuperar, a

no ser que estiguéssim disposats a llegir una gran quantitat d'articles en busca d'aquesta informació.

4.4.3. Special:SearchByProperty

Des d'aquesta pàgina especial es pretèn fer una cerca de totes aquelles entitats que continguin una determinada propietat i amb un valor determinat. Es pot accedir-hi escrivint `Special:SearchByProperty` directament a la barra de direccions del navegador. Un cop allà s'introdueix a un senzill formulari la propietat i el valor que es vol buscar. No obstant, la manera més senzilla i ràpida es fent clic en les icones   que es troben als factbox o a la pàgina `Special:Browse`.

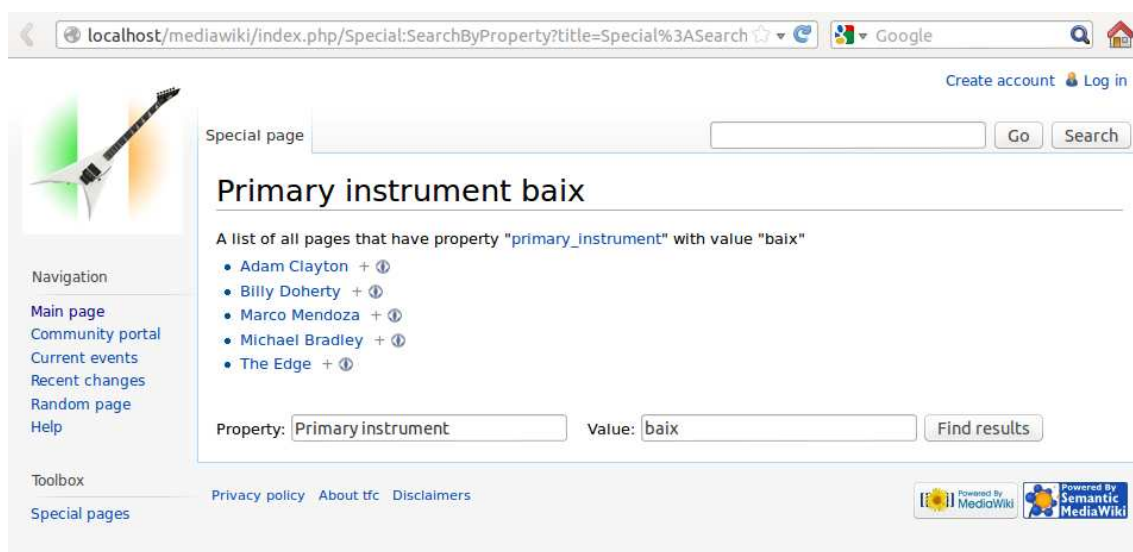


Figura 4-8. Pàgina `Special:SearchByProperty` per la propietat `primary_instrument` amb el valor “baix”.

A la **Figura 4-8** s'ha fet servir aquesta pàgina per buscar tots els artistes de la wiki que toquen el baix. Noteu l'ús de la propietat i el seu valor al formulari.

4.5. Cerca semàntica

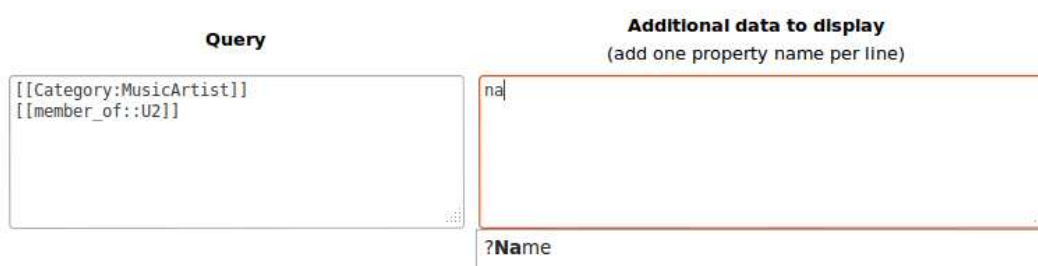
Per tal d'aprofitar encara més la informació semàntica afegida mitjançant anotacions, SMW inclou un llenguatge de consultes que permet accés “a mida” a aquest coneixement. La sintaxis d'aquest llenguatge és molt similar a la utilitzada en les anotacions i no requereix un gran esforç d'aprenentatge.

En una consulta semàntica s'especifiquen dues coses: quines pàgines es vol seleccionar i quina informació es vol obtenir d'aquestes pàgines. Indicant una o més condicions del que es demana, es poden seleccionar pàgines pel seu nom, espai de noms, categoria i, probablement el més important, pel valor de les seves propietats. Així, per exemple, la consulta per seleccionar tots els artistes que formen part del grup U2 seria la següent:

```
[[Category:MusicArtist]] [[Member_of::U2]]
```

La manera més senzilla de realitzar una consulta és mitjançant la pàgina especial `Special:Ask`. Aquesta disposa d'un formulari que dona assistència a l'usuari a l'hora de realitzar les consultes. A la part superior del formulari es poden trobar dos camps. El primer, anomenat "Query", fa la funció de seleccionar pàgines, mentre que el segon serveix per establir les propietats que es volen visualitzar de cada pàgina seleccionada. Aquest segon camp té la interessant característica de que en començar a escriure el nom d'una propietat es mostra un llistat d'opcions a escollir, com es pot veure a la **Figura 4-9**. Noteu que es pot indicar més d'una propietat al mateix temps, una per línia i sempre precedides del caràcter `?`.

