Web semàntica aplicada a la gestió i caracterització de dades, sensors i accés a serveis per a vaixells d'investigació oceanogràfica

Olivé Lopez, Joan
2n cicle d'Enginyeria en Informàtica

Geva Urbano, Felipe
Consultor

11/06/2013

Aquest treball està dedicat a la meva parella Montse i a la meva mare MªDolors, dos persones molt especials, molt importants per a mi i que sempre m'han recolzat en tot i més. Gràcies!

Aquesta obra està subjecta a una llicència de Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya de Creative Commons**Implementació del model conceptual amb Protégé**

FITXA DEL TREBALL FINAL

Títol del treball:	Web semàntica aplicada a la gestió i caracterització de dades, sensors i accés a serveis per a vaixells d'investigació oceanogràfica
Nom de l'autor:	Olivé Lopez, Joan
Nom del consultor:	Geva Urbano, Felipe
Data de lliurament:	06/2013
Àrea del Treball Final:	Web Semàntica
Titulació:	2n cicle d'Eng. Informàtica

Resum del Treball:

Amb el present projecte es proposa fer un estudi de la web semàntica aplicada a la descripció i modelització de la instrumentació marina operada en el context de campanyes d'investigació oceanogràfica.

Es considera que l'aproximació que proporciona la web semàntica pot ser de gran utilitat a l'hora de descriure un entorn complex i multidisciplinari com és el dels vaixells d'investigació, en els que operen de forma coordinada gran varietat de sensors, i en els que gran part d'aquesta informació es troba dispersa i poc estructurada. Per la qual cosa, el fet de poder comptar amb una eina, com és el cas de la web semàntica, que permeti accedir i compartir tant la informació com les diferents relacions que s'hi estableixen a nivell intern i extern pot resultar de gran utilitat.

En el context de la investigació oceanogràfica realitzada per vaixells intervenen una gran quantitat de sensors diferents que adquireixen dades molt variades. Aquest escenari dóna lloc a que la informació disponible es trobi dispersa de manera que moltes vegades resulta difícil accedir a la informació de cada sensor i respondre preguntes del tipus:

Quin tipus de dades proporciona?
Quins són els requeriments de treball?
Pot ser substituït per un altre sensor?
Necessita altres sensors associats per operar?
En quins escenaris opera?

Per resoldre aquestes qüestions i estructurar tota aquesta informació pot resultar molt apropiat l'ús d'una ontologia i per poder fer que sigui accessible a totes les persones i departaments als que pot resultar útil es proposa emprar una wiki semàntica.

Abstract:

This project aims to study the Semantic Web applied to the description and modelling of marine instrumentation operated in the context of oceanographic research campaigns.

It is considered that the Semantic Web approach can be very useful to describe a complex and multidisciplinary environment as in the case of the research vessels where there are operating a wide range of sensors. A lot of the information collected for these sensors are scattered and unstructured. Therefore, it can be very useful to have a tool, such as the semantic web, that allows to access and share both information and the different relationships established at internal and external level.

In the context of oceanographic research vessels are involved a lot of different sensors that acquire very different data. This scenario makes it difficult to access to information from each sensor and answer questions such as:

What type of data provides?

What are the job requirements?

Can be replaced by another sensor?

Does one sensor need other associated sensors to operate?

In which operating scenarios does it operate?

To solve these issues and to structure all this information it can be very helpful to use an ontology. And to make this ontology accessible to all people and departments that are interested in this information a Semantic Wiki is proposed.

Paraules clau:

Web semàntica, wiki, ITO, ontologia, oceanografia, sensor, SWE.

Índex

1. Introducció....1
1.1 Context i justificació del Treball....1
1.2 Objectius del Treball....2
1.3 Enfocament i mètode seguit....3
1.4 Planificació del projecte....4
Diagrames de Gantt....4
1.5 Breu sumari de productes obtinguts....9
1.6 Breu descripció dels altres capítols de la memòria....9
2. Conceptes bàsics....10
2.1 Web semàntica....10
2.2 XML....13
2.3 RDF i RDF-SCHEMA....15
2.4 OWL....17
2.5 Wikis Semàntiques....18
2.6 Projectes relacionats....19
SWE....19
Linked Sensor Data....22
3. Disseny de l'ontologia per vaixells d'investigació oceanogràfica....24
3.1 Anàlisi d'ontologies oceanogràfiques....25
The MMI Ontology Server....25
SeaDataNET....26
EMODNET....28
3.2 Eines de desenvolupament....29
Protégé....29
NeOn Toolkit....30
3.3 Requeriments de l'ontologia....31
Domini de l'ontologia: 31
Propòsit de l'ontologia:....31
Usuaris de l'ontologia:....34
3.4 Elaboració del model conceptual....35
Definició de termes de l'ontologia....35
Definició de les classes i propietats de les classes....37
Relacions entre objectes....38
Propietats de les dades....39
Implementació del model conceptual amb Protégé....40
Individuals....41
4. Wiki semàntica per a vaixells d'investigació oceanogràfica....49
4.2 OntoWiki....50
4.3 Semantic Media Wiki....51
4.4 Semantic Media Wiki Plus....52
4.5 Kiwi....54
4.6 Diferències i selecció de programari....54
4.7 Importació de l'ontologia....55
4.8 L'ontologia a SMW+....56
5. Conclusions i línies de futur....61
5.1 Línies de futur....62
6. Glossari....63
7. Bibliografia....64

Llista de figures

Figura 1: Diagrama de Gantt TFG....4
Figura 2: Diagrama de Gantt PAC 1....5
Figura 3: Diagrama de Gantt PEC 2....5
Figura 4: Diagrama de Gantt PEC 3....6
Figura 5: Diagrama de Gantt Entrega Final....7
Figura 6: Tasques....8
Figura 7: La web semàntica segons la va concebre Berners-Lee (Figura de Berns-Lee)....11
Figura 8: Evolució de la web [1]....12
Figura 9: XML....13
Figura 10: Capçalera XML....14
Figura 11: SensorML XSD....14
Figura 12: "Tripleta" RDF....15
Figura 13: Exemple de "Tripleta" del sistema de navegació....16
Figura 14: Esquema SensorML....19
Figura 15: Fragment de fitxer O&M....20
Figura 16: Estructura de fitxer O&M....20
Figura 17: Linked Sensor Data Workflow [17]....23
Figura 18: MMI Ontology Registry and Repository....25
Figura 19: Vocabulari SeaDataNet....26
Figura 20: Termes de SeaDataNet com a URNs dins de SensorML....27
Figura 21: Portal EMODNET....28
Figura 22: Protégé....29
Figura 23: NeOn Toolkit....30
Figura 24: Serveis OGC WMS, WFS i WCS al vaixell....32
Figura 25: Client Web SOS al vaixell....32
Figura 26: OpeNDAP a través del servidor Thredds al vaixell....33
Figura 27: Portal SeaDataNet de consulta de CDIs....33
Figura 28: Múltiples disciplines implicades (Font:neptunecanada.com)....35
Figura 29: Termosalinògraf....36
Figura 30: Diagrama de Classes....37
Figura 31: Classes i ús....37
Figura 32: Relacions i ús....38
Figura 33: Data Properties....39
Figura 34: Model conceptual de l'ontologia....40
Figura 35: Individuals....41
Figura 36: Instàncies de departament....41
Figura 37: CTD....42
Figura 38: Information_service....43
Figura 39: Measurements....44
Figura 40: Observable_property....45
Figura 41: Procedures....46
Figura 42: Weather Station....47
Figura 43: Instàncies de Thematic Discipline....48
Figura 44: Realitat–Ontologia–Wiki Semàntica....49
Figura 45: Semantic Media Wiki + Fitxa....52

Figura 46: Instal·lador SMW+.....53
Figura 47: Semantic Media Wiki +.....54
Figura 48: Importació Ontologia a SMW+.....55
Figura 49: Procés de importació de l'ontologia des de SMW+.....55
Figura 50: Ontologia a SMW+.....56
Figura 51: "Category Tree" a SMW+.....56
Figura 52: Funcionament SMW+.....57
Figura 53: Procedure a SMW+.....58
Figura 54: Query Interface.....59
Figura 55: Pàgina amb Consulta.....60

1. Introducció

1.1 Context i justificació del Treball

Amb el present projecte es proposa fer un estudi de la web semàntica aplicada a la descripció i modelització de la instrumentació marina operada en el context de campanyes d'investigació oceanogràfica.

Es considera que l'aproximació que proporciona la web semàntica pot ser de gran utilitat a l'hora de descriure un entorn complex i multidisciplinari com és el dels vaixells d'investigació, en els que operen de forma coordinada gran varietat de sensors, i en els que gran part d'aquesta informació es troba dispersa i poc estructurada. Per la qual cosa, el fet de poder comptar amb una eina, com és el cas de la web semàntica, que permeti accedir i compartir tant la informació com les diferents relacions que s'hi estableixen a nivell intern i extern pot resultar de gran utilitat.

El marc de treball correspon al dels vaixells d'investigació oceanogràfica. Concretament es descriu l'entorn, els vaixells i els sensors gestionats per la "Unidad de Tecnologia Marina", UTM, que pertany al "Consejo Superior de Investigaciones Científicas", CSIC.

La UTM gestiona 3 vaixells d'investigació: El Buque Oceanográfico Sarmiento de Gamboa (BO Sarmiento de Gamboa), Buque de Investigación Oceanográfica Hespérides (BIO Hespérides) i el Buque Oceanográfico García del Cid (BO García del Cid). Els dos primers tenen com a marc de treball tot el món i el García del Cid opera principalment al Mediterrani i en projectes de menys durada. Tant el BO Sarmiento de Gamboa com el BIO Hespérides compten amb un sistema de comunicació satèl·lit permanent que els permet l'enviament i recepció de dades en temps real.

Tots tres vaixells poden desenvolupar tasques molt variades en funció de la campanya que calgui realitzar. L'equipament disponible i els sensors desplegats en cada cas varien i s'adapten a les diferents necessitats de cada projecte en funció per exemple de si es tracta d'una campanya biològica, geofísica o geològica.

En el context de la investigació oceanogràfica realitzada per vaixells intervenen una gran quantitat de sensors diferents que adquireixen dades molt variades. Aquest escenari dóna lloc a que la informació disponible es trobi dispersa de manera que moltes vegades resulta difícil accedir a la informació de cada sensor i respondre preguntes del tipus:

Quin tipus de dades proporciona?
Quin són els requeriments de treball?
Pot ser substituït per un altre sensor?
Necessita altres sensors associats per operar?
En quins escenaris opera?

Per tal de resoldre aquestes qüestions i estructurar tota aquesta informació pot resultar molt apropiat l'ús d'una ontologia. Per poder fer que sigui accessible a totes les persones i departaments als que pot resultar útil es proposa emprar una wiki semàntica.

Per altra banda, també cal tenir en compte que les observacions que es porten a terme dins del marc de la investigació oceanogràfica són tant variades que fa que els sets de dades obtinguts impliquin diferents estàndards. Per tant, per tal de poder mantenir una coherència conceptual i entre termes resulta de gran utilitat emprar tecnologies i eines que facin possible la interoperabilitat semàntica.

1.2 Objectius del Treball

Amb el present treball es pretén aprofundir en el coneixement del que és la web semàntica, els seus principals conceptes i quines són les seves possibilitats actuals. Es proposa l'aplicació de la web semàntica a la problemàtica de la investigació oceanogràfica per tal d'avaluar l'estat de l'art de la web semàntica.

Es defineix una ontologia aplicada a l'àmbit marí d'investigació oceanogràfica que ha de descriure els diferents actors que hi participen, els sensors encarregats de l'adquisició de les dades, les pròpies dades i els serveis que s'ofereixen sobre aquestes dades.

De forma resumida els objectius són els següents:

1. Estudi dels conceptes bàsics de la web semàntica.
2. Estudi en profunditat de l'estat de la web semàntica aplicada a l'àmbit marí d'investigació oceanogràfica.
3. Disseny i desenvolupament d'una ontologia aplicada a la modelització de dades, sensors i serveis associats al marc de treball donat per vaixells d'investigació oceanogràfica.
4. Creació d'una wiki semàntica que permeti la consulta i explotació de l'ontologia creada anteriorment.

1.3 Enfocament i mètode seguit

La metodologia emprada consistirà en la realització de dos prototips:

1. Prototip de l'ontologia.
2. Prototip de la wiki semàntica.

En una primera part es defineix una fase de recerca d'informació sobre l'estat actual de la web semàntica. Per tal de portar a terme aquesta recopilació d'informació es realitzarà cerca d'informació a Internet, assistència a congressos o actes relacionats i en general es procedirà a recopilar tota aquella informació relativa a l'àmbit del projecte. Una vegada avaluat l'estat de la web semàntica i la idoneïtat de les diferents implementacions aplicades a la investigació oceanogràfica es procedirà al disseny i implementació d'una ontologia que descrigui el marc de treball triat. Per tal de realitzar el disseny d'aquesta ontologia es farà servir el programa Protégé [10]. Per últim és procedirà a muntar una wiki semàntica a partir de l'ontologia generada. De cara a realitzar aquesta tasca s'avaluaran les diferents opcions disponibles i els seus avantatges i desavantatges de manera que finalment es seleccioni una de les implementacions per realitzar l'exportació de l'ontologia.

1.4 Planificació del projecte

Diagrames de Gantt

WBS	Nom	Comença	Acaba	Feina
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**
1.1	Lectura documentació	27 feb	28 feb	2d
1.2	**Definició Projecte**	**1 mar**	**7 mar**	**5d**
1.2.1	Titol Provisional i Tema	1 mar	4 mar	2d
1.2.2	Establiment Objectius	5 mar	5 mar	1d
1.2.3	Planificació	6 mar	7 mar	2d
1.3	Revisió PAC1	8 mar	8 mar	1d
1.4	Entrega PAC1	11 mar	11 mar	1d
2	**PEC2**	**12 mar**	**15 abr**	**25d**
2.1	Recull Bibliografic	12 mar	13 mar	2d
2.2	Conceptes Bàsics Web Semàntica	14 mar	18 mar	3d
2.3	Comparació Wikis Semàntques	19 mar	21 mar	3d
2.4	**Protege**	**22 mar**	**26 mar**	**3d**
2.4.1	Instal·lació	22 mar	22 mar	1d
2.4.2	Aprenentatge entorn	25 mar	26 mar	2d
2.5	Disseny Ontologia	27 mar	11 abr	12d
2.6	Revisió PEC2	12 abr	12 abr	1d
2.7	Entrega PEC2	15 abr	15 abr	1d
3	**PEC3**	**16 abr**	**21 mai**	**26d**
3.1	Estat de l'art de Web Semàntica	16 abr	18 abr	3d
3.2	Proves Wikis Semàntques	19 abr	25 abr	5d
3.3	Importació Ontologia a Wiki Semàntica	26 abr	29 abr	2d
3.4	Proves Wiki Semàntica (Validació)	30 abr	7 mai	6d
3.5	Validació Wiki Semàntica	8 mai	10 mai	3d
3.6	Avaluació Resultats	13 mai	17 mai	5d
3.7	Revisió entrega PAC3	20 mai	20 mai	1d
3.8	Entrega PAC3	21 mai	21 mai	1d
4	**Entrega Final**	**22 mai**	**11 jun**	**15d**
4.1	Revisió memòria	22 mai	31 mai	8d
4.2	Revisió Wiki	3 jun	3 jun	1d
4.3	Elaboració presentació	4 jun	7 jun	4d
4.4	Revisió entrega final	10 jun	10 jun	1d
4.5	Entrega Final	11 jun	11 jun	1d
5	**Debat Virtual**	**25 jun**	**28 jun**	**4d**
5.1	Participació Debat	25 jun	28 jun	4d

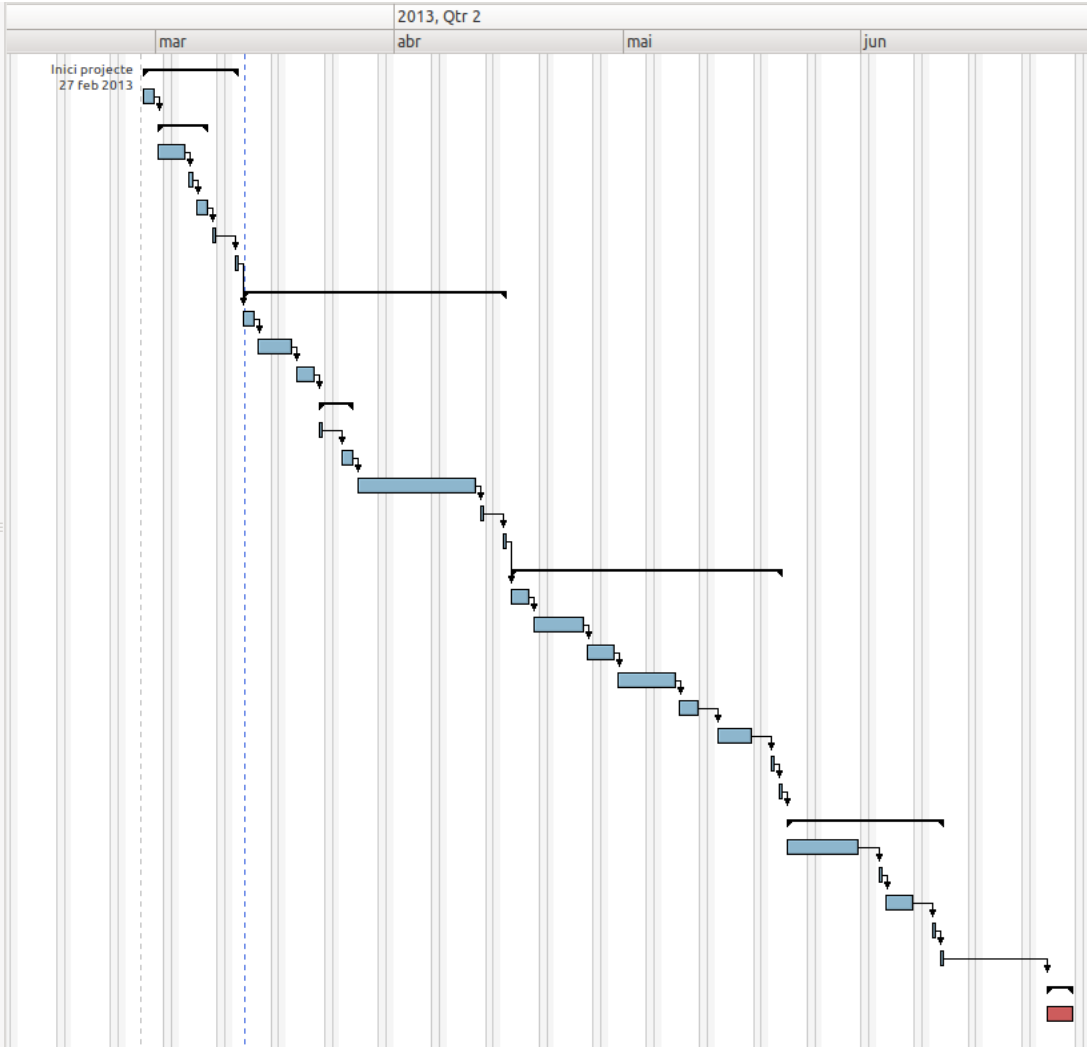


Figura 1: Diagrama de Gantt TFG

Planificació PAC 1

WBS	Nom	Comença	Acaba	Feina
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**
1.1	Lectura documentació	27 feb	28 feb	2d
1.2	**Definició Projecte**	**1 mar**	**7 mar**	**5d**
1.2.1	Titol Provisional i Tema	1 mar	4 mar	2d
1.2.2	Establiment Objectius	5 mar	5 mar	1d
1.2.3	Planificació	6 mar	7 mar	2d
1.3	Revisió PAC1	8 mar	8 mar	1d
1.4	Entrega PAC1	11 mar	11 mar	1d