Semantic search



The image shows the 'Special:Ask' interface for semantic search. It has two main sections: 'Query' and 'Additional data to display'. The 'Query' section contains a text box with the query '[[Category:MusicArtist]] [[member_of::U2]]'. The 'Additional data to display' section has a text box with 'na' and a dropdown menu showing '?Name'. Below the 'Additional data to display' text box is a small box with the text '?Name'.

Figura 4-9. Consulta semàntica amb `Special:Ask`.

Sota aquests dos camps es troben algunes opcions de configuració que es poden establir en la consulta. Per exemple, es pot indicar com ordenar els resultats, el màxim de resultats a mostrar, etc. Una menció important mereix l'apartat "Format as". Aquest permet escollir el format en que es vol obtenir el resultat de la consulta. El format per defecte és "Broad table", la típica taula de MediaWiki, però es pot escollir mostrar les dades en forma de taula simple, llistat, CSV, o RDF. Aquest últim genera un fitxer OWL/RDF amb les dades obtingudes. Veure l'apartat 4.7 per més informació.

A l'exemple anterior s'ha creat una consulta a partir d'una categoria concreta i una propietat concreta amb un valor concret (artistes membres d'U2, dels quals volem obtenir el nom). En molts casos, però, un usuari no busca només un valor concret, sinó tot un seguit de valors que compleixin certes condicions.

El llenguatge de consultes de SMW permet la utilització de **comparadors** i **comodins**. Els comodins s'escriuen com "+" i permeten qualsevol valor per una condició donada. Per exemple, `[[member_of::+]]` retorna totes les pàgines que inclouen aquesta propietat sigui quin sigui el seu valor.

Per la seva banda els comparadors són símbols que es col·loquen després dels caràcters `::` a les condicions de les propietats i que restringeixen els resultats. Es pot fer ús dels següents comparadors:

Comparador	Descripció	Exemple
>> i <<	'Major que' i 'menor que'	[[track_number::>>10]]
> i <	Per defecte 'Major o igual que' i 'menor o igual que', encara que es pot canviar el seu ús a l'arxiu LocalSettings.php	[[track_number::>10]]
≥ i ≤	'Major o igual que' i 'menor o igual que'	[[track_number::≥10]]
!	Distint de	[[track_number::!10]]
~	Comparació de cadenes de text, similar a like de SQL	[[name::~~David*]]
!~	Distint de per a cadenes de text	[[member_of::~!~U2]]

Taula 4-1. Comparadors disponibles en les consultes.

És també possible unir diferents consultes amb l'ús de l'operador OR. Així, podríem obtenir totes aquelles entitats que siguin, o bé cançons o bé discs, amb:

```
[[Category:Track]] OR [[Category:Record]]
```

L'operació "o" lògica també es pot aplicar per unir consultes que impliquin més d'un valor d'una propietat. En aquest cas, però, no es fa servir l'operador OR sinó els caràcters ||. La següent consulta recull els artistes que fan servir el teclat o el sintetitzador, o bé tots dos alhora:

```
[[primary_instrument::teclat||sintetitzador]]
```

Que és equivalent a:

```
[[primary_instrument::teclat]] OR  
[[primary_instrument::sintetitzador]]
```

Suposem ara que un usuari vol conèixer els músics que formen part del grup que va produir el disc "Positive Touch". Podríem executar una primera consulta que retornés el nom del grup per després utilitzar aquest resultat en una segona consulta i obtenir els seus membres. No obstant el nom del grup no ens interessa i, a més, hem hagut de fer dues consultes per obtenir la informació desitjada. SMW permet niar consultes, d'aquesta manera les dades que busquem les podem aconseguir amb la següent instrucció:

```
[[Category:MusicArtist]]  
[[member_of::<q>[[produced::Positive Touch]]</q>]]
```

La consulta anterior es pot escriure d'un mode més curt encadenant propietats:

```
[[Category:MusicArtist]]  
[[member_of.produced::Positive Touch]]
```


La pàgina `Special:Ask` presenta una interfície còmoda per crear consultes, però no és la única manera. Una consulta pot ser creada manualment i inserida en qualsevol pàgina de la wiki, el que a l'entorn de SMW rep el nom de **consulta en línia**. Els usuaris que llegeixin un article amb una consulta en línia no la poden veure, però sí el seu resultat.

Les consultes en línia són especialment útils quan es vol mostrar una informació que pot variar amb el temps. Per exemple, a l'article sobre un grup de música es podria fer servir per mostrar el llistat dels seus membres. Cada cop que s'accedeix a la pàgina la consulta s'executa i mostra la informació més recent. Així, si un membre marxa del grup, només cal esborrar l'anotació `member_of` de la seva pàgina per a que s'actualitzi la informació del grup. Les consultes en línia també són un mètode eficaç per evitar errors i inconsistència de les dades.

Existeixen dos tipus de consultes en línia, la funció `#ask` i la funció `#show`. La primera és la més completa de les dues, i funciona de la mateixa manera que les consultes creades amb el formulari d'`Special:Ask`. De fet, `Special:Ask` crea una consulta d'aquest tipus i en mostra el resultat.

El següent és un exemple d'utilització de funció `#ask` que selecciona els artistes de la wiki que formen part del grup U2 i en mostra el seu nom en una taula:

```
{{#ask: [[Category:MusicArtist]] [[member_of::U2]]
|?name = Nom
|format = table
}}
```

Es pot apreciar que els paràmetres que es fan servir són molt similars als que havíem vist fins ara:

- Primerament s'indiquen les pàgines que volem seleccionar, en aquest cas artistes que incorporin la propietat `member_of` amb el valor "U2".
- A la segona línia indiquem que la informació que volem extreure de cada pàgina és el nom de l'artista. Opcionalment podem especificar el text de capçalera que tindrà la columna a la taula amb "=".
- Finalment, a la tercera línia indiquem que el resultat el volem en forma de taula.

En realitat no és necessari escriure cada paràmetre en una línia diferent ni en l'ordre en el que s'ha fet, però és recomanable per facilitar-ne la lectura. Tanmateix sí és necessari l'ús del caràcter `|` per separar els arguments de la consulta. Un cop executada la instrucció, el resultat és el següent:

	Nom
Adam Clayton	Adam Clayton
Bono	Paul David Hewson
Larry Mullen	Larry Mullen
The Edge	David Howell Evans

Figura 4-10. Membres d'U2 extrets amb la funció #ask.

Totes les opcions que ofereix el formulari de la pàgina `Special:Ask` es poden aplicar a una funció #ask, això és, ordenar el resultat amb el paràmetre `sort` i indicar l'ordre d'ordenació amb `order`, limitar el resultat amb `limit`, etc. Al manual en línia de SMW es pot trobar el llistat complet dels paràmetres disponibles³⁹.

Per la seva banda, un ús típic de la funció #show es és el de mostrar el valor d'una propietat d'una determinada pàgina. Es pot pensar en #show com en una simplificació d'#ask per a fer servir en aquests casos. Per exemple, es podria voler inserir en un paràgraf l'instrument que fa servir en “Marco Mendoza” amb la instrucció:

```
{{#show: Marco Mendoza | ?primary_instrument}}
```

El resultat és una única paraula, en aquest cas “baix”, i per tant no cal indicar-ne cap format de sortida a la consulta.

4.6. Inferència

Com s'ha pogut contrastar en el capítol anterior, la cerca semàntica s'empra per a trobar pàgines en base a anotacions semàntiques introduïdes pels usuaris. Si les anotacions no s'han fet o són incorrectes la cerca semàntica no funcionarà com un es podria esperar. Per la seva banda, el terme inferència fa referència al fet que es pugui recuperar informació que no ha sigut afegida explícitament pels usuaris però que són deduïdes pel motor semàntic.

Suposem que un usuari crea la pàgina del músic Gary Moore i afegeix algunes anotacions semàntiques sobre el seu nom i el seu instrument. No obstant, l'usuari oblida incloure informació sobre la categoria, és a dir, la pàgina no conté l'etiqueta `[[Category:MusicArtist]]`. Amb una cerca semàntica que s'apliqués a entitats de tipus `MusicArtist`, Gary Moore no es trobaria entre els resultats. Tanmateix, un motor wiki semàntic amb capacitats d'inferència, en analitzar la pàgina sobre Gary Moore i trobar que conté l'etiqueta `primary_instrument` podria fer el següent raonament:

- La propietat `primary_instrument` pertany al domini de la classe `MusicArtist`.
- Gary Moore conté la propietat `primary_instrument`.

³⁹ http://semantic-mediawiki.org/wiki/Help:Inline_queries

- Per tant, Gary Moore és una instància de `MusicArtist`.

Lamentablement la capacitat d'inferència dels motors wiki semàntics és encara molt limitada i normalment no són capaços d'arribar a aquestes conclusions. SMW no és una excepció, i l'única capacitat d'inferència que suporta està basada en jerarquies de categories i de propietats. Aquesta inferència és molt senzilla: qualsevol cerca que inclogui una categoria específica retornarà també aquelles pàgines que pertanyin a una subcategoria d'aquesta categoria.

Amb MediaWiki és possible treballar amb jerarquies de categories. Per a fer-ho cal que a la pàgina que descriu una subcategoria s'afegeixi una etiqueta que indiqui la seva supercategoria. En el cas de la nostra wiki disposem de la categoria `SoloMusicArtist` que, segons la especificació de la Music Ontology és una subclasse de `MusicArtist`. Així doncs només cal modificar la pàgina `Category:SoloMusicArtist` per tal d'afegir la següent etiqueta:

```
[[Category:MusicArtist]]
```

Tornant a l'exemple sobre Gary Moore, com que la seva carrera musical es va desenvolupar majorment en solitari, sembla més adient incloure'l en la categoria `SoloMusicArtist`, ja que aquesta és més específica que `MusicArtist` (**Figura 4-11**).