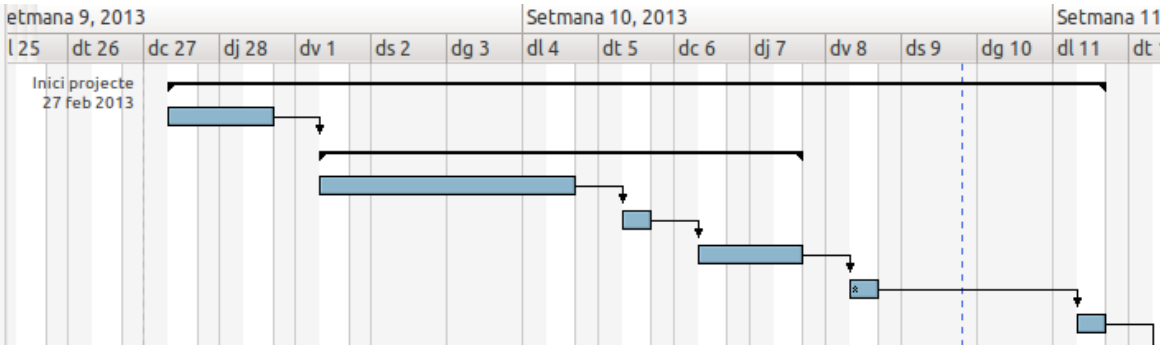


Figura 2: Diagrama de Gantt PAC 1

Planificació PEC 2

WBS	Nom	Comença	Acaba	Feina
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**
2	**PEC2**	**12 mar**	**15 abr**	**25d**
2.1	Recull Bibliografic	12 mar	13 mar	2d
2.2	Conceptes Bàsics Web Semàntica	14 mar	18 mar	3d
2.3	Comparació Wikis Semàntques	19 mar	21 mar	3d
2.4	**Protege**	**22 mar**	**26 mar**	**3d**
2.4.1	Instal·lació	22 mar	22 mar	1d
2.4.2	Aprenentatge entorn	25 mar	26 mar	2d
2.5	Disseny Ontologia	27 mar	11 abr	12d
2.6	Revisió PEC2	12 abr	12 abr	1d
2.7	Entrega PEC2	15 abr	15 abr	1d

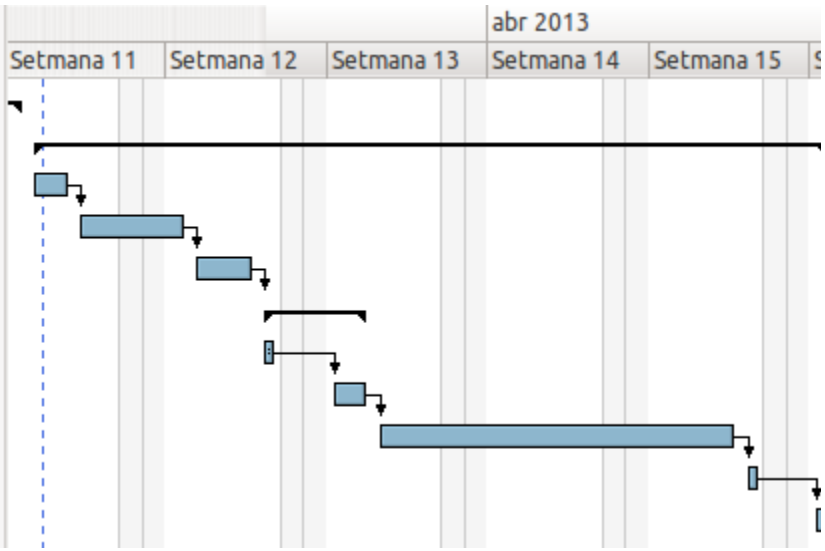


Figura 3: Diagrama de Gantt PEC 2

Planificació PEC 3

WBS	Nom	Comença	Acaba	Feina
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**
2	**PEC2**	**12 mar**	**15 abr**	**25d**
3	**PEC3**	**16 abr**	**21 mai**	**26d**
3.1	Estat de l'art de Web Semàntica	16 abr	18 abr	3d
3.2	Comparació Wikis Semàntques	19 abr	25 abr	5d
3.3	Importació Ontologia a Wiki Semàntica	26 abr	29 abr	2d
3.4	Proves Wiki Semàntica (Validació)	30 abr	7 mai	6d
3.5	Validació Wiki Semàntica	8 mai	10 mai	3d
3.6	Avaluació Resultats	13 mai	17 mai	5d
3.7	Revisió entrega PAC3	20 mai	20 mai	1d
3.8	Entrega PAC3	21 mai	21 mai	1d

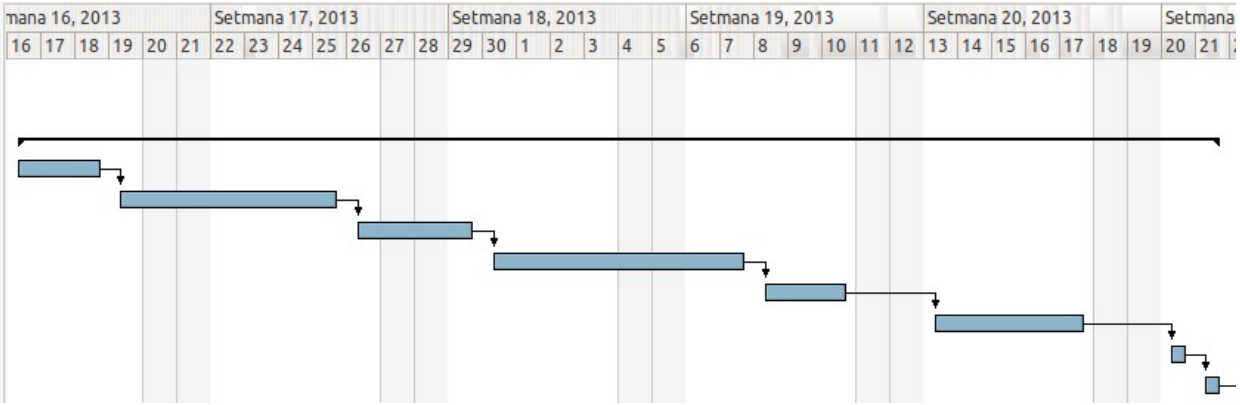


Figura 4: Diagrama de Gantt PEC 3

Planificació entrega final

WBS	Nom	Comença	Acaba	Feina
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**
2	**PEC2**	**12 mar**	**15 abr**	**25d**
3	**PEC3**	**16 abr**	**21 mai**	**26d**
4	**Entrega Final**	**22 mai**	**11 jun**	**15d**
4.1	Revisió memòria	22 mai	31 mai	8d
4.2	Revisió Wiki	3 jun	3 jun	1d
4.3	Elaboració presentació	4 jun	7 jun	4d
4.4	Revisió entrega final	10 jun	10 jun	1d
4.5	Entrega Final	11 jun	11 jun	1d

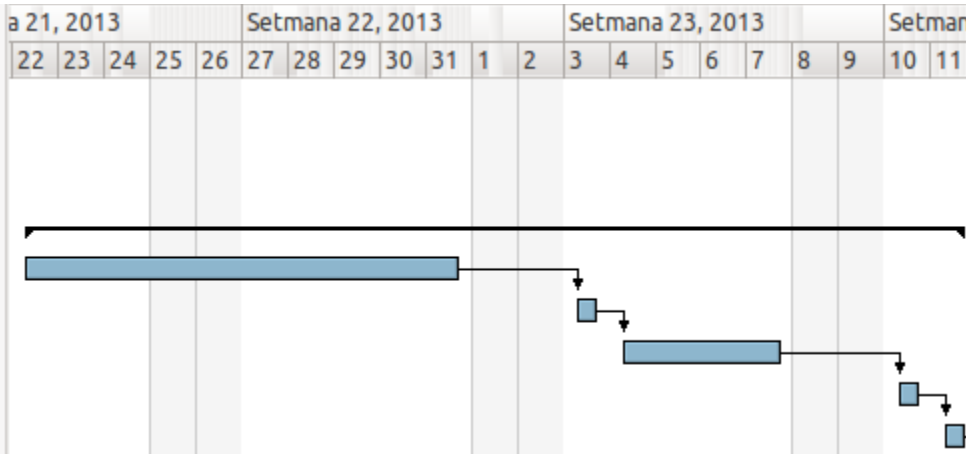


Figura 5: Diagrama de Gantt Entrega Final

WBS	Nom	Comença	Acaba	Feina	Durada
1	**PAC1-Pla de Treball**	**27 feb**	**11 mar**	**9d**	**9d**
1.1	Lectura documentació	27 feb	28 feb	2d	2d
1.2	**Definició Projecte**	**1 mar**	**7 mar**	**5d**	**5d**
1.2.1	Titol Provisional i Tema	1 mar	4 mar	2d	2d
1.2.2	Establiment Objectius	5 mar	5 mar	1d	1d
1.2.3	Planificació	6 mar	7 mar	2d	2d
1.3	Revisió PAC1	8 mar	8 mar	1d	1d
1.4	Entrega PAC1	11 mar	11 mar	1d	1d
2	**PEC2**	**12 mar**	**15 abr**	**25d**	**25d**
2.1	Recull Bibliografic	12 mar	13 mar	2d	2d
2.2	Conceptes Bàsics Web Semàntica	14 mar	18 mar	3d	3d
2.3	Comparació Wikis Semàntques	19 mar	21 mar	3d	3d
2.4	**Protege**	**22 mar**	**26 mar**	**3d**	**3d**
2.4.1	Instal·lació	22 mar	22 mar	1d	1d
2.4.2	Aprenentatge entorn	25 mar	26 mar	2d	2d
2.5	Disseny Ontologia	27 mar	11 abr	12d	12d
2.6	Revisió PEC2	12 abr	12 abr	1d	1d
2.7	Entrega PEC2	15 abr	15 abr	1d	1d
3	**PEC3**	**16 abr**	**21 mai**	**26d**	**26d**
3.1	Estat de l'art de Web Semàntica	16 abr	18 abr	3d	3d
3.2	Proves Wikis Semàntques	19 abr	25 abr	5d	5d
3.3	Importació Ontologia a Wiki Semàn	26 abr	29 abr	2d	2d
3.4	Proves Wiki Semàntica (Validació)	30 abr	7 mai	6d	6d
3.5	Validació Wiki Semàntica	8 mai	10 mai	3d	3d
3.6	Avaluació Resultats	13 mai	17 mai	5d	5d
3.7	Revisió entrega PAC3	20 mai	20 mai	1d	1d
3.8	Entrega PAC3	21 mai	21 mai	1d	1d
4	**Entrega Final**	**22 mai**	**11 jun**	**15d**	**15d**
4.1	Revisió memòria	22 mai	31 mai	8d	8d
4.2	Revisió Wiki	3 jun	3 jun	1d	1d
4.3	Elaboració presentació	4 jun	7 jun	4d	4d
4.4	Revisió entrega final	10 jun	10 jun	1d	1d
4.5	Entrega Final	11 jun	11 jun	1d	1d
5	**Debat Virtual**	**25 jun**	**28 jun**	**4d**	**4d**
5.1	Participació Debat	25 jun	28 jun	4d	4d

Figura 6: Tasques

1.5 Breu sumari de productes obtinguts

Com a resultat del present treball de final de grau s'han obtingut els següents productes:

- Una ontologia en format OWL que modela la problemàtica del domini proposat: sensors i serveis.
- Integració de l'ontologia anterior en una wiki semàntica de manera que aquesta sigui fàcilment consultable.

1.6 Breu descripció dels altres capítols de la memòria

En el primer capítol es fa una introducció al treball, quins són els seus objectius i es presenta quina és la planificació del mateix. A continuació, al segon capítol es presenten els conceptes bàsics relacionats. En el tercer capítol es procedeix amb el disseny de l'ontologia proposada i finalment es presenta la integració de l'ontologia a una wiki semàntica.

2. Conceptes bàsics

En primer terme es passa a introduir tota una sèrie de conceptes rellevants en el context del present treball: “Web semàntica aplicada a la gestió i caracterització de dades, sensors i accés a serveis per a vaixells d'investigació oceanogràfica”. Aquests conceptes constitueixen la base teòrica sobre la que es realitza la resta del treball. Es considera rellevant realitzar una introducció històrica al que és la web semàntica, en que consisteix i que és el que aporta, a continuació es passa a fer una breu introducció a tota una sèrie de conceptes i tecnologies rellevants en aquest context i que fa possible la seva implementació. Així es parla de XML, RDF, OWL, ontologia i wiki semàntica.

2.1 Web semàntica

Definició:

> “The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C. It is based on the Resource Description Framework (RDF).”
>
> W3C, http://www.w3.org/2001/sw/

Des dels orígens de la “world wide web”, www, als finals dels anys 80 per part de Tim Berners-Lee i Robert Cailliau mentre treballaven al CERN, la web s'ha estès de tal manera que ha revolucionat la manera en que entenem la societat avui en dia. Posteriorment el mateix Tim Berners-Lee és el primer en fer servir la expressió web semàntica.

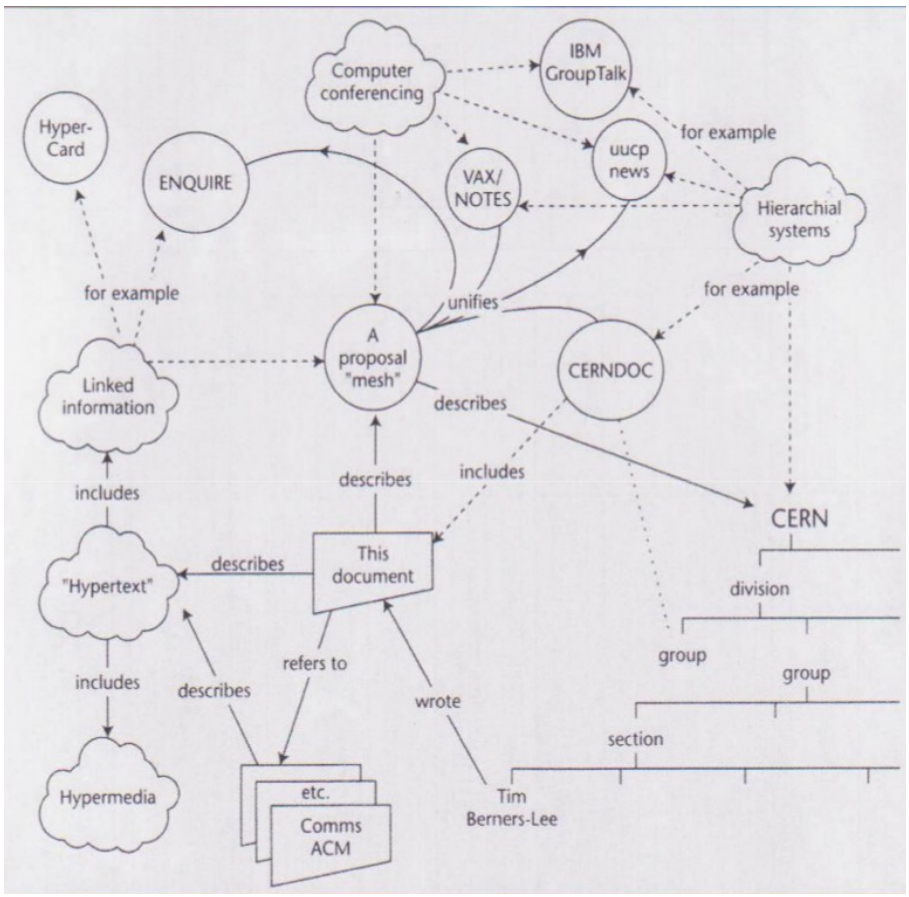


Figura 7: La web semàntica segons la va concebre Berners-Lee (Figura de Berns-Lee)

La quantitat d'informació disponible ha crescut de forma exponencial i ja fa temps que ha deixat d'estar confinada al món dels ordinadors i és accessible des de tot tipus de dispositius, ja siguin telèfons intel·ligents, "smartphones", o la televisió, per posar només dos exemples. Podem dir que el mateix èxit de la web ha estat un dels seus problemes degut a la sobrecàrrega d'informació i a la seva heterogeneïtat.

Encara ara la majoria dels continguts de la web estan construïts emprant HTML, l'objectiu del qual és la representació de la informació en un format més o menys amigable de cara als usuaris. L'HTML per contra no resulta una eina útil per estructurar la informació ni les relacions que hi pot haver entre les dades.

Per tal de donar solucions a les limitacions inherents al HTML anteriorment enunciades es fa evident que cal buscar nous llenguatges dissenyats des del primer moment per tal de treballar amb dades, tal i com és el cas de Resource Description Framework (RDF), Web Ontology Language (OWL) i de Extensible Markup Language (XML).

RDF, OWL i XML són algunes de les tecnologies emprades per la web semàntica.

L'objectiu que es pretén assolir amb la web semàntica és fer evolucionar la web actual formada per documents semiestructurats cap a un nou paradigma web que es pot anomenar com “Web of data” [12]. Segons el World Wide Web Consortium (W3C) [13]:

"The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries." [14].

Cal que les dades siguin accessibles fent servir l'arquitectura general de la web, tal i com és el cas de URIs. Fonamentalment del que es tracta és de que a l’introduir semàntica a la web, sigui el propi programari el que processi el contingut de la web.

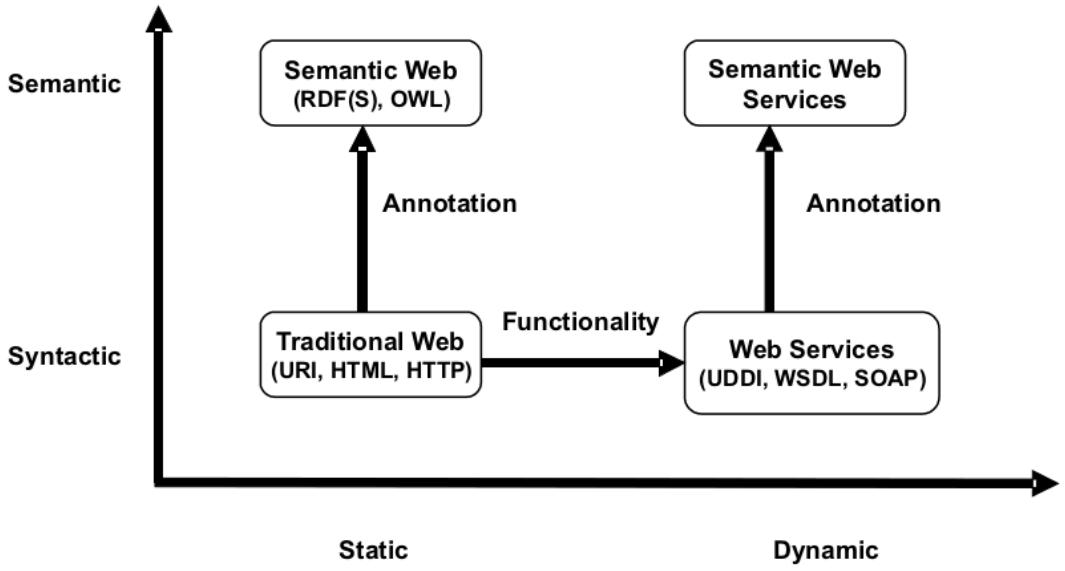


Figura 8: Evolució de la web [1]

A la figura anterior es pot veure com la web tradicional evoluciona d’una web estàtica a una web dinàmica a l’estendre les seves funcionalitats (evolució en el sentit de les X). Per altra banda evoluciona donant contingut semàntic mitjançant la utilització d'anotacions (evolució en el sentit de les Y), donant lloc a les tecnologies RDF(S) i OWL. Per últim, si s'ajunten les dues evolucions s'arribaria als serveis web semàntics, que conjuga les dues evolucions.

És important tenir en compte que en el seu origen la web està pensada per a l'intercanvi d'informació entre humans. Per contra, la web semàntica està pensada com una web de dades que pot ser processada directa i indirectament per màquines. El que es pretén amb la web semàntica, entre d'altres qüestions, és fer possible que les màquines puguin entendre i respondre preguntes formulades pels usuaris basant-se en el significat. Per tal de fer possible aquesta comprensió cal que la informació estigui semànticament estructurada.

2.2 XML

Per tal del poder assolir els objectius de la web semàntica cal un llenguatge tecnològicament neutre que permeti l'intercanvi d'informació entre sistemes i d'aquesta manera faciliti la interoperabilitat. El millor candidat per aquesta tasca és XML (eXtensible Markup Languatge) [15]. XML deriva del llenguatge "Standard Generalized Markup Language", SGML [27].

Així com HTML és un llenguatge pensat per presentar la informació (capa de presentació), XML permet representar l'estructura de les dades (capa de dades). Entre els avantatges de XML destaca que sigui fàcilment extensible, simplement afegint noves etiquetes.

XML fa possible que els documents que hi ha a la web tinguin una estructura però aquesta no expressa res de la semàntica.

```
- <om:featureOfInterest>
  - <gml:DynamicFeature gml:id="Vessel_Site">
    - <gml:description>
        The vessel site seen as the shape of the observation (a 3D box ) in a specific time. Is the Volume of air
        been chose to describe the FOI
      </gml:description>
    - <gml:location>
      - <gml:Point gml:id="Vessel_Site_Position">
          <gml:pos srsName="http://www.opengis.net/def/crs/EPSG/0/4326">54.9 10.52</gml:pos>
        </gml:Point>
      </gml:location>
    - <gml:validTime>
      - <gml:TimeInstant gml:id="Vessel_Site_Time">
          <gml:timePosition>2012-06-29T17:22:25</gml:timePosition>
        </gml:TimeInstant>
      </gml:validTime>
    </gml:DynamicFeature>
  </om:featureOfInterest>
```

Figura 9: XML

Un document XML no deixa de ser un arxiu de text. Cal que estigui codificat emprant Unicode (del qual el joc de caràcters ASCII és un subconjunt).

Els documents XML han de mantenir una certa estructura i complir una sèrie de normes per tal de ser un document XML vàlid. Així tenim que ha d'estar format únicament per marques o etiquetes i de dades de caràcters. Les etiquetes es troben dins dels delimitadors "<" i ">".

<etiqueta>

XML, a l'igual que molts llenguatges de programació és sensible a majúscules ("Case sensitive"), per la qual cosa diferencia majúscules i minúscules. Tot document XML consta de dues parts: una capçalera i el cos del document. A la capçalera es poden trobar metadades que descriuen el propi document, tal i com és el cas de la versió de XML que fa servir i el tipus de document. Al cos del document s'hi troba la informació que s'emmagatzema seguint una estructura jeràrquica.

```
<?xml version="1.0" encoding="UTF-8"?>
<sml:SensorML xmlns:sml="http://www.opengis.net/sensorML/1.0.1"
    xmlns:gml="http://www.opengis.net/gml"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:swe="http://www.opengis.net/swe/1.0.1"
    xmlns:xlink="http://www.w3.org/1999/xlink"
    xsi:schemaLocation="http://www.opengis.net/sensorML/1.0.1
       http://schemas.opengis.net/sensorML/1.0.1/sensorML.xsd"
    version="1.0.1">
```

Figura 10: Capçalera XML

A la figura anterior es pot veure un exemple de capçalera de document XML. Apareix la següent informació:

- A la primera línia s'indica que la **versió** de XML és la 1.0 i a continuació la declaració de la codificació del document, que és la UTF-8.
- A la següent línia s'observa com es tracta d'un document que descriu un sensor emprant el llenguatge de modelat de sensors (**SensorML**) [16].
- Es passa a enumerar tota una sèrie d'espais de noms amb l'identificador **xmlns**, que es faran servir al document.

A l'hora de treballar amb documents XML complexos resulta de gran utilitat fer servir esquemes XML, XSD. XSD, a l'igual que la resta de llenguatges de esquema per XML, permet expressar tot un conjunt de regles que ha de complir un document XML de manera que es pugui considerar com a vàlid per aquell esquema. Un exemple el podem trobar a l'hora de treballar amb SensorML on resulta difícil saber quines etiquetes hi ha disponibles en cada cas o quines són obligatòries.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema targetNamespace="http://www.opengis.net/sensorML/1.0.1" xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:gml="http://www.o
www.opengis.net/sensorML/1.0.1" xmlns:swe="http://www.opengis.net/swe/1.0.1" elementFormDefault="qualified" attributeFormDefault="unqu
        <xs:annotation>
                <xs:documentation>SensorML document root definition

                        SensorML is an OGC Standard.
                        Copyright (c) 2007,2010 Open Geospatial Consortium, Inc. All Rights Reserved.
                        To obtain additional rights of use, visit http://www.opengeospatial.org/legal/ .
                </xs:documentation>
        </xs:annotation>
        <!-- ================================================== -->
        <xs:import namespace="http://www.opengis.net/gml" schemaLocation="http://schemas.opengis.net/gml/3.1.1/base/gml.xsd"/>
        <xs:import namespace="http://www.opengis.net/swe/1.0.1" schemaLocation="http://schemas.opengis.net/sweCommon/1.0.1/swe.xsd"/>
        <xs:include schemaLocation="./system.xsd"/>
        <!-- ================================================== -->
        <xs:element name="SensorML">
                <xs:annotation>
                        <xs:documentation>SensorML document root</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:group ref="sml:metadataGroup"/>
                                <xs:element name="member" maxOccurs="unbounded">
                                        <xs:complexType>
                                                <xs:choice>
                                                        <xs:element ref="sml:_Process"/>
                                                        <xs:element ref="sml:DocumentList"/>
                                                        <xs:element ref="sml:ContactList"/>
                                                </xs:choice>
                                                <xs:attributeGroup ref="gml:AssociationAttributeGroup"/>
                                        </xs:complexType>
                                </xs:element>
                        </xs:sequence>
                        <xs:attribute name="version" type="xs:token" use="required" fixed="1.0.1"/>
                </xs:complexType>
        </xs:element>
```

Figura 11: SensorML XSD

Alguns llenguatges que es promulguen com alternatives a XML són “JavaScript Object Notation”, JSON [28] i “YAML Ain't Another Markup Language”, YAML [29].

2.3 RDF i RDF-SCHEMA

Una vegada que amb XML ja hi ha un llenguatge que permet representar la estructura de les dades, “Resource Description Framework”, RDF [30] permet afegir contingut semàntic als seus elements. RDF és un “Framework” per metadades a la web desenvolupat per W3C. RDF va ser desenvolupat per Ramanathan V. Guha quan treballava en Apple Computer.

RDF és la primera aproximació a un llenguatge que permeti expressar contingut semàntic a la web i descriure recursos a la web.

La idea de RDF és que les declaracions dels recursos siguin expressions de la forma Subjecte-Predicat-Objecte. El conjunt d'aquests tres elements en RDF es coneix com a “**Tripleta”**.

Figura 12: "Tripleta” RDF

Una “tripleta” consta dels següents elements:

- El subjecte: és el recurs, el que es descriu, s'ha de poder identificar inequívocament (amb URI). La URI pot ser una URL o un URN. En el marc de treball del projecte seria un sensor, com l’estació meteorològica o el vaixell per exemple.
- El predicat: és la propietat o relació que es descriu sobre el subjecte o recurs en relació a l'objecte. Associa el valor d'una propietat a un recurs o subjecte. En el marc de treball seria per exemple: “observat per”, “calculat per”.
- L'objecte: és el valor de la propietat a la que es refereix el predicat o bé un altre subjecte. Té associada una URI i un significat concret.

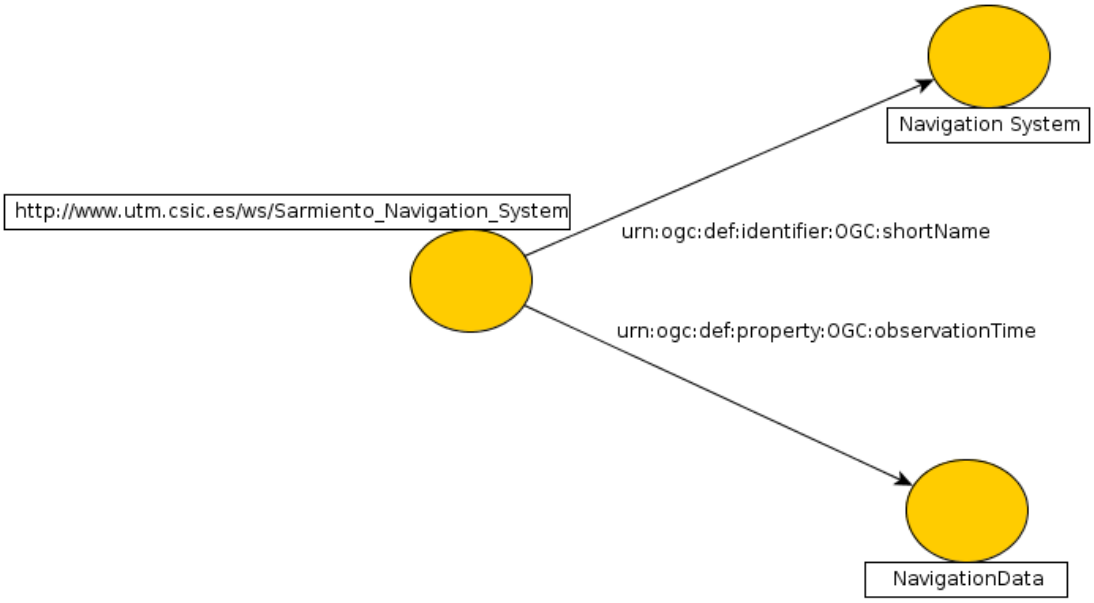


Figura 13: Exemple de "Tripleta" del sistema de navegació

El model RDF es representa en forma de grafs dirigits on el node origen és una URI, l'arc també és una URI i el node destí pot ser una URI o un literal.

RDF-Schema manté les característiques de RDF però soluciona algunes de les limitacions d'aquest, a l'afegir-li semàntica de manera que el converteix en una ontologia.

2.4 OWL

OWL (Web Ontology Language) [31] és un llenguatge per a ontologies a la web i ja des del seu origen està dissenyat per utilitzar-lo a la web en general i a la web semàntica en particular. OWL va un pas més enllà que RDF-Schema de manera que permet descriure relacions complexes entre classes i establir restriccions a les classes i a les propietats.

OWL aprofita les capacitats de RDF i les estén amb les següents capacitats:

- Capacitat de ser distribuït entre sistemes diferents.
- Escalable de manera que dóna resposta a les necessitats de la web.
- És compatible amb els estàndards web d'accessibilitat i internacionalització.
- Permet crear ontologies obertes i extensibles.

OWL partint de RDF i RDF Schema afegeix més vocabulari per descriure les propietats i les classes. Seria el cas de les relacions entre classes (per exemple: "disjointness"), cardinalitat (per exemple: "exactly one").

A través d'una ontologia és possible descriure tots el conceptes relacionats amb un determinat domini i les relacions que poden haver entre aquests conceptes.

Hi ha tres "sabors" o nivells de OWL:

1. **OWL-Lite**: és la més senzilla i limitada, fa servir només algunes de les capacitats de OWL. Per exemple, en OWL-Lite les classes només poden ser definides en termes de noms de superclasses i només es poden fer servir algunes de les restriccions de classe. La única cardinalitat que permet definir de forma explicita és 0 o 1.
2. **OWL-DL** ("Description Logic"): és el nivell indicat si són necessàries majors capacitats d'expressió del llenguatge i a la vegada mantenir la integritat computacional.
3. **OWL-Full**: és el nivell indicat si són necessàries les màximes capacitats d'expressió del llenguatge però sense mantenir les garanties computacionals.

2.5 Wikis Semàntiques

La paraula wiki prové del Hawaia i significa ràpid. Una wiki és un conjunt de pàgines web que poden ser editades a través d'un navegador web. Això facilita l'intercanvi de coneixement i correspon a un model de web on l'usuari no és passiu i contribueix d'una forma directa al contingut d'aquestes pàgines web, corresponent al que es coneix com a Web 2.0.

Una **wiki semàntica** és una wiki que té un model de coneixement subjacent descrit a les seves pàgines [32]. Aquest model de coneixement cal que estigui disponible en un llenguatge formal, de manera que sigui possible que màquines o sistemes puguin accedir a aquest coneixement i processar-lo en un model de base de dades relacional. Com a llenguatge formal s'entén aquell llenguatge on els símbols primitius i les regles per unir aquests símbols estan formalment especificats.

A diferència de les wikis normals o sintàctiques, les wikis semàntiques permeten capturar i identificar informació sobre les dades que hi ha a les seves pàgines i les relacions que poden haver entre diferents pàgines. Tot això fa possible que les wikis semàntiques siguin consultades i exportades tal i com si es tractés d'una base de dades.

La gran expansió que han experimentat les wikis fa que cada vegada més usuaris estiguin familiaritzats amb el seu funcionament i això fa que usuaris no especialitzats, a través d'un simple navegador web, puguin accedir a la informació de la wiki d'una forma amigable i intuïtiva.

2.6 Projectes relacionats

SWE

Sensor Web Enablement, SWE [19] és una iniciativa del “Open Geospatial Consortium”,OGC [18]. que busca la definició d'un marc de serveis sobre els sensors que facilitin la interoperabilitat, en dos aspectes. En primer lloc en la definició del propi sensor:

- Que fa el sensor.
- A quina organització pertany.
- Quina és la freqüència de refresc de les dades.
- Es tracta d'un sensor que està en moviment o bé es tracta d'un sensor fix.

En segon lloc, en l'accés a les dades adquirides pel sensor, ja sigui mitjançant preguntes simples del tipus retornem les 10 observacions més recents o bé més complexes en les que es pot interrogar per exemple quines dades hi ha en una regió determinada de l'espai. Es tracta d'un camp de treball molt nou encara ara però que ha d'experimentar una gran expansió de l'estil del que han tingut els serveis sobre mapes: “web map services” i “web features services” (WMS i WFS).

Per descriure el sensor dintre de SWE es fa servir el llenguatge de modelat de sensors, SensorML [20] i pel que fa a la descripció de les observacions es fa servir “observations and measurements”, O&M [21].

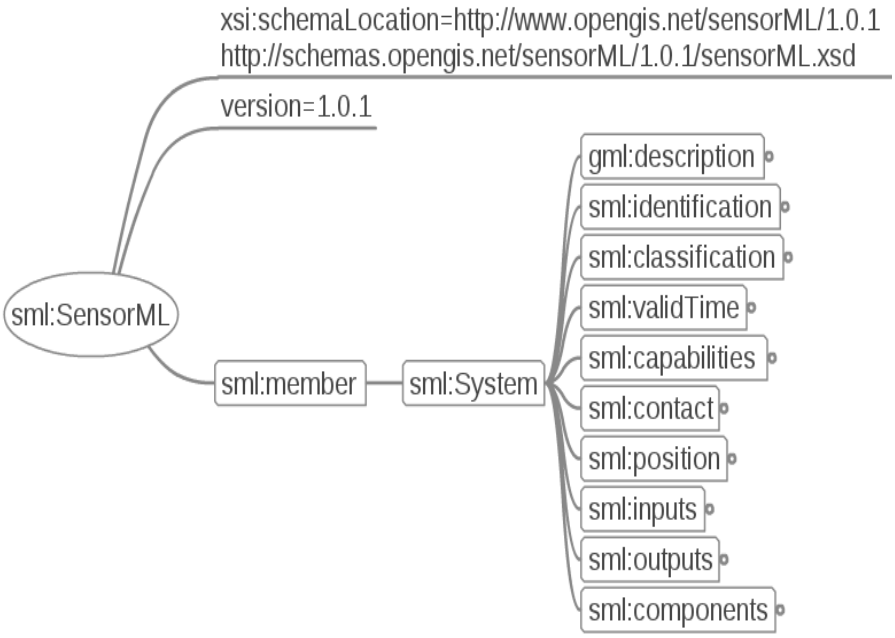


Figura 14: Esquema SensorML

En el context de la investigació oceanogràfica realitzada per vaixells intervenen una gran quantitat de sensors diferents que adquireixen dades molt variades. Aquest escenari dóna lloc a que les descripcions dels sensors en SensorML i de les pròpies dades en O&M, donin lloc extensos fitxers XML.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<om:OM_Observation gml:id="SARMIENTO_WEATHER_STATION_2012-11-02-00-00-00"
    xmlns:om="http://www.opengis.net/om/2.0" xmlns:swe="http://www.opengis.net/swe/2.0"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:gml="http://www.opengis.net/gml/3.2"
    xsi:schemaLocation="http://www.opengis.net/om/2.0 http://schemas.opengis.net/om/2.0/observation.xsd
    http://www.opengis.net/swe/2.0 http://schemas.opengis.net/sweCommon/2.0/swe.xsd">

    <!-- Description  -->
    <gml:description>Instant Weather Observation made onboard Sarmiento de Gamboa vessel</gml:description>

    <!-- Type -->
    <om:type
        xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_ComplexObservation"/>

    <!-- Phenomenon Time -->
    <om:phenomenonTime>
        <gml:TimeInstant gml:id="phenomenonTime">
            <gml:timePosition>2012-06-29T17:22:25</gml:timePosition>
        </gml:TimeInstant>
    </om:phenomenonTime>
    <om:resultTime xlink:href="#phenomenonTime"/>
```

Figura 15: Fragment de fitxer O&M

Figura 16: Estructura de fitxer O&M

Tot i que les descripcions dels sensors en SensorML i de les dades en O&M són molt flexibles i es pot arribar fàcilment al grau de detall desitjat, aquestes no contenen cap tipus de semàntica. Per tant, poden ser de gran interès les iniciatives que acosten el món de SWE i el de la web semàntica, tal com és el cas de ”Linked Sensor Data”.

Com a exemple de l'interès d'aproximar el món de SWE i de la web semàntica es pot citar l'article "Semantically-Enabled Sensor Plug & Play for the Sensor Web" [47]. En aquest article els autors presenten una aproximació que permet:

- Ampliar SWE amb funcionalitats semàntiques.
- Presentar un mecanisme de publicació - subscripció per sobre de SWE.
- Presentar un model que permeti declaracions descriptives de les interfícies dels sensors.

Linked Sensor Data

És important tenir en conte el projecte “Linked Sensor Data” [17], que fa tota una sèrie de recomanacions de cara a exposar, compartir i connectar fragments de dades, informació i coneixement que hi ha a la web, fent servir URIs i RDF.

Actualment podem trobar una gran varietat de sensors adquirint dades, ja sigui en el marc dels vaixells d'investigació oceanogràfica o a la vida quotidiana. Avui en dia els sensors han passat de ser utilitzats només per usuaris experts com poden ser els científics a formar part de la vida quotidiana dels ciutadans, de manera que de forma natural ha calgut buscar com fer accessibles les dades proporcionades pels sensors a través de la web.

Si examinem per exemple un dels projectes que aplica els principis de “linked data” com és el de Harshal Patni, Cory Henson, Amit P. Sheth del Ohio Center of Excellence in Knowledge enabled Computing (Kno.e.sis) [17] tenim que els objectius que es pretenen assolir amb aquest projecte són:

- Fer accessibles conjunts de dades provinents d’estacions meteorològiques.
- Que aquests conjunts de dades siguin fàcilment consultables, “querys”.
- Fer possible la consulta de dades dels sensors en base a noms de llocs d'on es troba situat el sensor.
- Implantar el programari que transformi els dades en format O&M a RDF, “O&M2RDFConverter”, de manera que s'estableixi aquest enllaç entre el món OGC de SWE amb el de RDF de la web semàntica.

Del que es tracta, és fer servir RDF com a llenguatge per representar aquests sets de dades. Concretament el que es busca és descriure els sensors que adquireixen les dades i les observacions realitzades per aquests sensors. Aquest dos objectius, a grans trets és el mateix que es busca amb la iniciativa de SWE [19] de OGC [18], però amb ”Linked Sensor Data” es tracta d'aprofitar les possibilitats de la web semàntica, de manera que es defineix tot un flux de dades:

1. En un primer moment les dades són adquirides pel sensor.
2. Aquestes dades s'incorporen al sistema en forma de O&M de SWE.
3. Els fitxers O&M es converteixen a instàncies RDF (O&M2RDFCONVERTER).
4. Per passar finalment a estar emmagatzemades a una base de dades semàntica, Virtuoso.