Editing Gary Moore

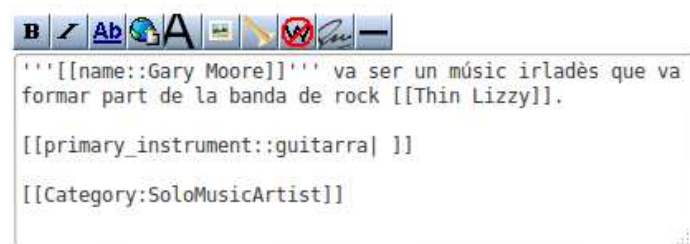


Figura 4-11. Gary Moore és una instància de `SoloMusicArtist`.

A partir d'aquest moment la pàgina sobre Gary Moore es trobarà en els resultats de tota consulta que involucri entitats de tipus `MusicArtist`.

En quant a les capacitats d'inferència basades en jerarquies de propietats són molt similars, ja que amb SMW es poden definir subpropietats. Per exemple, la Music Ontology incorpora la propietat `download` que relaciona una cançó o un disc amb una web des d'on es pot descarregar. La ontologia defineix també tres especialitzacions d'aquesta en funció de si la descàrrega pot ser gratuïta (`freedownload`), és una previsualització (`previewdownload`) o és de pagament (`paiddownload`).


Per a especificar en SMW que `freedownload` és una subpropietat de `download`, només cal que a la pàgina `Property:download` s'afegeixi:

```
[[subproperty of::Property:download]]
```

Així, qualsevol consulta que incorpori la propietat `download` retornarà també les pàgines que continguin la propietat `freedownload`.

4.7. Exportació de les dades

A partir de les anotacions semàntiques en la wiki, SMW pot generar documents en format OWL/RDF amb la finalitat d'exportar les dades i que aquestes puguin ser llegides i processades per computadors. Per a això disposa d'una pàgina especial anomenada `Special:ExportRDF`, que conté un petit formulari on indicar quines pàgines es volen exportar.

Existeix una dreuera per generar aquest fitxer directament des dels articles de la wiki: es tracta de la icona **RDF feed**  que es troba a la cantonada superior dreta del factbox. Fent clic en aquest enllaç el fitxer es genera automàticament a partir de les anotacions semàntiques de l'article i és visualitzat en el navegador. A la **Figura 4-12** es pot veure el fitxer OWL/RDF que SMW ha creat a partir de l'article sobre el cantant d'U2, i que tot seguit s'analitza:

```
<?xml version="1.0" encoding="UTF-8"?>
```

La primera línia indica que es tracta d'un arxiu XML, i és que tot arxiu OWL/RDF està basat en aquesta sintaxis.

```
<!DOCTYPE rdf:RDF [  
  <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>  
  <!ENTITY rdfs 'http://www.w3.org/2000/01/rdf-schema#'>  
  <!ENTITY owl 'http://www.w3.org/2002/07/owl#'>  
  <!ENTITY swivt 'http://semantic-mediawiki.org/swivt/1.0#'>  
  ...  
>
```

El bloc `<!DOCTYPE>` és un DTD intern que defineix algunes entitats. Les entitats són variables especials que substitueixen algunes cadenes de text. P.ex., cada cop que apareix el text `&owl;` el podem substituir per `http://www.w3.org/2002/07/owl#`. Tot i que es defineixen, les entitats no es fan servir al llarg del document.

```
<rdf:RDF  
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"  
  xmlns:owl="http://www.w3.org/2002/07/owl#"  
  xmlns:swivt="http://semantic-mediawiki.org/swivt/1.0#"  
  ...
```

L'element arrel amb que comença tot document OWL/RDF, indicador de que es tracta de la representació escrita d'un graf RDF. Defineix una sèrie d'espais de noms però amb un nou significat. Mentre que en un document XML tradicional els espais de nom serveixen per evitar ambigüitats, en RDF indiquen vocabularis existents a la web i que són importats i utilitzats al document.



Figura 4-12. Fitxer OWL/RDF de l'article sobre Bono.

```

<owl:Ontology
  ...
  <owl:imports
    rdf:resource="http://semantic-mediawiki.org/swivt/1.0"/>
</owl:Ontology>

```

Sota l'element arrel, el primer element que es troba en un document OWL és `owl:Ontology`. La única informació significativa que aquest element aporta a l'arxiu és que s'importa una ontologia que es farà servir al llarg d'aquest. La ontologia s'anomena SWiVT⁴⁰, i és pròpia de SMW. D'aplicació interna, inclou termes bàsics emprats en el sistema d'anotacions, com `Subject`, `page` o `creationDate`.

```

<swivt:Subject

```

Especifica un subjecte del graf RDF, i del qual es defineixen algunes propietats. Dels dos subjectes que incorpora aquest document, el primer és el que parla sobre el article pròpiament. En quant a l'aspecte que realment ens interessa, que és la compartició d'informació amb l'exterior, podem veure la utilització dels termes de la Music Ontology `mo:member_of` i `mo:primary_instrument`, així el com `foaf:name` de la ontologia FOAF. Noteu també com amb `rdf:type` s'indica que es tracta d'un `MusicArtist`.