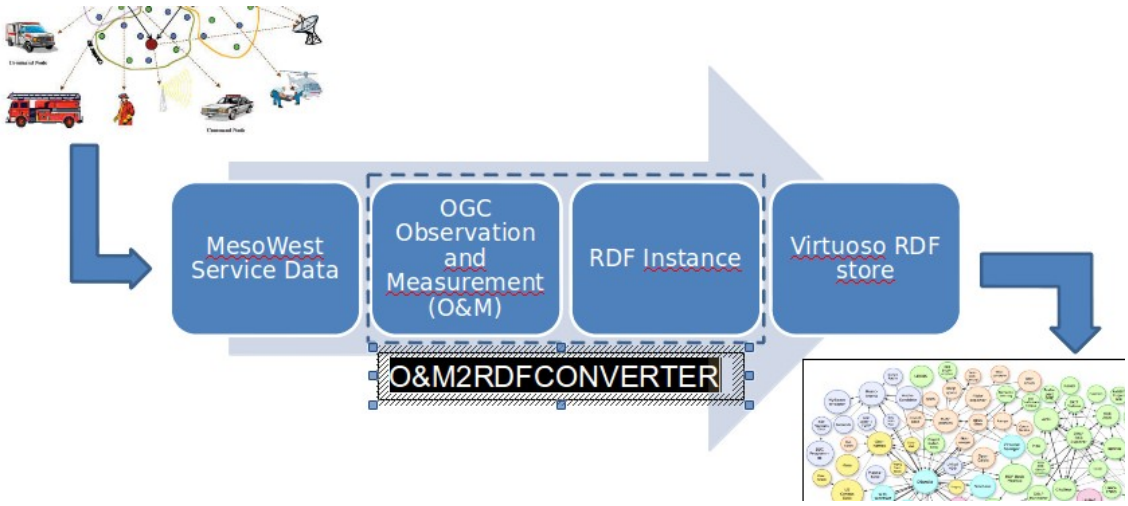


Figura 17: Linked Sensor Data Workflow [17]

A través d'aquest flux per a cada observació del sensor se l'hi associa un lloc (“ObservationLocation”), un moment temporal de mostreig (“SamplingTime”), un conjunt de dades (“result”) i un objecte d’observació (“Feature”).

3. Disseny de l'ontologia per vaixells d'investigació oceanogràfica

Una ontologia pot ser una eina molt útil per vaixells d'investigació oceanogràfica ja que permet:

- Definir termes i conceptes: es poden definir vocabularis i espais de noms segons disciplines i necessitats. Resulta especialment útil emprar definicions estàndards tal i com pot ser el cas de CF [45] o UOM [46] o bé desenvolupar vocabularis personalitzats que permetin descriure les particularitats de cada escenari concret.
- Estructurar relacions entre termes: de manera que es puguin establir relacions entre variables i conceptes. Un exemple podria ser:

"La temperatura superficial de l'aigua és un subset de la temperatura de l'aigua"

- Fer que aquesta informació sigui directament interoperable entre màquines i sistemes: emprant URIs que apunten directament a les definicions.
- Informació estable i reutilitzable: permet fer servir les mateixes definicions per fitxers NetCDF i SensorML.

A l'hora d'avaluar la idoneïtat d'emprar ontologies a l'àmbit marí cal tenir en compte diferents aspectes:

1. Consideracions tècniques: quin tipus d'ontologia o vocabulari és el més adient, ja que hi han moltes opcions.
    1. Tipus:
        1. Plans: diccionaris, glossaris, "code list".
        2. Multinivell: taxonomies.
        3. Relacionals: Thesaurus, **ontologies**.
    2. Formats:
        1. Simples com CSV i XML.
        2. RDF.
        3. SKOS.
        4. **OWL**.
    3. Mètodes d'accés:
        1. Usuari final.
        2. "Machine to machine".

En aquest treball s'ha considerat que per la potència i flexibilitat que proporciona resulta més adient una aproximació emprant ontologies expressades en OWL, de manera que es faciliti la seva explotació de cara a l'usuari final a través de wikis semàntiques, però tal i com es pot veure al resum anterior les solucions poden ser molt variades.

3.1 Anàlisi d'ontologies oceanogràfiques

A continuació es repassen algunes de les principals iniciatives per crear ontologies en el camp de la oceanogràfica:

The MMI Ontology Server

A la URL: http://mmisw.org/orr/#b es pot consultar online "The MMI Ontology Registry and Repository". Consisteix en una aplicació web que permet crear, actualitzar ontologies i termes dins de l'àmbit marí. A través de pàgina web es pot accedir al catàleg d'ontologies i fer cerques de manera que finalment es pot consultar una fitxa amb metadades de l'ontologia seleccionada i accedir a la definició de la mateixa.

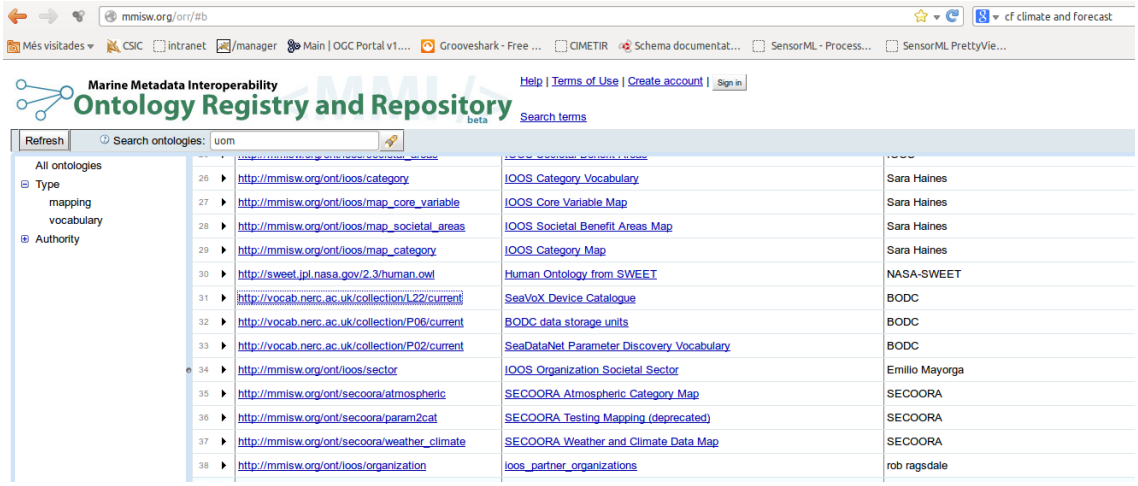
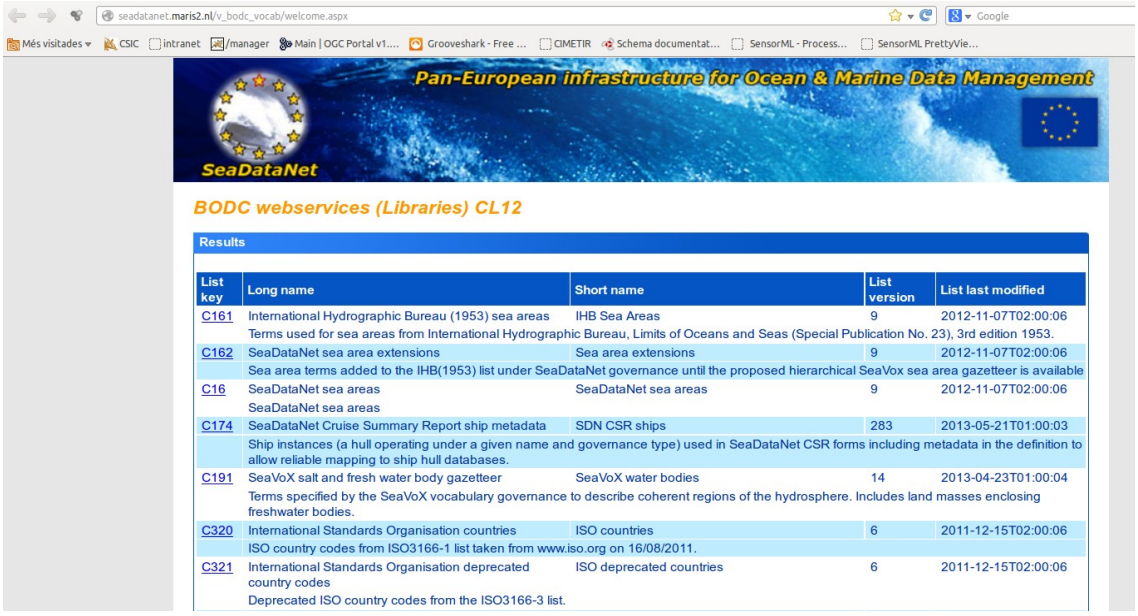


Figura 18: MMI Ontology Registry and Repository

SeaDataNET

SeaDataNet és un important projecte Europeu dins de la investigació oceanogràfica que pretén crear un marc de treball conjunt europeu en el que diferents institucions hi puguin intercanviar i compartir les dades i les metadades obtingudes en els decurs de campanyes d'investigació oceanogràfica.

SeaDataNet ha desenvolupat un potent i ampli vocabulari [48] que per contra es difícil de navegar i on resulta complicat trobar els termes, que per altra banda tenen noms molt críptics que resulta difícil relacionar amb els conceptes que defineixen. També és remarcable que com a URI es fa servir molt més URNs, tal i com es pot veure a la següent figura on es fa servir el vocabulari de SeaDataNet per descriure els "inputs" dins d'un document de SensorML que descriu un termosalinògraf.



List key	Long name	Short name	List version	List last modified
C161	International Hydrographic Bureau (1953) sea areas Terms used for sea areas from International Hydrographic Bureau, Limits of Oceans and Seas (Special Publication No. 23), 3rd edition 1953.	IHB Sea Areas	9	2012-11-07T02:00:06
C162	SeaDataNet sea area extensions Sea area terms added to the IHB(1953) list under SeaDataNet governance until the proposed hierarchical SeaVox sea area gazetteer is available	Sea area extensions	9	2012-11-07T02:00:06
C16	SeaDataNet sea areas SeaDataNet sea areas	SeaDataNet sea areas	9	2012-11-07T02:00:06
C174	SeaDataNet Cruise Summary Report ship metadata Ship instances (a hull operating under a given name and governance type) used in SeaDataNet CSR forms including metadata in the definition to allow reliable mapping to ship hull databases.	SDN CSR ships	283	2013-05-21T01:00:03
C191	SeaVoX salt and fresh water body gazetteer Terms specified by the SeaVoX vocabulary governance to describe coherent regions of the hydrosphere. Includes land masses enclosing freshwater bodies.	SeaVoX water bodies	14	2013-04-23T01:00:04
C320	International Standards Organisation countries ISO country codes from ISO3166-1 list taken from www.iso.org on 16/08/2011.	ISO countries	6	2011-12-15T02:00:06
C321	International Standards Organisation deprecated country codes Deprecated ISO country codes from the ISO3166-3 list.	ISO deprecated countries	6	2011-12-15T02:00:06

Figura 19: Vocabulari SeaDataNet

Al document “Semantic Interoperability and SeaDataNet” presentat a l’ICAN Workshop [49] es fa palesa la preocupació per part de SeaDataNet per dotar al seu projecte dels avantatges i potencials de la web semàntica. Així es parla de que molts dels camps de metadades que es fan servir s'omplen de vocabularis controlats, però aquests tenen diversos problemes:

- Són difícils de gestionar
- Falten definicions

Els esquemes de metadades de SeaDataNet incorporen “Schematron” de manera que faciliten la validació dels continguts contra uns “vocabularis mestres”.

```
<sml:inputs>
    <sml:InputList>
        <sml:input name="Electrical conductivity of the water sample">
            <swe:ObservableProperty definition="urn:SDN:P021:51:CNDC"/>
        </sml:input>
        <sml:input name="Temperature of the water sample">
            <swe:ObservableProperty definition="urn:SDN:P021:73:TEMP"/>
        </sml:input>
        <sml:input name="Chlorophyll concentration in the water sample">
            <swe:ObservableProperty definition="urn:SDN:P021:73:CPWC"/>
        </sml:input>
    </sml:InputList>
```

Figura 20: Termes de SeaDataNet com a URNs dins de SensorML

A la figura anterior es poden veure els elements que composen els termes de SeaDataNet o “SDN Semantic Markup”. Així si ens fixem en la cadena:

SDN:P021:21:TEMP

Podem distingir:

- SDN: identificador de namespace.
- P021: identificador de recurs.
- 21: identificador del número de versió del recurs.
- TEMP: identificador del component del recurs.

Aquesta URN es pot transformar en URL amb les següents substitucions (actualment hi ha una certa tendència a preferir utilitzar URLs enfront de URNs, ja que resulta més senzill navegar fins les definicions):

- http://vocab.ndg.nerc.ac.uk/ = namespace
- term/P021 = recurs.
- /21 = versió.
- /TEMP = identificador del component del recurs.

De manera que expressat en forma de URL queda de la següent manera:

http://vocab.ndg.nerc.ac.uk/term/P021/21/TEMP

Al fer la petició al servidor de vocabularis NDG amb l'identificador del recurs, aquest retornarà un document SKOS on es descriu el concepte i els diferents mappings.

EMODNET

“European Marine Observation and Data Network”, EMODNET és una important iniciativa que pretén donar un marc de treball per tal d'intercanviar dades marines per tal que es promogui la interoperabilitat entre organitzacions de recerca Europea.

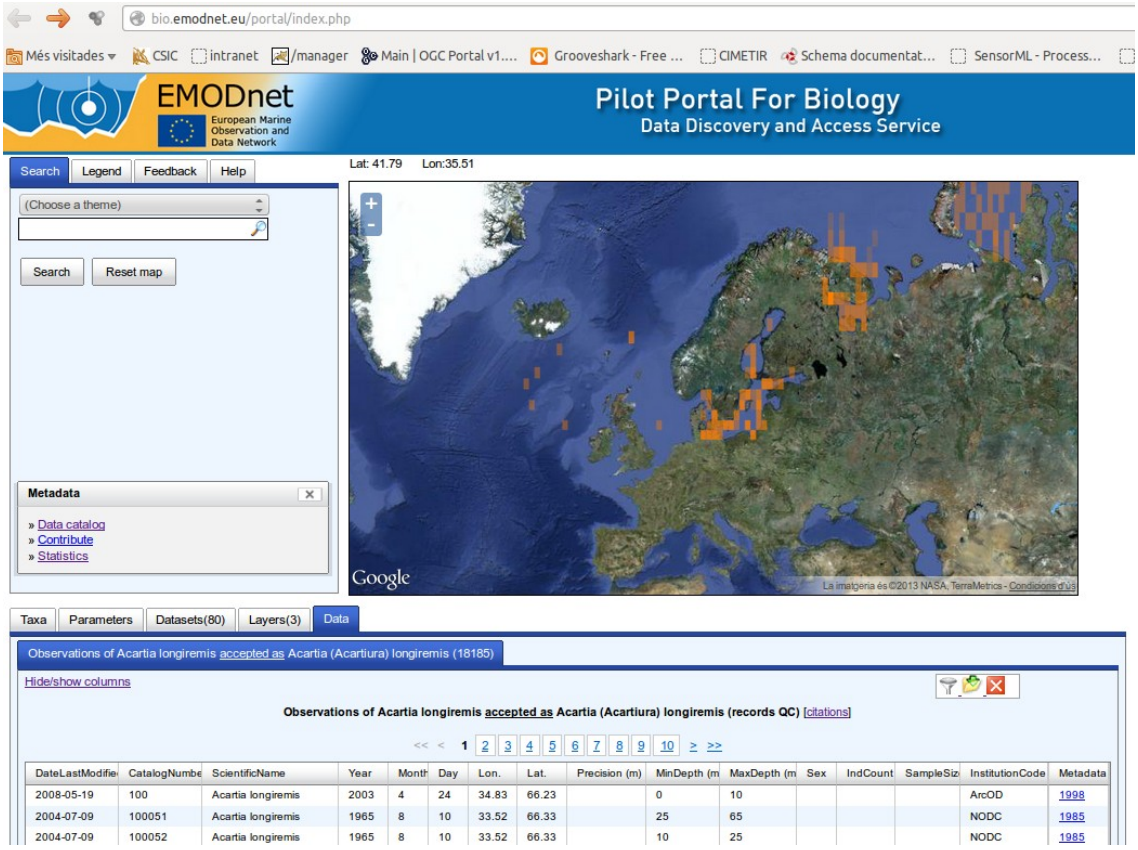


Figura 21: Portal EMODNET

3.2 Eines de desenvolupament.

Protégé

Per tal de treballar amb les ontologies s'ha optat per l'eina Protégé [22], concretament es fa servir la versió 3.4.8 i la 4.3. Protégé es distribueix com a programari lliure sota una llicència "Mozilla Public License". També s'ha fet servir l'extensió OWLVIZ com a eina de visualització de les classes. OWL es distribueix sota llicència GNU.

Hi ha dues maneres de modelar amb Protégé, a través de Protégé-Frames o de Protégé-OWL i en tots dos casos es pot exportar el model a RDF(S) , OWL i a XML Shema.

Protégé té una àmplia comunitat d'usuaris i desenvolupadors, doncs la salut del projecte és molt bona.

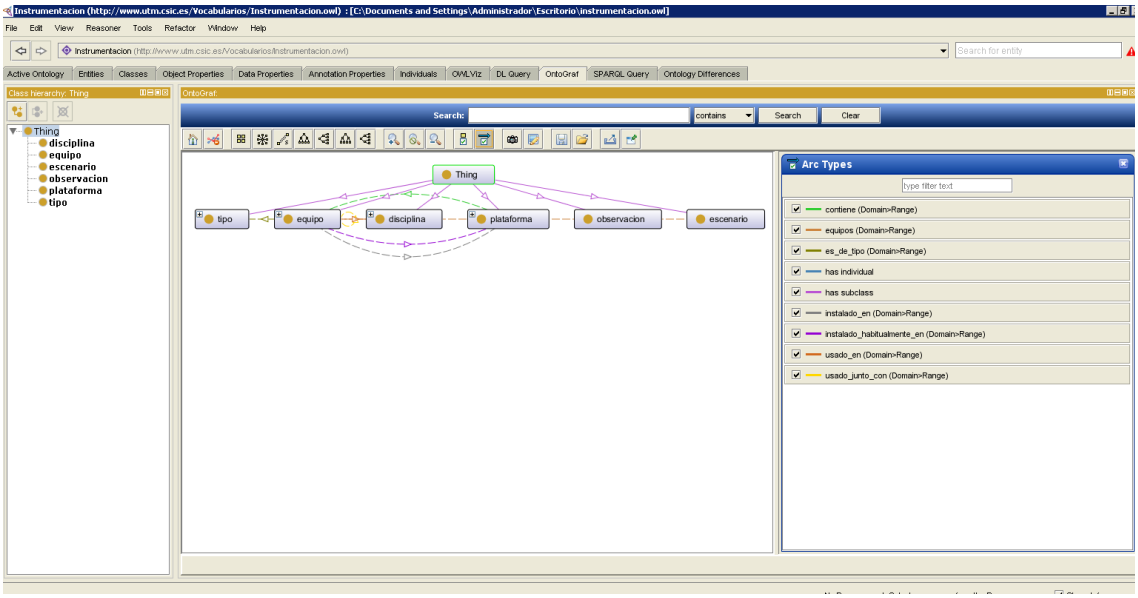


Figura 22: Protégé

S'ha trobat dificultats a l'hora de treballar amb la versió 4.2 de Linux (Ubuntu 12.04). Protégé quedava penjat a l'intentar obrir l'ontologia. Aquest comportament no es produïa sempre i a vegades aconseguia obrir-la. S'ha optat per instal·lar la mateixa versió 4.2 en una màquina virtual de Windows on no ha aparegut aquest problema. Per poder fer servir l'extensió OWLViz ha estat necessari instal·lar graphviz.

També s'ha fet servir l'eina "Generate OWLDoc" que permet generar un conjunt de pàgines web amb la informació de l'ontologia.

NeOn Toolkit

També s'ha avaluat l'eina Neon Toolkit [50] que tot i presentar menys funcionalitats que Protégé, potser per aquest mateix fet es mostra senzilla d'utilitzar i útil per la qual cosa s'ha combinat amb Protégé per treballar amb l'ontologia quan s'ha estimat oportú.

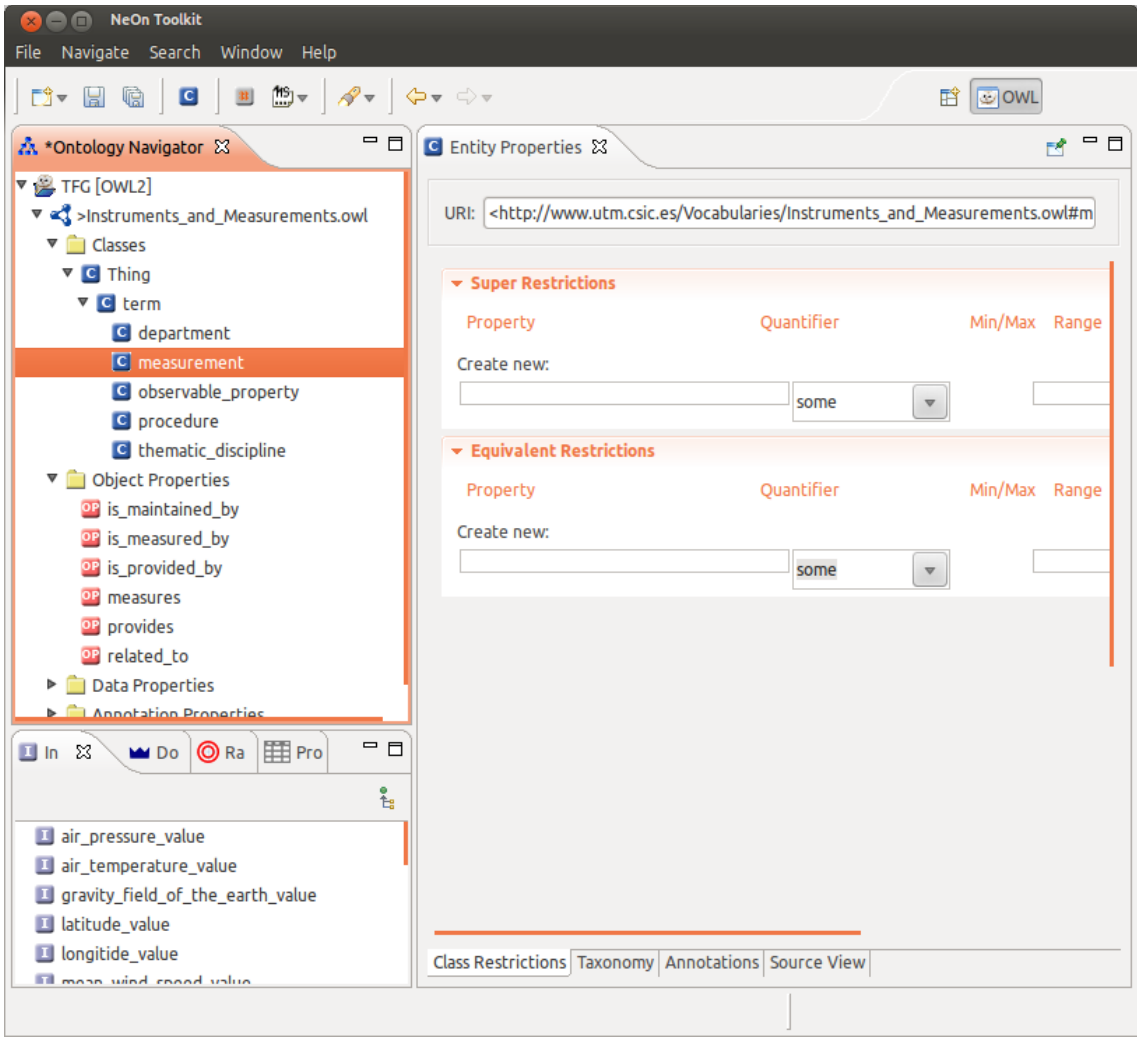


Figura 23: NeOn Toolkit

3.3 Requeriments de l'ontologia

A continuació es passa a descriure els requeriments que ha de complir l'ontologia. La recollida de requisits és un procés clau en tot projecte l'objectiu del qual és identificar i gestionar els requisits del programari a desenvolupar.

Domini de l'ontologia:

L'ontologia ha de descriure i modelitzar la instrumentació marina operada en el context de campanyes d'investigació oceanogràfica. El present treball es limita a la modelització de la instrumentació oceanogràfica desplegada en vaixells d'investigació, tot i que poden haver altres escenaris que queden fora de l'abast del present treball, tal i com és el cas dels observatoris submarins.

Propòsit de l'ontologia:

L'ontologia ha de permetre accedir a la informació rellevant de la instrumentació, saber a quina plataforma es troba desplegada, si es troba o no operativa, amb quina disciplina està relacionada o quins departaments la gestionen. Aquesta informació es clau per gestionar dos dels esdeveniments més importants en la gestió de la instrumentació:

- Campanyes oceanogràfiques
- Manteniment i/o varada

[veure Glossari]

Cal preveure la integració de tots els serveis de dades i metadades que hi pot haver en l'àmbit de la investigació oceanogràfica. Entre aquestes serveis hi ha:

- **Web Map Service** (WMS) [39]: Servei OGC que produeix mapes de dades referenciades espacialment de forma dinàmica a partir d'informació geogràfica.
- **Web Feature Service** (WFS) [40]: Servei OGC que defineix operacions web de interfície per la consulta i edició d'entitats geogràfiques ("features") vectorials.
- **Web Coverage Service** (WCS) [43]: Servei OGC també conegut com a servei de cobertura web, proporciona una interfície que permet realitzar peticions de cobertura geogràfica a través de la web.

Figura 24: Serveis OGC WMS, WFS i WCS al vaixell

- **Sensor Observation Service**: SOS [41]. Servei web que permet un accés estandarditzat a les observacions adquirides pel sensor i a les descripcions del mateix sensor.

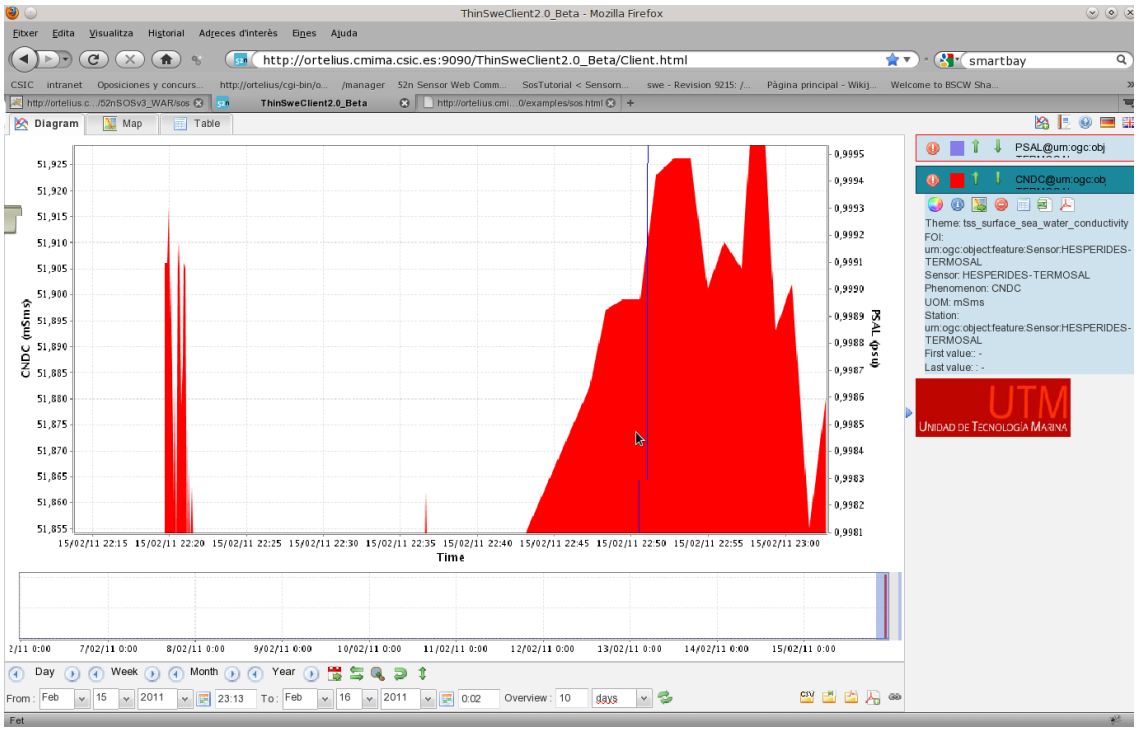


Figura 25: Client Web SOS al vaixell

- **Catalog Service for the Web:** CSW [42]. Servei OGC que defineix operacions de descobriment, cerca i consulta de metadades relacionades amb dades, serveis i recursos de tipus geogràfic.
- **OPeNDAP** [38]: acrònim de "Open-source Project for a Network Data Access Protocol". Es tracta d'un protocol i una arquitectura de transport de dades àmpliament utilitzat en el món científic.

Figura 26: OpeNDAP a través del servidor Thredds al vaixell

- SeadataNet **CDI/CSR** [44] Services: Seadatanet és un projecte europeu l'objectiu del qual és compartir dades i metadades entre països i institucions de manera que quedi emmagatzemat en un servei central la informació de quines campanyes s'han realitzat en cada moment, amb quins sensors i quin tipus de dades s'ha recollit.

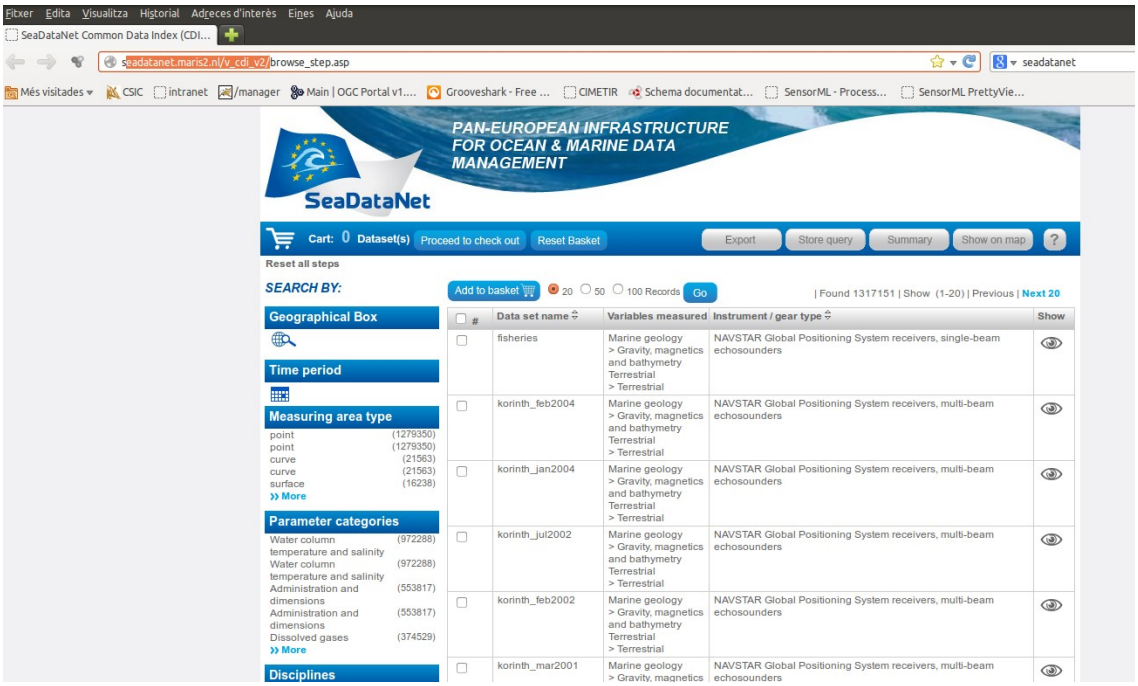


Figura 27: Portal SeaDataNet de consulta de CDIs

Un dels propòsits de l'ontologia a desenvolupar és poder arribar als serveis que proporcionen les dades i metadades.

Usuaris de l'ontologia:

Els usuaris de l'ontologia o "stakeholders", entenent com a "stakeholders" aquelles persones o grups de persones que tenen algun interès en el sistema., són els següents:

- Tècnics UTM: són els propis tècnics encarregats de la gestió i manteniment, operen els instruments o sensors.

- Personal científic: són els membres de la comunitat científica en general, siguin de la organització que sigui i de qualsevol país i que hagi de consultar les dades obtingudes en una campanya o veure quines capacitats o necessitats pot tenir un instrument de cara a la preparació d'una campanya d'investigació oceanogràfica.

3.4 Elaboració del model conceptual

A continuació es passa a definir un petit glossari de termes que es fan servir àmpliament dins de l'àmbit de treball del present projecte i que resulta clau deixar ben acotats de cara a fer-los servir posteriorment.

Definició de termes de l'ontologia

- **thematic_discipline**: la ciència o camps de la ciència interessada en la investigació oceanogràfica i en un determinat punt de vista a l'hora de realitzar les observacions i fer servir uns sensors en concret.

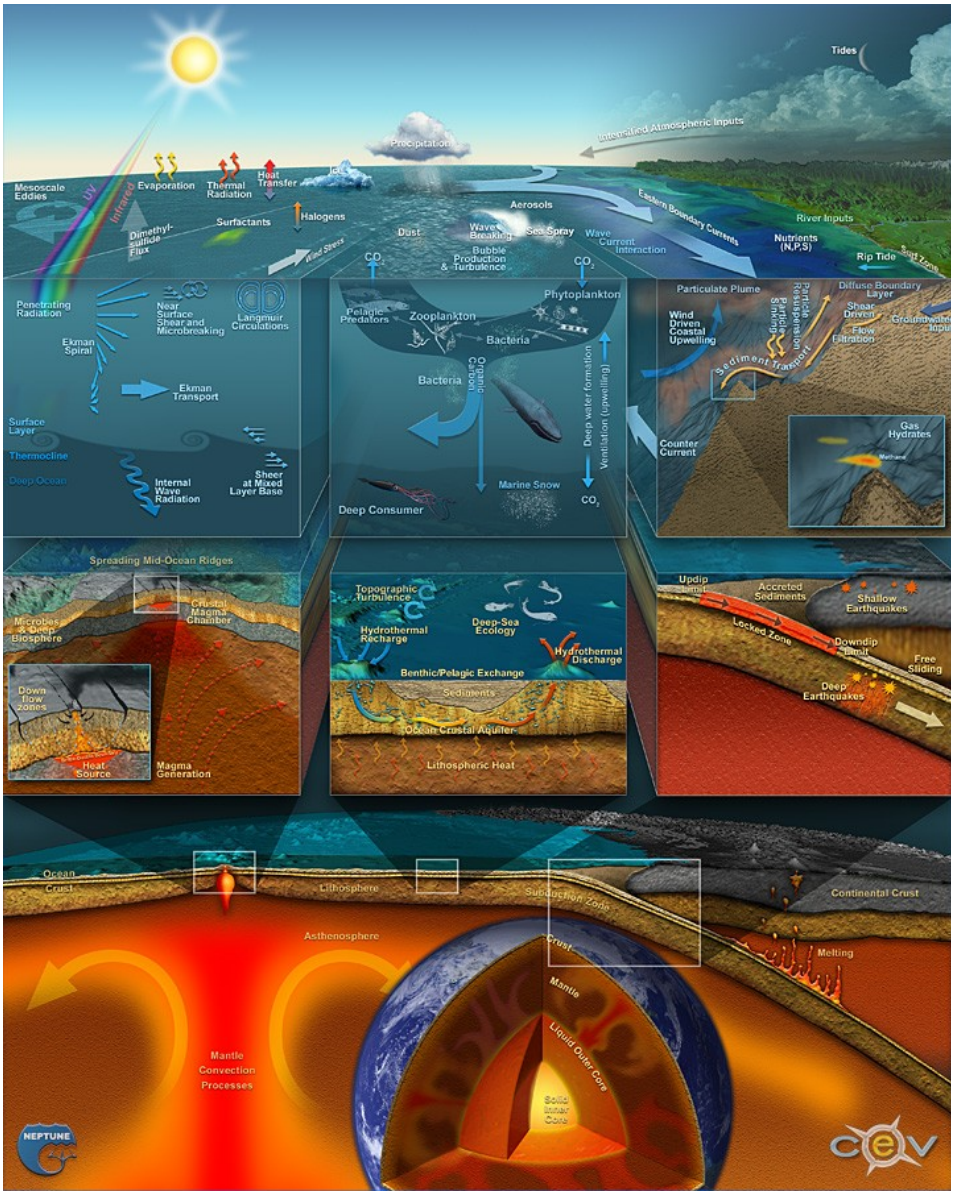


Figura 28: Múltiples disciplines implicades (Font:neptunecanada.com)

- **procedure**: com a “procedure” entenem el sensor o instrument encarregat de l’adquisició de dades. Un sensor a la seva vegada pot estar format per més altres sensors (sistema de sistemes). Seria per exemple el cas del termosalinògraf, integrat per sensors de temperatura i conductivitat.

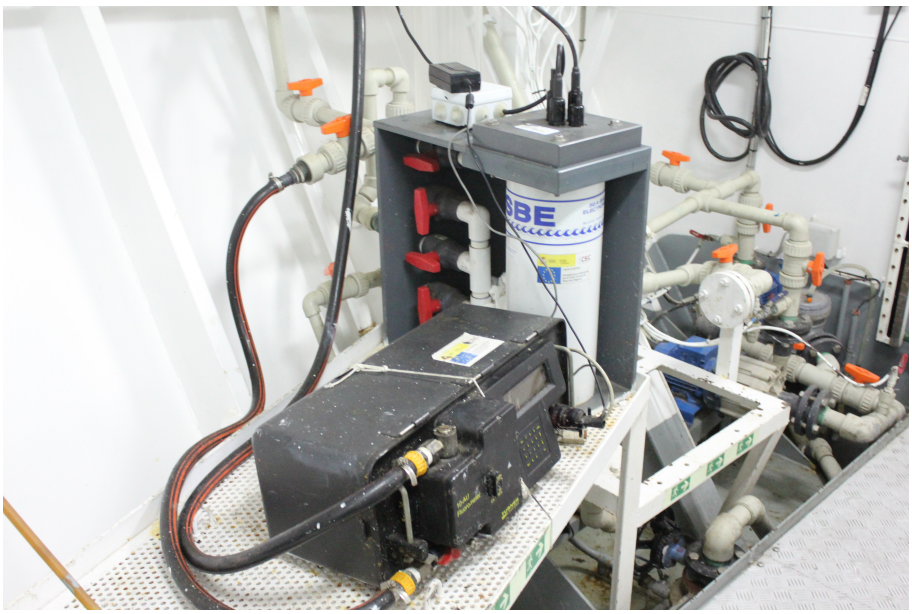


Figura 29: Termosalinògraf

- **observable_property**: variable física observada. Seria el cas de la temperatura, conductivitat o profunditat per exemple.
- **departament**: correspon a organització interna dins de la UTM en departaments especialitzats, de manera que cadascun d'ells té assignades una sèrie de tasques i equips (“procedures“) a realitzar i mantenir.
- **measurement**: al mesurar un instrument (“procedure”) una determinada propietat física (“observable_property”) dóna lloc a una mesura. És el cas de la velocitat sobre l'aigua, la velocitat sobre el fons, o el valor de la temperatura de l'aigua, per exemple.
- **information_service**: associat a cada sensor o “procedure” es defineix un servei de dades, que és accessible a través de una URI i mitjançant el que es pot accedir a les dades adquirides per aquest instrument. La natura de les dades adquirides i el tipus de processat que es realitza posteriorment fa que s'utilitzi un o altre servei o fins i tot que la mateixa informació es pugui trobar en varis serveis diferents.

Definició de les classes i propietats de les classes

Partint dels termes definits a l'apartat anterior, es realitza una primera jerarquia de classes:

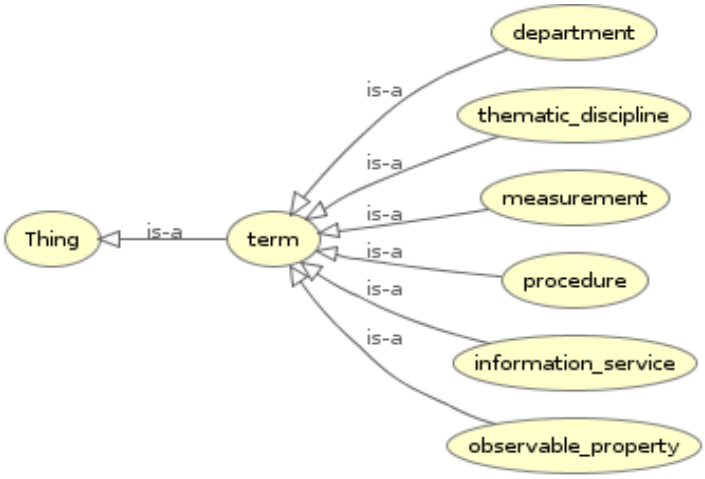


Figura 30: Diagrama de Classes

A la següent figura es poden veure les classes representades de l'ontologia en forma de núvol. La mida de les etiquetes és proporcional al seu ús i ordenat alfabèticament.