Aquest bloc del document pot ser representat d'una manera simplificada amb el següent graf:

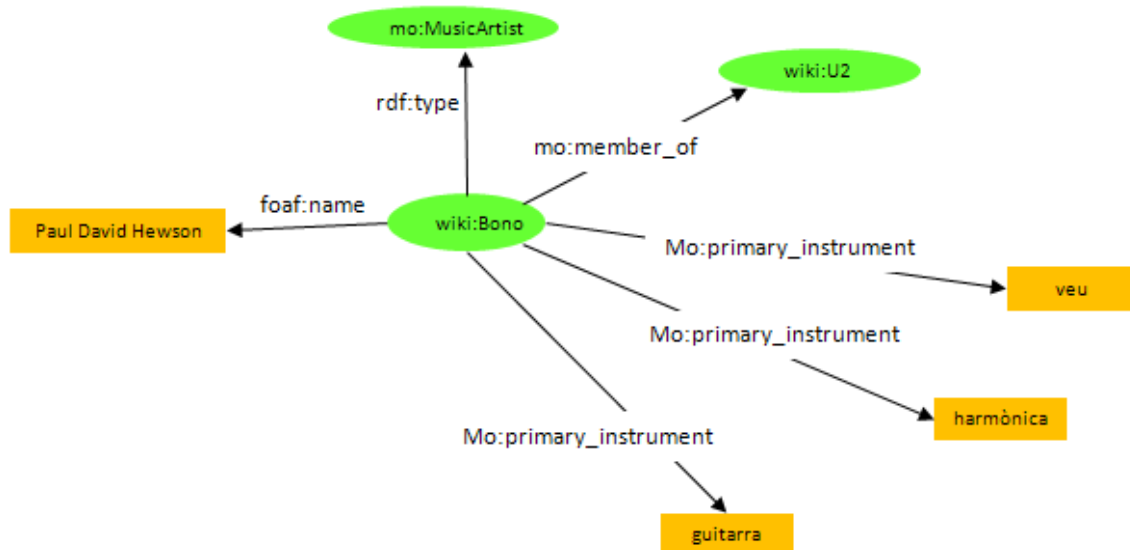


Figura 4-13. Graf simplificat sobre Bono.

L'altre subjecte del document descriu breument l'altre article de la wiki amb el que la pàgina de Bono està relacionada: l'article sobre U2. Per finalitzar, el document RDF també descriu la propietat `mo:member`, que és la propietat amb que U2 es relaciona amb Bono.

⁴⁰ <http://semantic-mediawiki.org/swivt/>

La informació RDF associada a una pàgina no només és accessible des de dins de la pròpia wiki, sinó que aquesta es pot enllaçar amb l'exterior per mitjà d'una URL de tipus:

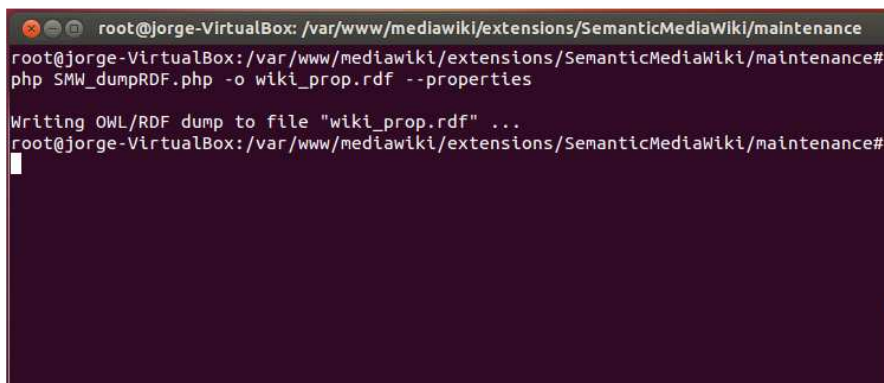
```
url/Special:ExportRDF/Pàgina
```

On `url` és la URL on està allotjada la wiki a la web. A la **Figura 4-12**, on es pot apreciar la URL per a les dades RDF associades a Bono a la barra de direccions. En aquest cas, però, la URL és un directori local en tractar-se d'un servidor instal·lat i funcionant en un PC domèstic. Si haguéssim publicat la nostra wiki en un servidor de Internet amb un nom com, per exemple, `http://www.irishrock.com/`, una eina remota podria accedir a al graf de Bono apuntant directament a:

```
http://www.irishrock.com/Special:ExportRDF/Bono
```

Adicionalment, SMW incorpora un script que permet exportar totes les dades semàntiques que conté la wiki. L'script s'anomena `SMW_dumpRDF.php` i es troba dins de la carpeta `maintenance` que s'ha creat en instal·lar SMW. Mitjançant les opcions es poden exportar totes les dades o només les categories, les propietats, tipus de dades, etc, a un arxiu destí.

Un exemple d'ús d'aquest script és el següent, en que des de la línia de comandes d'Ubuntu s'exporten les dades semàntiques referents a les propietats a un arxiu anomenat `wiki_prop.rdf`:



```
root@jorge-VirtualBox: /var/www/mediawiki/extensions/SemanticMediaWiki/maintenance
root@jorge-VirtualBox: /var/www/mediawiki/extensions/SemanticMediaWiki/maintenance#
php SMW_dumpRDF.php -o wiki_prop.rdf --properties

Writing OWL/RDF dump to file "wiki_prop.rdf" ...
root@jorge-VirtualBox: /var/www/mediawiki/extensions/SemanticMediaWiki/maintenance#
```

Figura 4-14. Exportació de les propietats de la wiki amb `SMW_dumpRDF.php`.

El resultat és un arxiu RDF amb tot un seguit d'elements `<owl:DatatypeProperty>` que defineixen aquestes propietats.

```

- <owl:DatatypeProperty rdf:about="http://purl.org/ontology/mo/track">
  <rdfs:label>Track</rdfs:label>
  <swivt:page rdf:resource="http://localhost/mediawiki/index.php/Property:Track"/>
  <rdfs:isDefinedBy rdf:resource="http://localhost/mediawiki/index.php/Special:ExportRDF/Property:Track"/>
  <swivt:wikiNamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">102</swivt:wikiNamespace>
  <swivt:specialProperty_IMPO rdf:datatype="http://www.w3.org/2001/XMLSchema#string">mo track
  http://purl.org/ontology/mo/</swivt:specialProperty_IMPO>
  <swivt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2012-11-24T09:47:22Z</swivt:wikiPageModificationDate>
  <property:Modification_date-23aux rdf:datatype="http://www.w3.org/2001/XMLSchema#double">2456255.9078935</property:Modification_date-23aux>
  <swivt:wikiPageSortKey rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Track</swivt:wikiPageSortKey>
  <swivt:type rdf:resource="http://semantic-mediawiki.org/swivt/1.0#_wpg"/>
</owl:DatatypeProperty>

```

Figura 4-15. Fragment de l'arxiu wiki_prop.rdf.

5. Conclusions

Al llarg d'aquest treball s'ha pogut contrastar com les tecnologies de la Web Semàntica estableixen les regles base per a la compartició del coneixement a la web. Les wikis són una eina que poden jugar un paper molt important en aquest sentit. Aquestes, amb la incorporació d'ontologies, fan possible que la gran quantitat de dades disgregades en diferents repositoris a Internet puguin ser "enteses", a més de per l'ésser humà, pels computadors.

Durant la creació de la wiki sobre música rock s'han pogut ajuntar aquestes peces i s'ha comprovat el potencial dels motors wikis semàntics en general i de Semantic MediaWiki en particular, un dels objectius principals d'aquest projecte. Aquests motors suposen una millora en quant a la gestió de la informació respecte als motors tradicionals, facilitant enormement la navegació i la cerca gràcies a la inclusió de marcat semàntic.

Semantic MediaWiki presenta una interfície i un mode de funcionament relativament senzill per a propòsits generals, i està dirigit a aquells usuaris que vulguin desenvolupar una wiki i compartir-la amb el món, ja sigui amb persones com amb processos automàtics. Treballar amb un motor semàntic ha facilitat la consolidació de conceptes que en començar el treball podien semblar certament difusos (ontologia, inferència,...) , i que també era un dels objectius del projecte.

El fet de no estar familiaritzat amb les wikis semàntiques ni amb la participació amb wikis tradicionals va dificultar la planificació del projecte, sobretot pel que fa al Capítol 4. Si bé la resta de capítols es van poder seguir segons el guió i la temporització previstes, la realització d'aquest es va veure alterada des d'un principi per les dificultats que es van presentar en la instal·lació de les eines requerides, i que va esdevenir en un endarreriment dels temps d'execució de les tasques. A més, en començar a treballar amb Semantic MediaWiki es va alterar l'enfocament inicial del capítol perquè es va començar a ser conscient de les seves possibilitats reals.

El treball acaba amb l'exportació d'articles de la wiki en format RDF. Ha mancat veure però com altres eines són capaços d'entendre aquestes dades i incorporar-les al seu procés, com per exemple, un cercador orientat a recopilar dades sobre música.