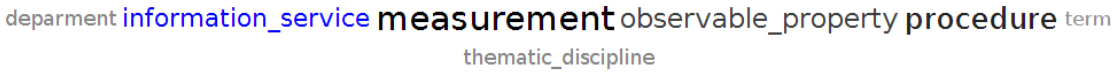


Figura 31: Classes i ús

Alguns dels termes emprats, com per exemple "procedure" i "observedProperty" provenen del món de SWE. Però les relacions entre entitats de vocabulari no es poden representar explícitament als fitxers de SensorML o O&M, tot i que això podria ser de gran utilitat de cara a poder accedir a la informació d'aquests fitxers.

Relacions entre objectes

A partir d'aquestes classes s'han definit les següents **relacions entre objectes:**

- **contains:** un sensor pot contenir altres sensors. Així un CTD conté sensors de conductivitat, temperatura i profunditat.
- **informations_service_at:** un "procedure" té associat un servei d'informació per consultar les dades que adquireix.
- **interested_in:** una disciplina temàtica està interessada en observar unes determinades propietats físiques relacionades amb el seu camp de treball.
- **is_contained_in:** un sensor pot estar format per altres sensors.
- **is_measured_by:** una propietat física és mesurada per un determinat sensor.
- **is_provided_by:** una mesura la proporciona un determinat sensor.
- **measures:** un instrument mesura una determinada propietat física.
- **Provides:** un instrument proporciona tota una sèrie de mesures.
- **related_to:** una propietat física esta relacionada amb una o més disciplines.
- **references_to:** un servei de dades (SOS, WMS, CSW, etcètera) fa referència a les dades adquirides per un o més sensors.
- **is_maintained_by:** el manteniment d'un sensor es realitzat per un departament especialitzat encarregat de realitzar les diferents tasques associades de reparació, calibratge i operació del sensor.
- **mantains:** per últim, un departament és l'encarregat de mantenir un sensor.

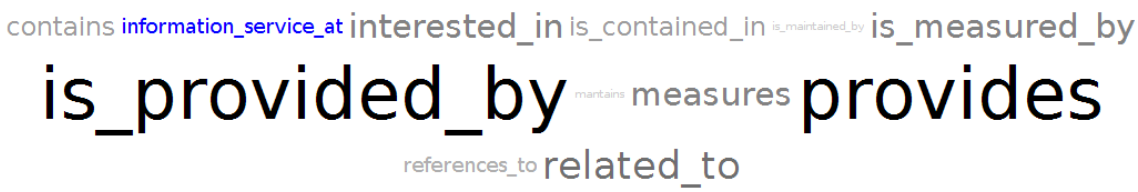


Figura 32: Relacions i ús

A la figura anterior es poden veure les relacions representades en forma de núvol. La mida de les etiquetes és proporcional al seu ús a l'ontologia i ordenat alfabèticament.

Així una “Thematic discipline” està “interested in” una “observed property” i una “observed property” està “related to” una “Thematic discipline”. També tenim que una “observed property” és “measured by” un “procedure” i un “procedure” “masures” una “observed property”.

Propietats de les dades

Com a propietats de les dades s'ha definit:

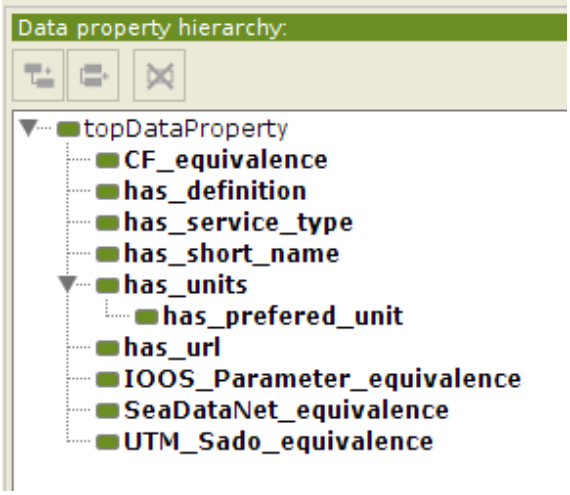


Figura 33: Data Properties

- **CF_equivalence**: indica l’equivalència del terme amb els termes de “Climate and Forecast” o CF [51].
- **has_definition**: definició del terme.
- **has_service_type**: tipus de servei: WMS, WFS, SOS per exemple.
- **has_short_name**: nom curt o abreviat.
- **has_units**: unitats de representació.
  - **has_prefered_units**: les unitats de representació preferides entre vàries de vàlides.
- **has_url**: URL del servei.

http://ortelius.cmima.csic.es:8080/geoserver/SDG/wms?service=WMS&version=1.1.0&request=GetMap&layers=SDG:termosal%20SDG&styles=&bbox=-79.952,17.357,14.56,43.158&width=1208&height=330&srs=EPSG:4326&format=application/openlayers

- **IOOS_Parameter_equivalence**: equivalència del terme a IOOS [52].
- **SeaDataNet_equivalence**: equivalència del terme a SeaDataNet. Seria del tipus: “SDN:P021:51:CDTA”.
- **UTM_Sado_equivalence**: equivalència del terme a nivell intern a la UTM, a nivell de programa o taules de base de dades. Seria del tipus:”presion_atm”.

Implementació del model conceptual amb Protégé

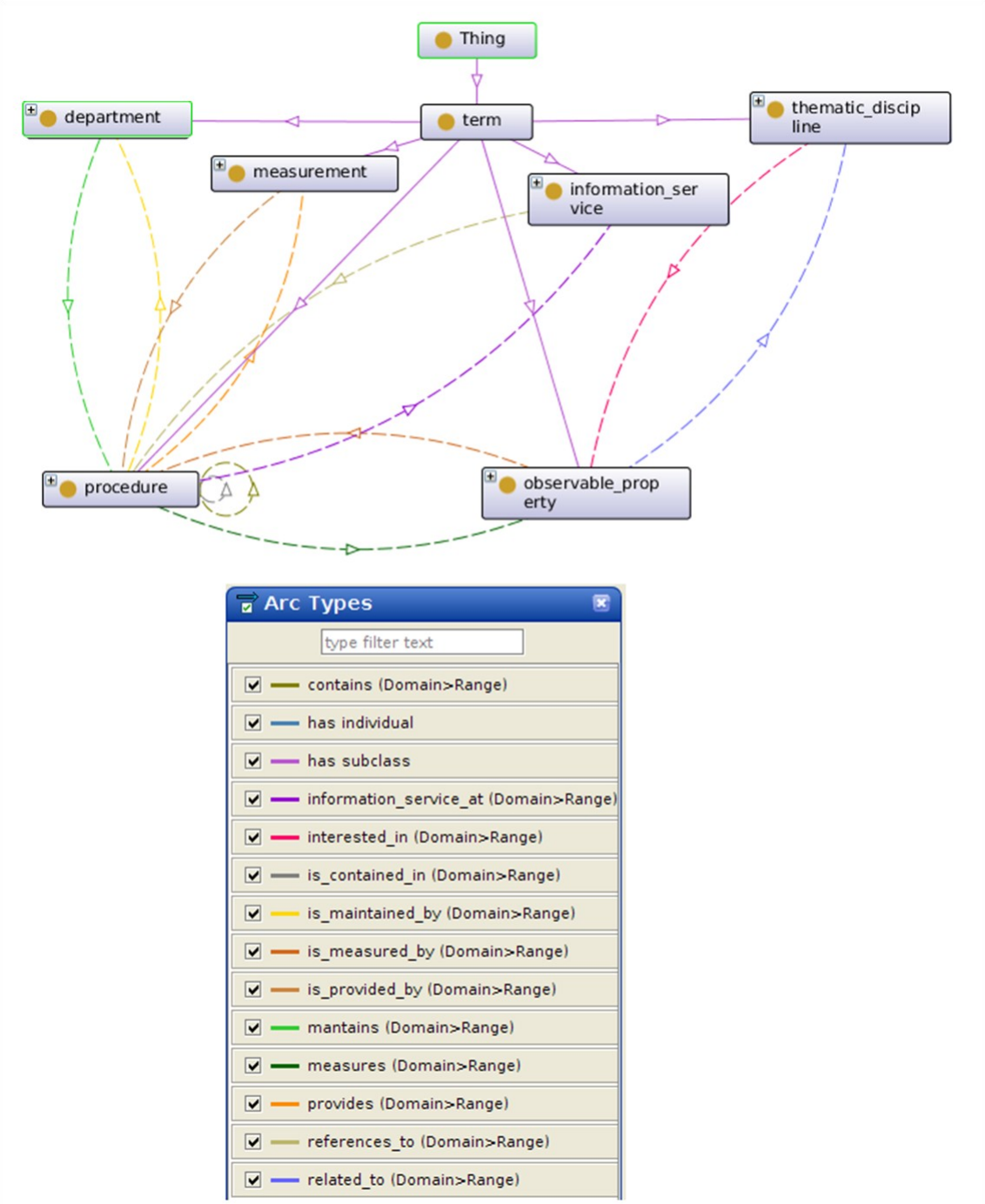


Figura 34: Model conceptual de l'ontologia

A la figura anterior es pot veure una representació de la implementació de l'ontologia amb Protégé. Tal i com s'observa hi ha una gran quantitat de relacions entre les diferents classes que fa possible l'extracció posterior d'informació.

Individuals

Per tal de poder realitzar una correcta avaluació de l'ontologia s'han creat instàncies de les diferents classes anteriorment citades de manera que sigui possible valorar com funciona l'ontologia i tenir una millor idea de la seva validesa un cop importada a la wiki. Es citen alguns exemples de cada classe per il·lustrar la gran varietat de sensors i conceptes que intervenen.

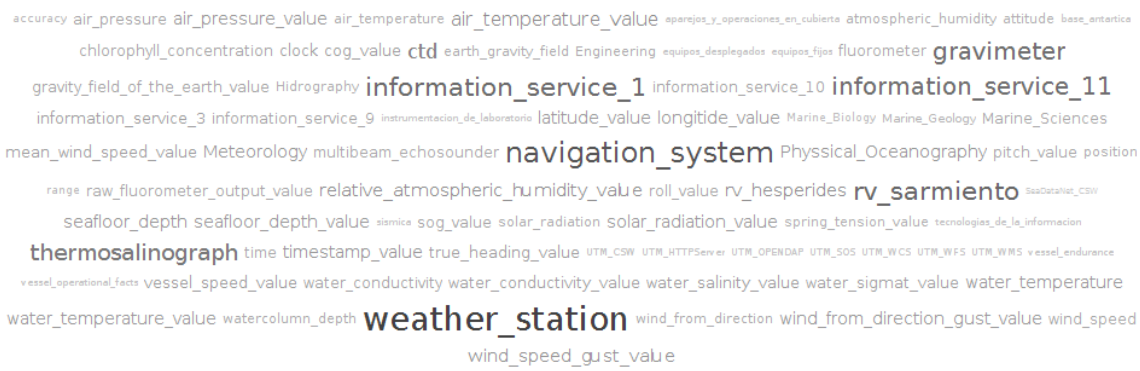


Figura 35: Individuals

Les instàncies creades són les següents:

1. **<u>Departament</u>**:

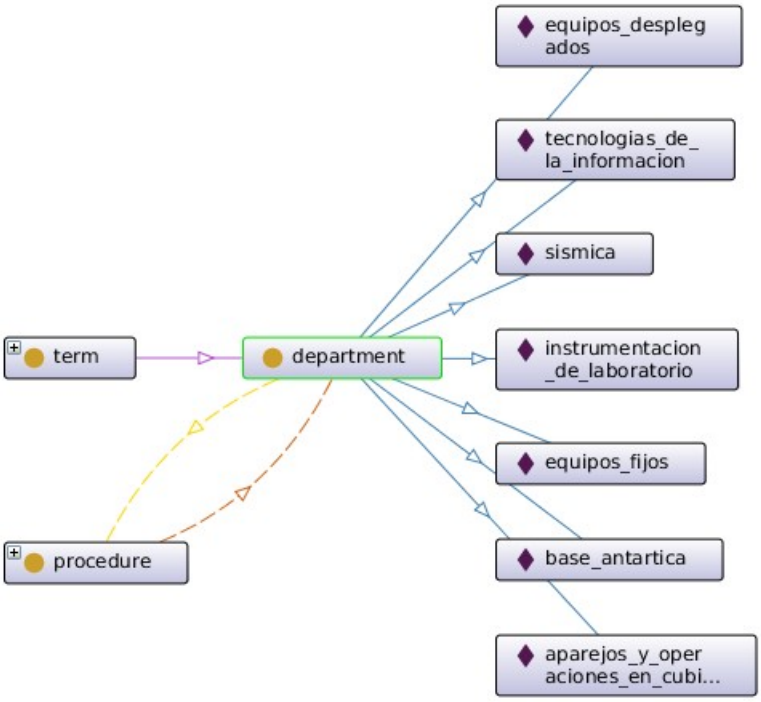


Figura 36: Instàncies de departament

Els noms que s'han utilitzat per anomenar cada departament dins de la estructura de gestió de la UTM són els següents:

- **Equipos_desplegados**: departament responsable d'aquells equips que no sempre es troben operatius i que es “despleguen” o “repleguen” en funció de la campanya. Per exemple CTD o xarxes.

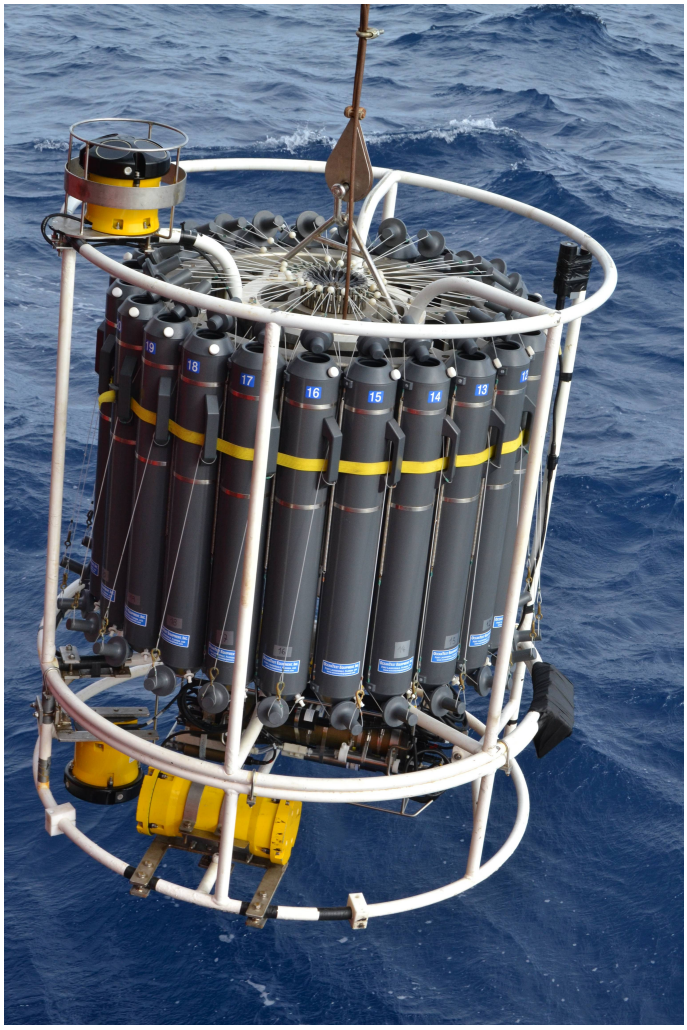


Figura 37: CTD

- **Tecnologias_de_la_información**: departament responsable de la gestió de l’adquisició i gestió de dades i, de les comunicacions de dades i veu.
- **Sísmica**: departament responsable de les campanyes de sísmica.
- **Instrumentación_de_laboratorio**: departament responsable de la gestió del laboratori i el seu instrumental.
- **Equipos_fijos**: departament responsable de tots aquells instruments que en tot moment es troben operatius als vaixells. És el cas de la sonda de profunditat.

- **Base_antartica**: departament responsable de la gestió de la base antàrtica Juan Carlos I.
- **Aparejos_y_operaciones_en_cubierta**: departament responsable de la gestió i manteniment de la part mecànica.

Tal i com es veu la divisió en departaments no segueix sempre el mateix criteri, així pel cas de la base antàrtica respon a una plataforma d'investigació en concret i per "equipos_fijos" la separació s'ha fet tenint en compte la natura dels instruments. En tot cas aquesta és la divisió en departaments que en el moment actual es troba en funcionament a la UTM.

2. **<u>Information_service</u>**:

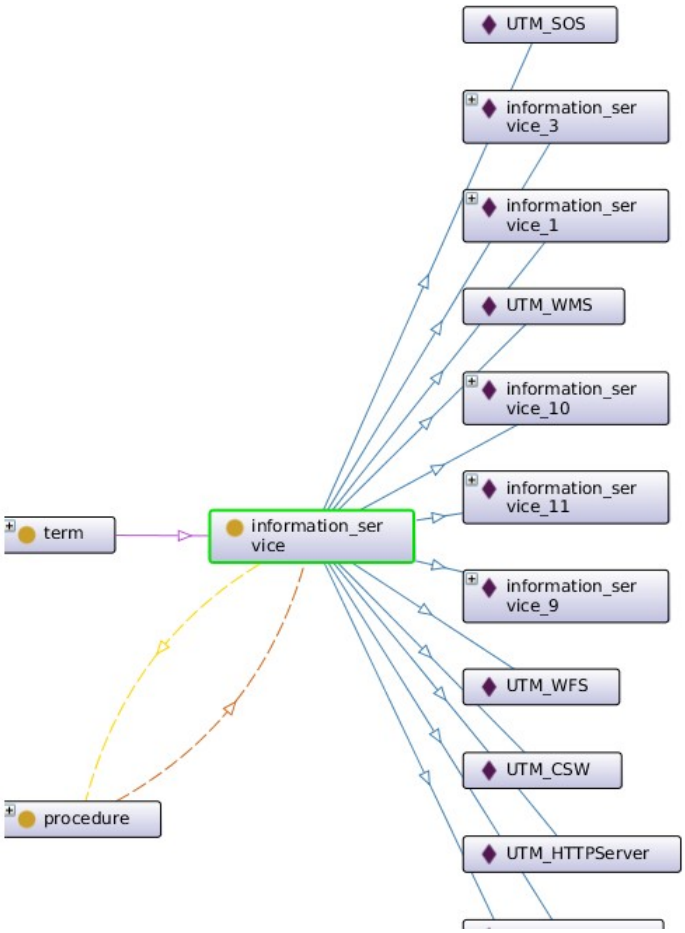


Figura 38: Information_service

S'han creat diferents serveis per mirar de representar la heterogeneïtat de serveis de dades que interven: SOS, WMS, WFS, WCS, CSW entre d'altres.