6. Glossari

- **Classe.** en Programació Orientada a Objectes, una classe descriu un grup d'entitats amb característiques comunes del món real, que es fa servir com un model per a crear objectes d'aquest tipus.
- **DTD.** Document Type Definition, és un llenguatge que es fa servir per definir l'estructura d'un document XML.
- **Folcsonomia.** Categorització col·laborativa per mitjà d'etiquetes.
- **Graf.** Estructura constituïda per objectes i les relacions que es poden establir entre ells.
- **Inferència.** Raonament mitjançant el qual hom passa d'una o més proposicions, acceptades com a veres o falses, a una altra proposició, la veritat o falsedat de la qual hom suposa que depèn de la veritat o falsedat de la primera o primeres proposicions.
- **Instància.** En Programació Orientada a Objectes, una instància descriu un membre concret del grup d'entitats al que pertany la classe a partir del qual s'ha creat.
- **Interfície.** Conjunt d'elements gràfics que permeten als usuaris interactuar de forma intuïtiva amb un sistema informàtic.
- **Metadades.** Dades que proveeixen informació sobre un o més aspectes d'unes altres dades, com la estructura, el propòsit, etc.
- **Namespace.** Espai de noms, prefix per un conjunt de noms generalment emprat per evitar ambigüitats.
- **Objecte.** Veure instància.
- **Ontologia.** Representació d'un àrea de coneixement, anomenat domini, per mitjà d'un conjunt de conceptes i les relacions entre ells.
- **OWL/XML.** Sintaxis basada en XML per expressar ontologies.
- **OWL.** Web Ontology Language, llenguatge per a expressar ontologies.
- **Programació orientada a objectes.**
- **RDF Schema.** Llenguatge que serveix per especificar vocabularis per a ser emprats en triplets RDF.

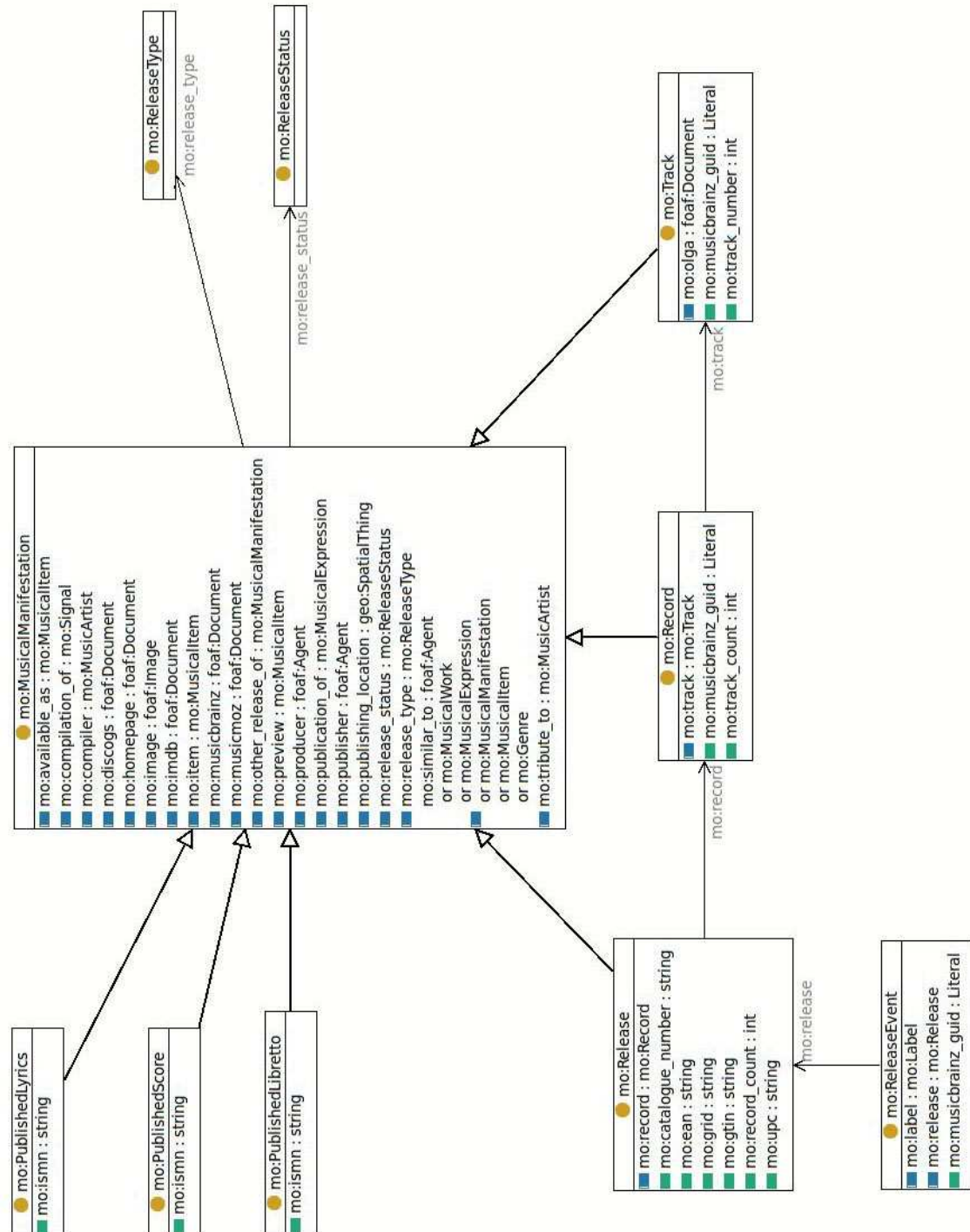
- **RDF.** Resource Frame Description, model de dades per representar informació sobre els recursos de la web i les seves propietats.
- **RDF/XML.** Sintaxis basada en XML per representar grafs RDF.
- **Script.** Fitxer de text que conté un programa.
- **Semàntica.** Significat de les paraules.
- **SPARQL.** SPARQL Protocol And RDF Query Language, llenguatge de consulta de dades RDF.
- **Subclasse.** Classe que deriva d'una altra classe i de la que hereta les seves propietats.
- **Superclasse.** Classe de la qual deriven una o més subclasses.
- **Triplet.** Element bàsic d'un graf RDF, format per un subjecte, un predicat i un objecte.
- **W3C.** World Wide Web Consortium, comunitat internacional formada per Tim Berners Lee amb el propòsit de desenvolupar estàndards per a la web.
- **WYSIWYG.** What You See Is What You Get, es un terme que es refereix a la tecnologia informàtica que permet que el que es veu durant l'edició o programació es correspongui més o menys acuradament amb el resultat final.
- **XML.** Extensible Markup Language, conjunt de regles per representar informació d'una manera estructurada.
- **XML Schema.** Llenguatge basat en XML que es fa servir per definir l'estructura d'un document XML.
- **XPath.** XML Path Language, llenguatge de consulta per a dades XML.
- **XQuery.** XML Query Language, llenguatge de consulta per a dades RDF.

7. Bibliografía

- **Antoniou, G.; Van Harmelen, F.** (2003). A Semantic Web Primer. The Mit Press.
- **Ducharne, B.** (2011). Learning SPARQL. O'Reilly.
- **Eckstein, R.; Casabianca, M.** (2001). XML Pocket Reference, 2nd Edition. O'Reilly.
- **Jiménez Mavillard, A.** (2011). Wikis y tecnologías semánticas, una propuesta para la Externalización del Conocimiento. Universidad de Sevilla.
- **Ray, E. T.** (2001). Learning XML. O'Reilly.
- **Walmsley, P.** (2007). XQuery. O'Reilly.
- **Yu, L.** (2011). A Developer's Guide to the Semantic Web. Springer.
- **Hoenderboom, B.; Peng, L.** (2009). A Survey of Semantic Wikis for Requirements Engineering. University of Groningen.
- **Dello, K.; Tolksdorf, R.; Simperl, E.** Creating and using Semantic Web information with Makna. Freie Universität Berlin.
- **Dello, K.; Nixon, L.; Tolksdorf, R.** Extending the Makna Semantic Wiki to support Workflows. Freie Universität Berlin.
- **Bry, F.** (2012). Semantic Wikis: Approaches, Applications, and Perspectives. A 8th Reasoning Web Summer School.
- **Schaffert, S.; Bry, F.; Baumeister, J. Kiesel, M.** Semantic Wikis.
- **Krötzsch, M.; Vrandečić, D.; Völkel, M.; Haller, H.; Studer, R.** (2007). Semantic Wikipedia. Institut AIFB, Universität Karlsruhe (TH).
- **Buffa, M.; Gandon, F.** SweetWiki: Semantic Web Enabled Technologies in Wiki.
- **Kuhn, T.** AceWiki: A Natural and Expressive Semantic Wiki. University of Zurich.
- **Schaffert, S.** IkeWiki: A SemanticWiki for Collaborative Knowledge Management. Salzburg Research Forschungsgesellschaft.

- **Auer, S.; Dietzold, S.; Riechert, T.** OntoWiki – A Tool for Social, Semantic Collaboration. University of Pennsylvania.
- **MediaWiki.** <http://www.mediawiki.org/wiki/MediaWiki>, 2012.
- **Music Ontology class schemas.**
<https://github.com/motools/musicontology/wiki/Class-Schemas>, 2012.
- **Music Ontology Specification.** <http://musicontology.com/>, 2012.
- **OWL 2 Document Overview.** <http://www.w3.org/TR/owl2-overview/>, 2012.
- **OWL 2 Primer.** <http://www.w3.org/TR/owl2-primer/>, 2012.
- **RDF Primer.** <http://www.w3.org/TR/rdf-primer/>, 2012.
- **Semantic MediaWiki.** <https://www.semantic-mediawiki.org/>, 2012.
- **Semantic Web.** <http://semanticweb.org>, 2012.
- **SPARQL.** <http://www.w3.org/TR/rdf-sparql-query/>, 2012.
- **XML Schema Primer.** <http://www.w3.org/TR/xmlschema-0/>, 2012.
- **XPath.** <http://www.w3.org/TR/xpath/>, 2012.
- **XQuery.** <http://www.w3.org/TR/xquery/>, 2012.
- **W3Schools.** <http://www.w3schools.com/>, 2012.

8.1. Diagrama de classes de la MusicOntology (MusicalManifestation)



8.2. Diagrama de classes de la MusicOntology (MusicArtist, MusicGroup, SoloMusicArtist)