3. **Measurement**:

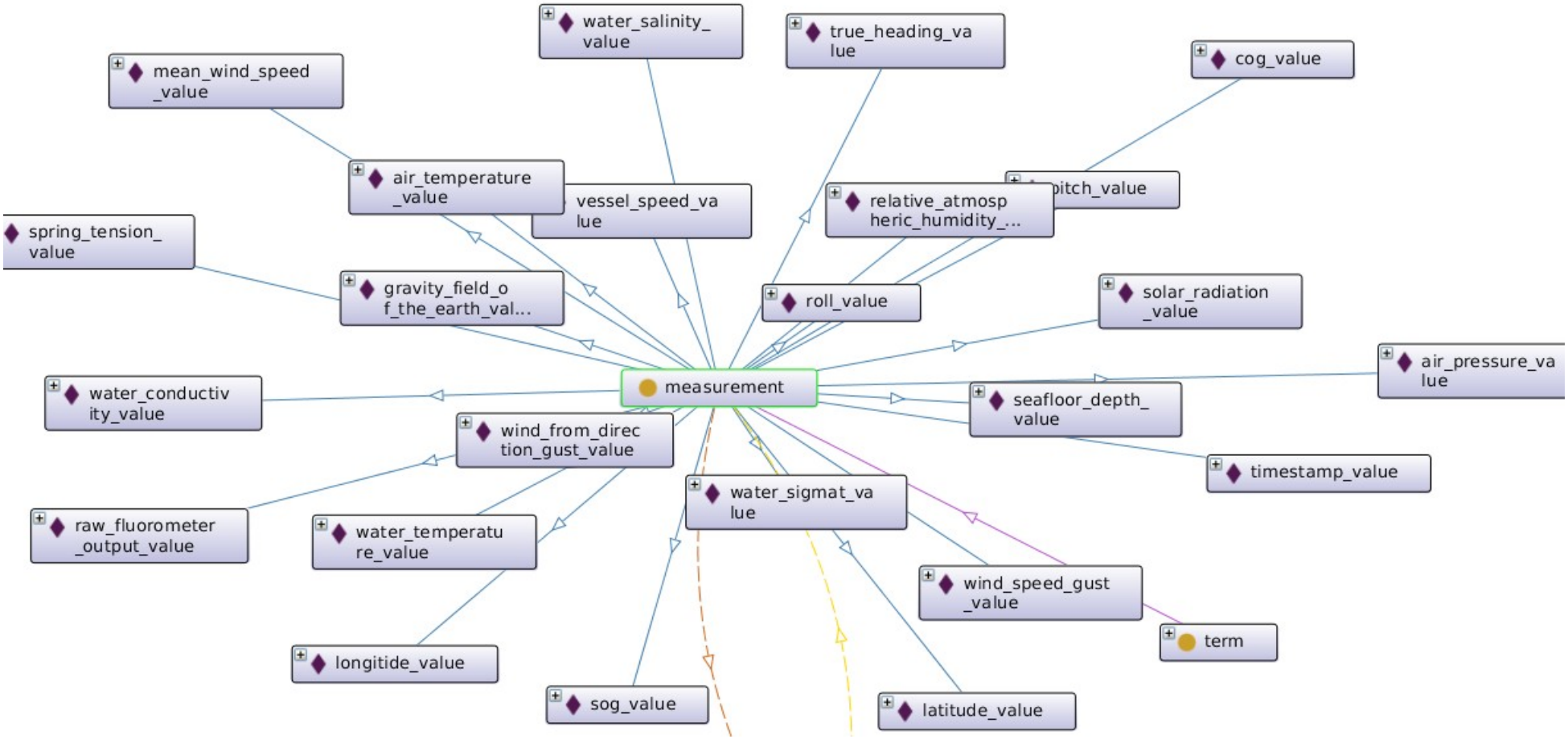


Figura 39: Measurements

A tall d'exemple s'han creat algunes de les mesures més emprades.

4. **<u>Observable_property</u>**:

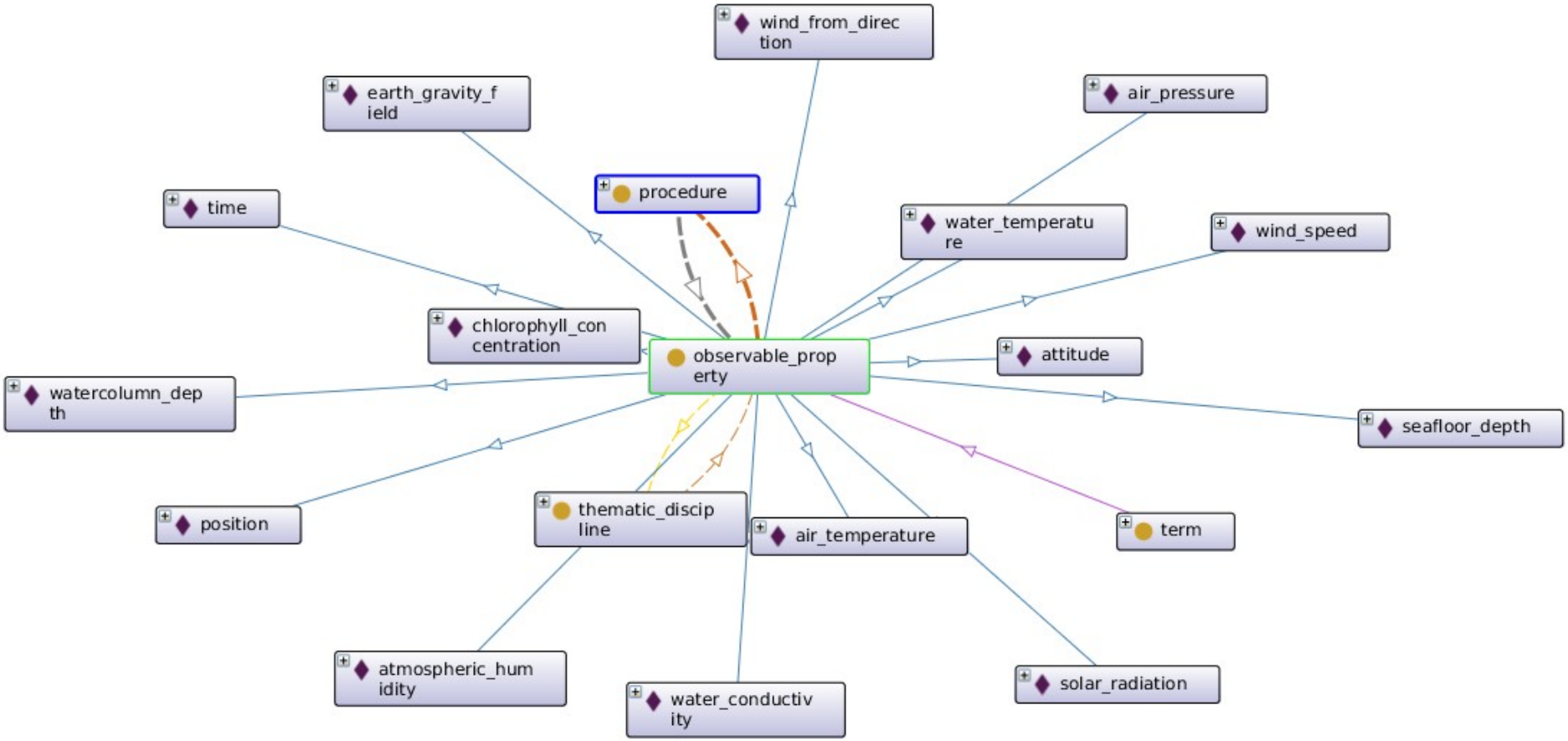


Figura 40: Observable_property

S'ha seguit el mateix criteri i s'han creat algunes de les propietats més utilitzades.

5. **Procedure**:

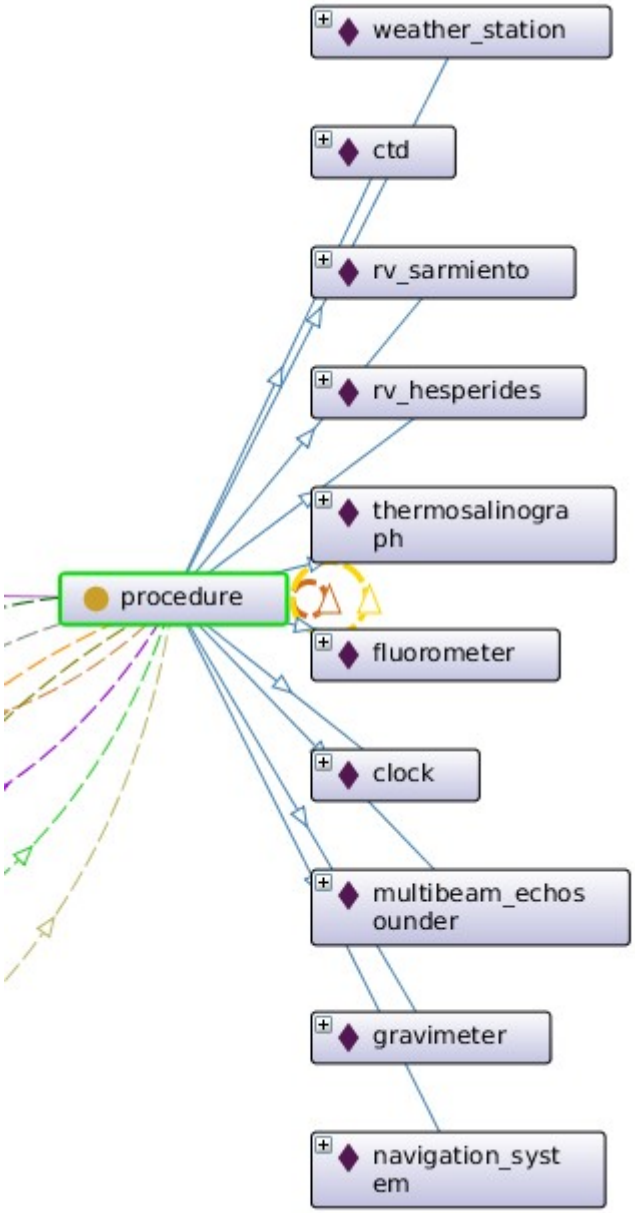


Figura 41: Procedures

S'han creat les següents instàncies, que corresponen a alguns dels instruments d'utilització més habitual en el decurs d'una campanya oceanogràfica:

- **weather_station**: estació meteorològica.

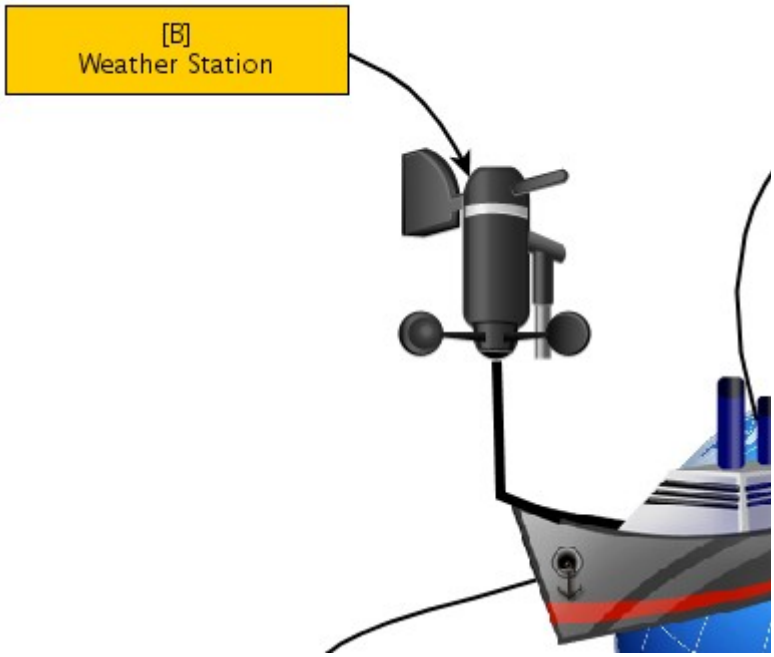


Figura 42: Weather Station

En tota campanya oceanogràfica hi ha una sèrie d'instruments que funcionen en tot moment i que proporcionen informació d'interès general o de context que també s'anomena "Underway data". Aquesta informació de propòsit general està formada per l'estació meteorològica, el termosalinògraf, el fluorímetre i el sistema de posició.

- **CTD**: instrument que mesura la temperatura, conductivitat i pressió a diferents profunditats de manera contínua (fins a 24 vegades per segon), entre d'altres paràmetres.
- **rv_sarmiento**: vaixell d'investigació oceanogràfica Sarmiento de Gamboa. Es pot considerar tot el vaixell com un sensor o una plataforma formada per diferents sub-sistemes que a la seva vegada tornen a estar formats per altres sistemes.
- **rv_hesperides**: vaixell d'investigació oceanogràfica Hespérides. S'aplica el mateix raonament anterior. Es tracta d'un sistema de sensors.
- **thermosalinograf**: instrument que de manera contínua mesura la temperatura i conductivitat de l'aigua a una certa profunditat en el cas dels vaixells que ens ocupa (entre 2 i 5 metres). A partir d'aquestes mesures també proporciona la salinitat.
- **fluorometer**: instrument que de manera contínua mesura la fluorimetria de l'aigua de mar que recull a una certa profunditat. Treballava de forma conjunta amb el termosalinògraf.
- **clock**: rellotge global a tot el sistema que permet situar cada mostra o mesura en una posició concreta. Normalment es tracta d'un servidor extern.

- **multibean_echosounder**: proporciona la profunditat que hi ha sota el vaixell en cada cas. Aquesta informació es proporciona de forma conjunta amb les dades de longitud i latitud que proporciona el sistema de navegació.
- **gravimeter:** instrument que proporciona el valor del camp gravitatori local a cada punt per on navega el vaixell.
- **navigation_system**: sistema de navegació que proporciona la longitud i latitud a mesura que el vaixell navega. Resulta de vital importància per situar cada observació en el punt geogràfic que li correspon.

6. **Thematic_discipline**:

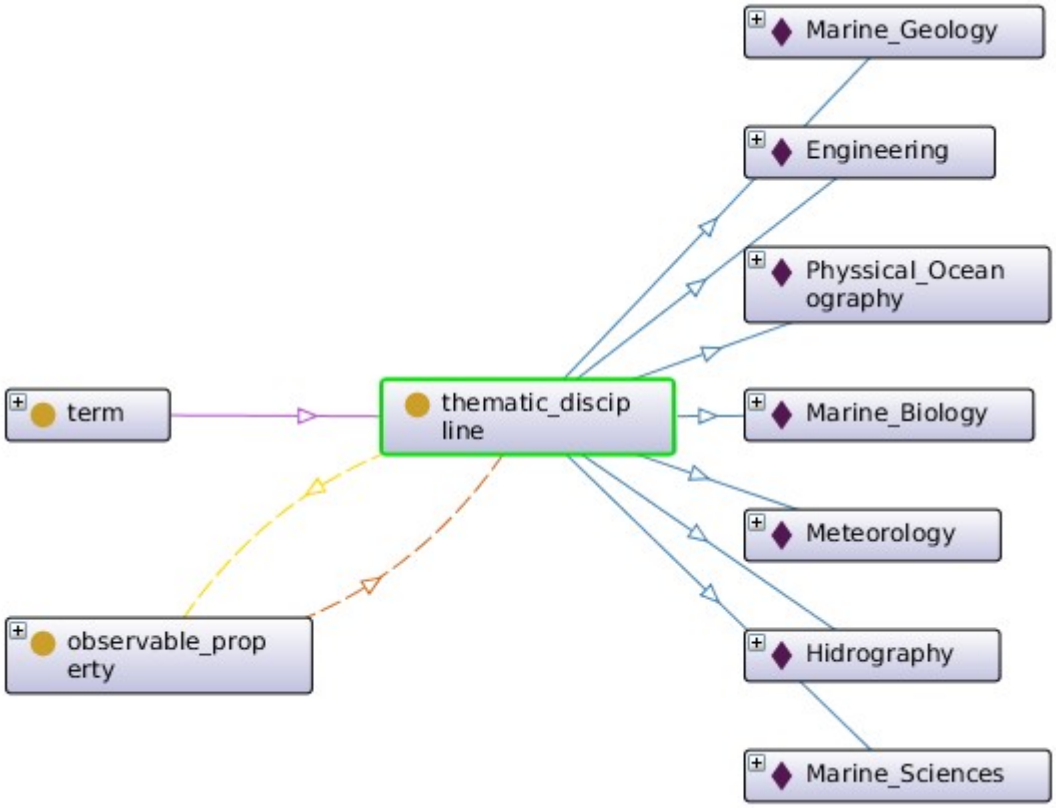


Figura 43: Instàncies de Thematic Discipline

S'han creat instàncies de les principals disciplines interessades, directa o indirectament en l'àmbit de la investigació oceanogràfica.

4. Wiki semàntica per a vaixells d'investigació oceanogràfica

Una vegada elaborada l'ontologia per a vaixells d'investigació oceanogràfica s'ha optat per utilitzar una wiki per fer que tota la informació semàntica estigui disponible per a la major quantitat possible de persones. Sense importar el seu grau de coneixement o especialització i a través d'un simple navegador web poden consultar l'ontologia creada.

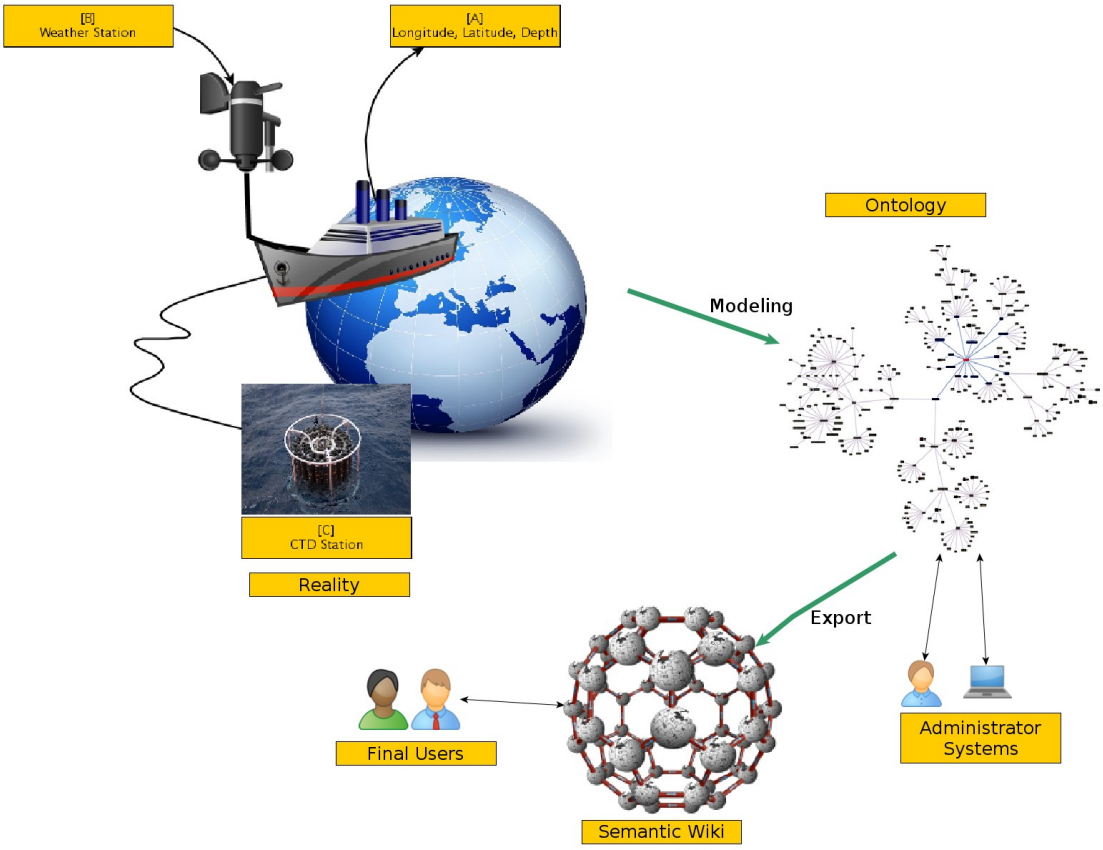


Figura 44: Realitat–Ontologia–Wiki Semàntica

A la figura anterior es pot veure com en primer terme hi ha un procés de modelització de la realitat dels vaixells d'investigació oceanogràfica on aquests són uns sistemes complexos formats per diferents sensors. Com a resultat d'aquest procés de modelització s'obté una ontologia, que és consultable per usuaris més especialitzats i per intercanviar informació amb d'altres sistemes. Finalment, com a resultat d'un procés d'exportació de l'ontologia a una wiki semàntica, tot aquest coneixement és accessible per tota la resta de persones susceptibles d'estar-hi interessades, a través d'un navegador web i d'una manera el més entenedora possible.

A l'hora d'exportar l'ontologia creada a una wiki semàntica en primer lloc s'ha procedit a avaluar les diferents implementacions que estan disponibles en el moment actual, per tal de decidir quina és la més adient per assolir els objectius del present projecte. S'han avaluat les següents wikis semàntiques:

- OntoWiki

.

- Semantic Media Wiki
- Semantic Media Wiki Plus
- Kiwi

4.2 OntoWiki

Ontowiki [33] és una aplicació de wiki semàntica de programari lliure llicenciat sota GNU General Public License Versió 2 impulsada per la Unió Europea.

Ontowiki pot fer servir tant MySQL com "Virtuoso triple store" [34]. El suport per màquines Linux és força bo i hi ha la possibilitat d'afegir un repositori i instal·lar un d'aquests dos paquets:

- Ontowiki-mysql.
- Ontowiki-virtuoso.

Per tal d'avaluar les possibilitats d'Ontowiki s'opta per instal·lar i avaluar el paquet de MySQL, ja que és un SGBD que sol estar més present a les organitzacions i pot permetre que la wiki semàntica a desenvolupar s'integri a d'altres aplicacions ja presents al sistema informàtic de marc del present treball. Val a dir que la instal·lació resulta complicada i poc documentada.

OntoWiki no obliga a fer servir una sintaxis especial per fer les anotacions RDF ja que utilitza el concepte de "mapes d'informació" de manera que a cada node li correspon una pàgina que enllaça amb els recursos amb els que està relacionada. Mitjançant els "mapes d'informació" es poden modelar els diferents elements d'una ontologia OWL de forma directa.

OntoWiki està basada en formularis i no en sintaxis, tractant així d'amagar als usuaris les complexitats dels formalismes de la representació del coneixement que emmagatzema.

OntoWiki presenta gran nombre de funcionalitats de cerca semàntica, com cerca per "text complert" per tots els camps de tipus literal i accepta l'ús de **SPARQL**.

4.3 Semantic Media Wiki

Semantic Media Wiki (SMW) també és programari lliure. **Es tracta d'una extensió de la wiki que permet emmagatzemar dades semàntiques dins de les planes de la wiki. Permet fer cerques semàntiques i mostrar els resultats d'aquestes en forma de mapes, calendaris i gràfics que es poden exportar a d'altres formats com RDF i CSV. SMW guarda les dades a MySQL i a "RDF triplestore" tal i com pot ser el cas de 4store i Virtuoso (tot i que sempre es segueix fent servir MySQL).**

Tot i que SPARQL és el **llenguatge de consulta més estès a la web semàntica SMW no implementa aquesta possibilitat ja que no es pot accedir a les dades emmagatzemades a MySQL.**

S'instal·la la versió Semantic MediaWiki 1.8.0.4. Els requeriments són els següents:

- MediaWiki [36] >= 1.17.
- **Validator MediaWiki extension >= 0.5.**
- PHP >= 5.2**.**
- **MySQL >= 4.0.14**

Primer cal instal·lar MediaWiki, la instal·lació és senzilla. Per instal·lar SMW només cal copiar els directoris de SMW a la carpeta extensions de MediaWiki.

4.4 Semantic Media Wiki Plus

Semantic MediaWiki Plus (SMW+) [37] ja s'anuncia a la seva pàgina com una wiki semàntica empresarial per compartir coneixement dintre d'equips. Es distribueix sota llicència GNU GPL v2. La versió actual és la 1.7.0.

Semantic MediaWiki Plus	
www.smwplus.net	
Status:	stable
Last release:	1.7.0 (2012/04/24)
License:	GPL
Language:	PHP
Affiliation:	DIQA Projektmanagement GmbH

Figura 45: Semantic Media Wiki + Fitxa

Enfront de les dificultats trobades tant en la instal·lació com en la importació de l'ontologia a la wiki en les anteriors wikis avaluades (Ontowiki i Semantic Media wiki), les funcionalitats i facilitat de configuració de Semantic Media Wiki Plus la fan molt atractiva. Malauradament actualment SMW+ ja no es continua desenvolupant des del 2012 per manca de finançament. Han sorgit com mínim dos projectes que tenen voluntat d'aprofitar les extensions i esforç fet en SMW+. Aquest projectes són: DataWiki [53] i OntoLTS[54].

- **DataWiki** és una solució comercial, tot i que té una versió d'avaluació. El cost de DataWiki oscil·la entre els 5.000 i 25.000 € depenent de la llicència.

- **OntoLTS** busca poder seguir utilitzant moltes de les extensions de SMW+ en versions més actuals de Semantic Media Wiki, versions posteriors a la 1.17.

La versió amb instal·lador únicament es troba disponible per a Windows (versions Windows 7 o Windows Server 2008). Es tracta de la versió: **SMW+ Community Edition-1.7.0_b125.exe**.

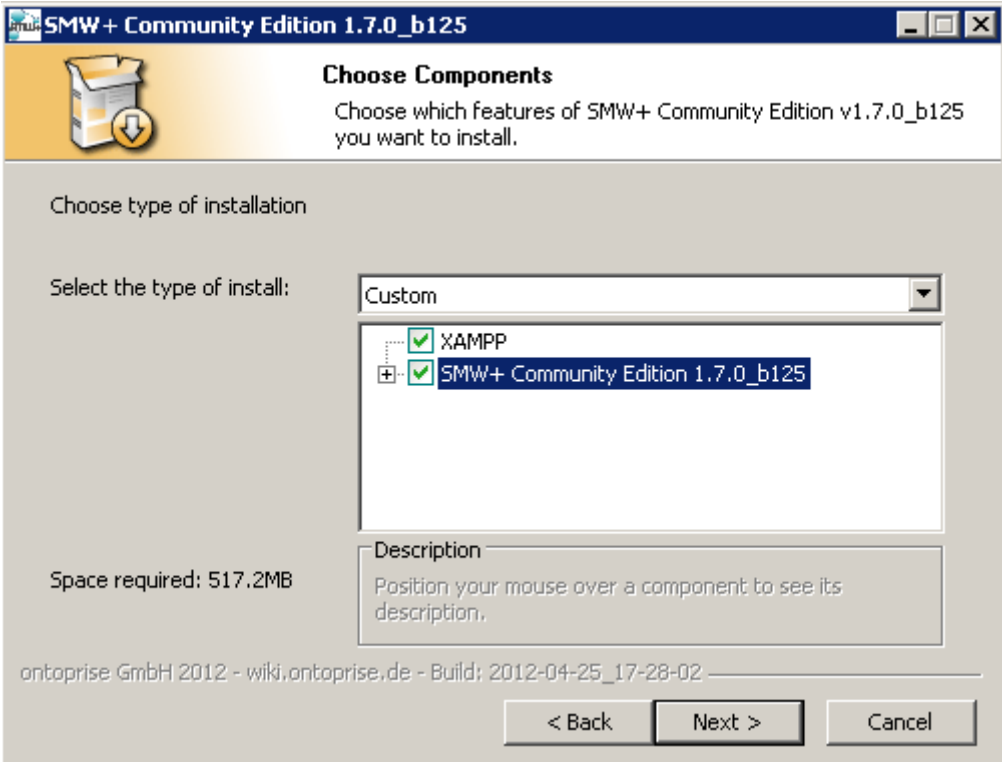


Figura 46: Instal·lador SMW+

L'instal·lador permet muntar tot l'entorn per tal que la wiki sigui totalment operativa de manera que dóna l'opció de configurar un XAMPP (paquet format pel servidor HTTP Apache, base de dades MySQL i PHP).

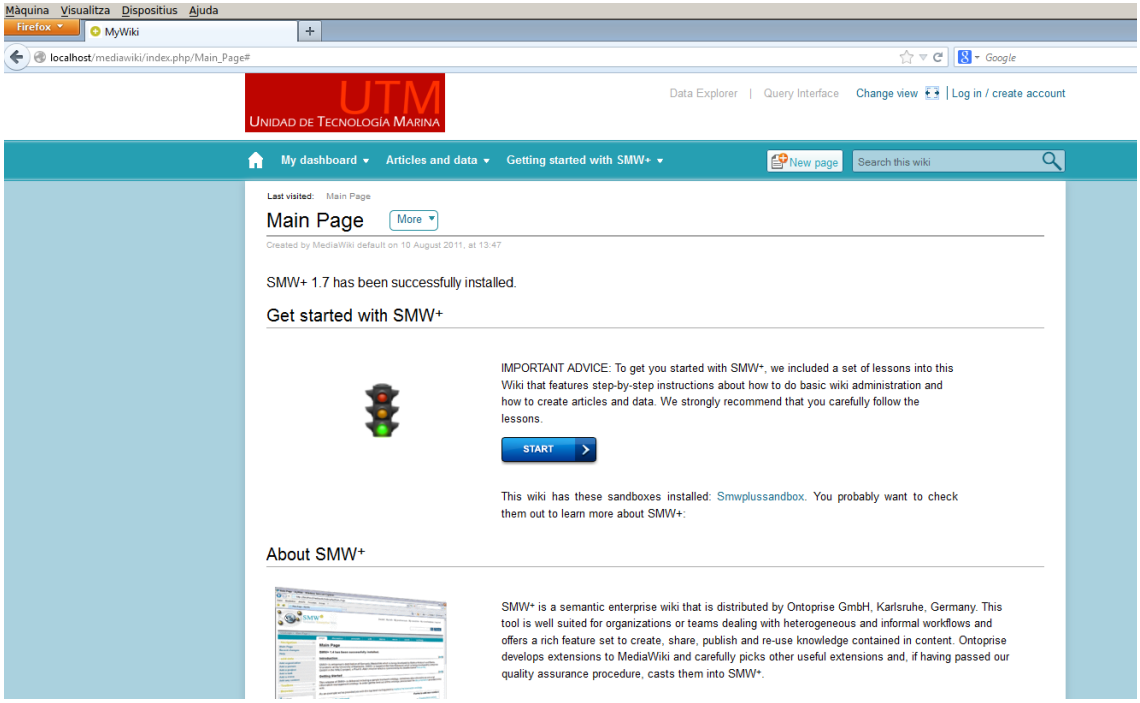


Figura 47: Semantic Media Wiki +

A la figura anterior es pot veure la pantalla de benvinguda de Semantic Media Wiki Plus. La instal·lació s'ha fet en una màquina virtual amb Windows 7, per tal de complir amb els requeriments del programari.

4.5 Kiwi

Tot i que és un projecte de Wiki Semàntica que segons sembla ha recollit opinions molt favorables en el moment de realitzar el present treball no ha estat possible descarregar-lo per a realitzar una avaluació.

4.6 Diferències i selecció de programari

OntoWiki presenta grans avantatges com el fet de poder fer servir SPARQL per a les consultes i el fet de poder importar l'ontologia a través d'una eina especifica. Per contra la instal·lació ha resultat força complicada. El paquet que hi ha per Ubuntu no ha estat possible fer-lo funcionar.

Després d'avaluar les diferents wikis amb capacitats semàntiques s'ha optat per treballar amb SMW+ donades les facilitats d'importació de l'ontologia i la facilitat d'instal·lació de tot el paquet sencer de requeriments dins d'un sol executable.

4.7 Importació de l'ontologia

A través de la eina de importació de SMW+ és trivial importar l'ontologia.

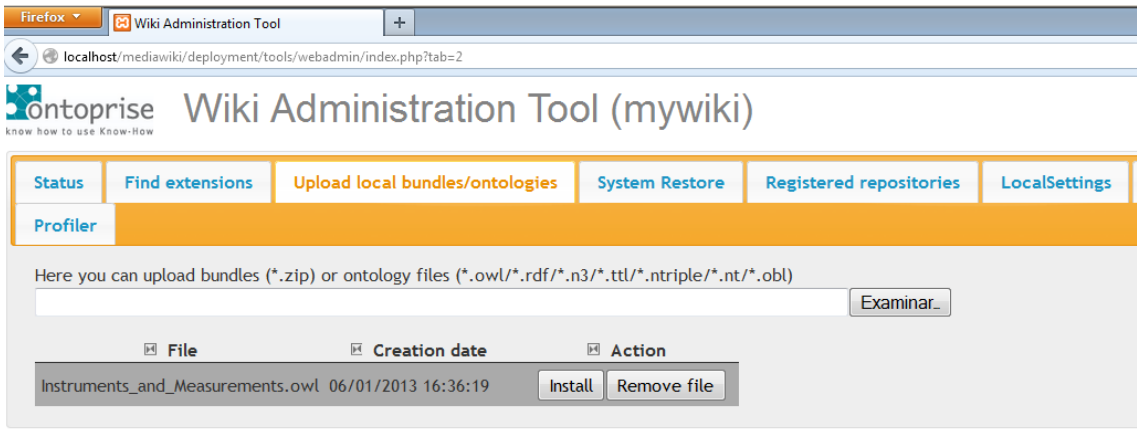


Figura 48: Importació Ontologia a SMW+

Una vegada seleccionada es procedeix amb la instal·lació de l'ontologia a SMW+. L'únic punt amb el que cal parar compte és amb la instal·lació de **TripleStore Basic** i una vegada instal·lat indicar el path al fitxer **onto2mxml.bat** per tal que al importar pugui trobar el fitxer **onto2mxml.exe.**

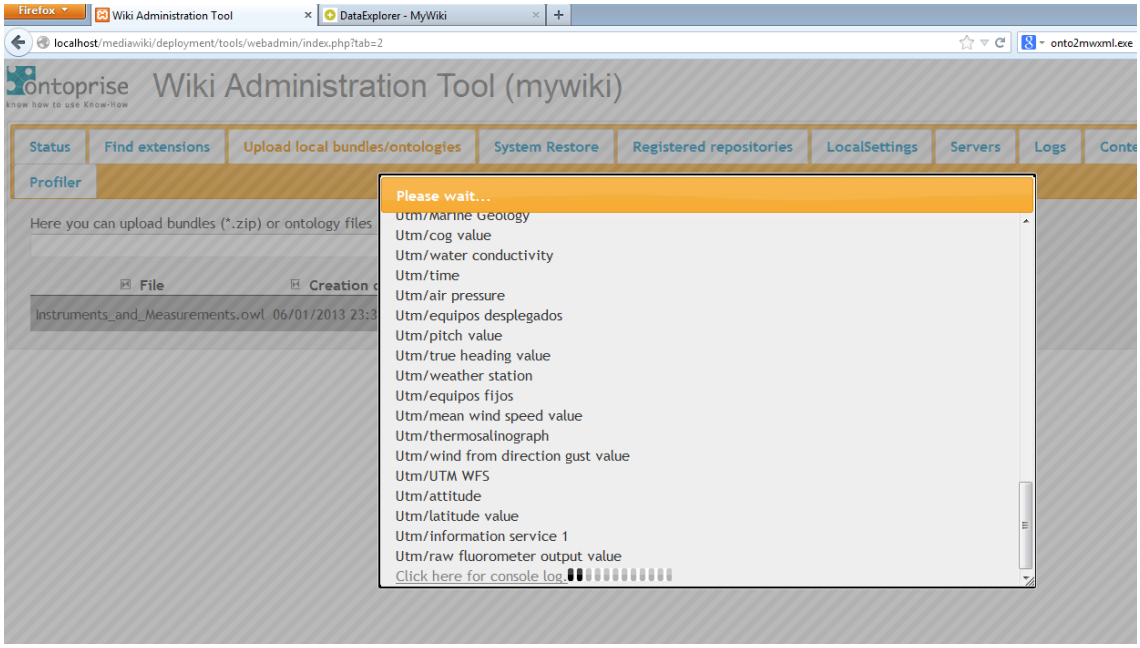


Figura 49: Procés de importació de l'ontologia des de SMW+

4.8 L'ontologia a SMW+

Una vegada s'ha importat l'ontologia a SMW+ aquesta ja és accessible des del navegador.

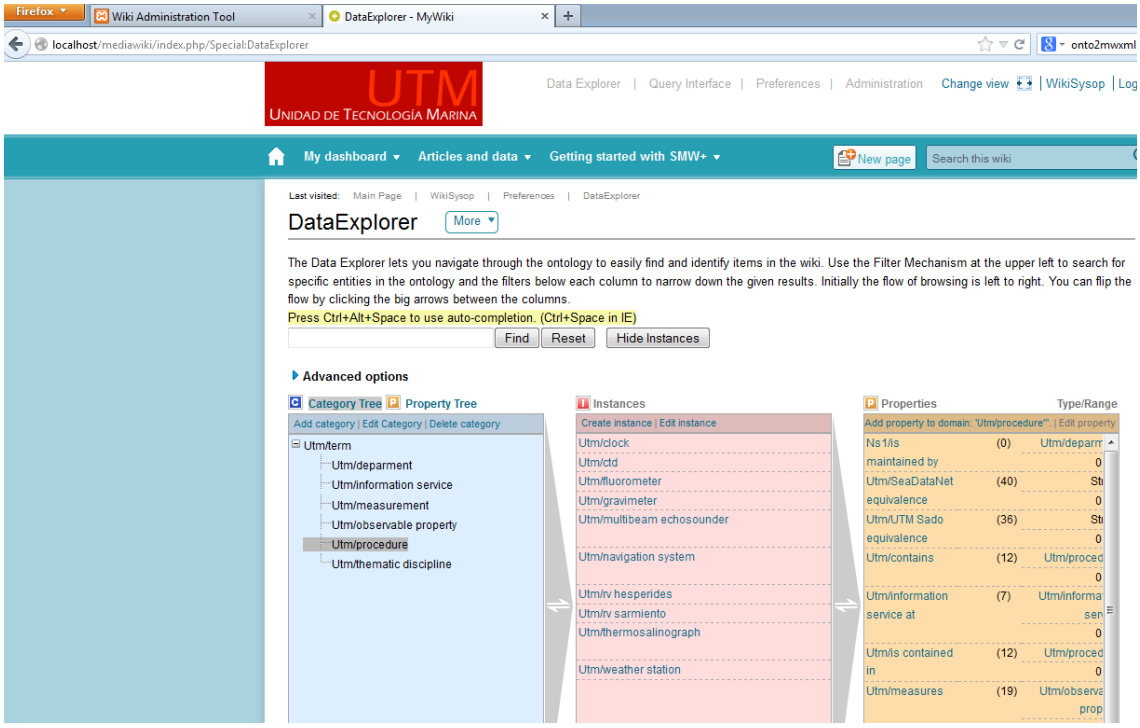


Figura 50: Ontologia a SMW+

A la pantalla inicial si se selecciona "**Data Explorer**" apareixen tres arbres de visualització. Al primer apareixen les categories que s'han creat. Es pot triar entre "**Category Tree**" i "**Property Tree**". La següent figura correspon a "**Category Tree.**"

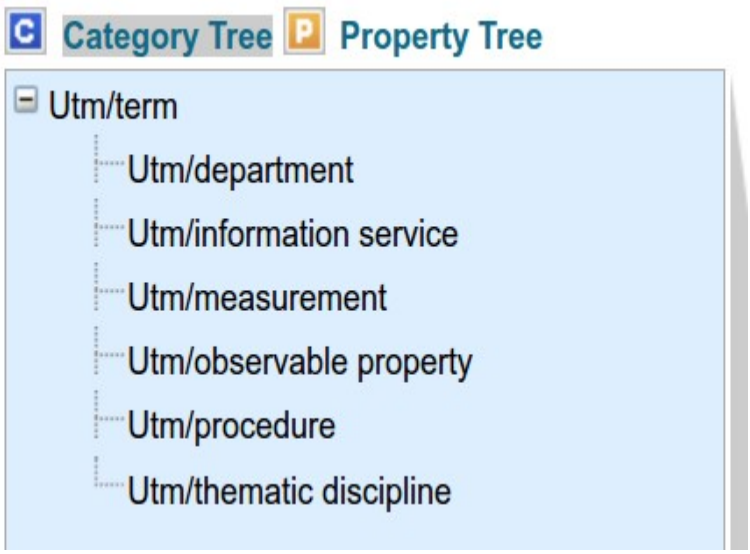


Figura 51: "Category Tree" a SMW+

A la part superior a l'igual que a la resta d'arbres sempre hi ha la possibilitat d'afegir un nou element, editar-lo i esborrar-lo. Si comparem la facilitat amb la que qualsevol usuari pot interactuar amb l'ontologia a través de la wiki en relació a les relatives dificultats de realitzar aquestes tasques directament a través de Protégé es pot veure que de cara a tenir l'ontologia actualitzada resulta de gran ajuda donar a tots els usuaris la possibilitat de realitzar aquestes tasques a través d'una interfície com la Wiki, que en molts casos ja els hi resultarà familiar. Per altra banda, no hi ha la necessitat de realitzar cap instal·lació de programari especialitzat a les seves màquines. Alguns d'aquests avantatges també es poden assolir amb WebProtege, tot i que segueix sent necessari que l'usuari es familiaritzi amb l'entorn.

Figura 52: Funcionament SMW+

Tal i com es pot veure a la figura anterior els usuaris de l'ontologia a través de SMW+ poden seleccionar una categoria, a continuació seleccionar una instància i veure les propietats d'aquesta instància.

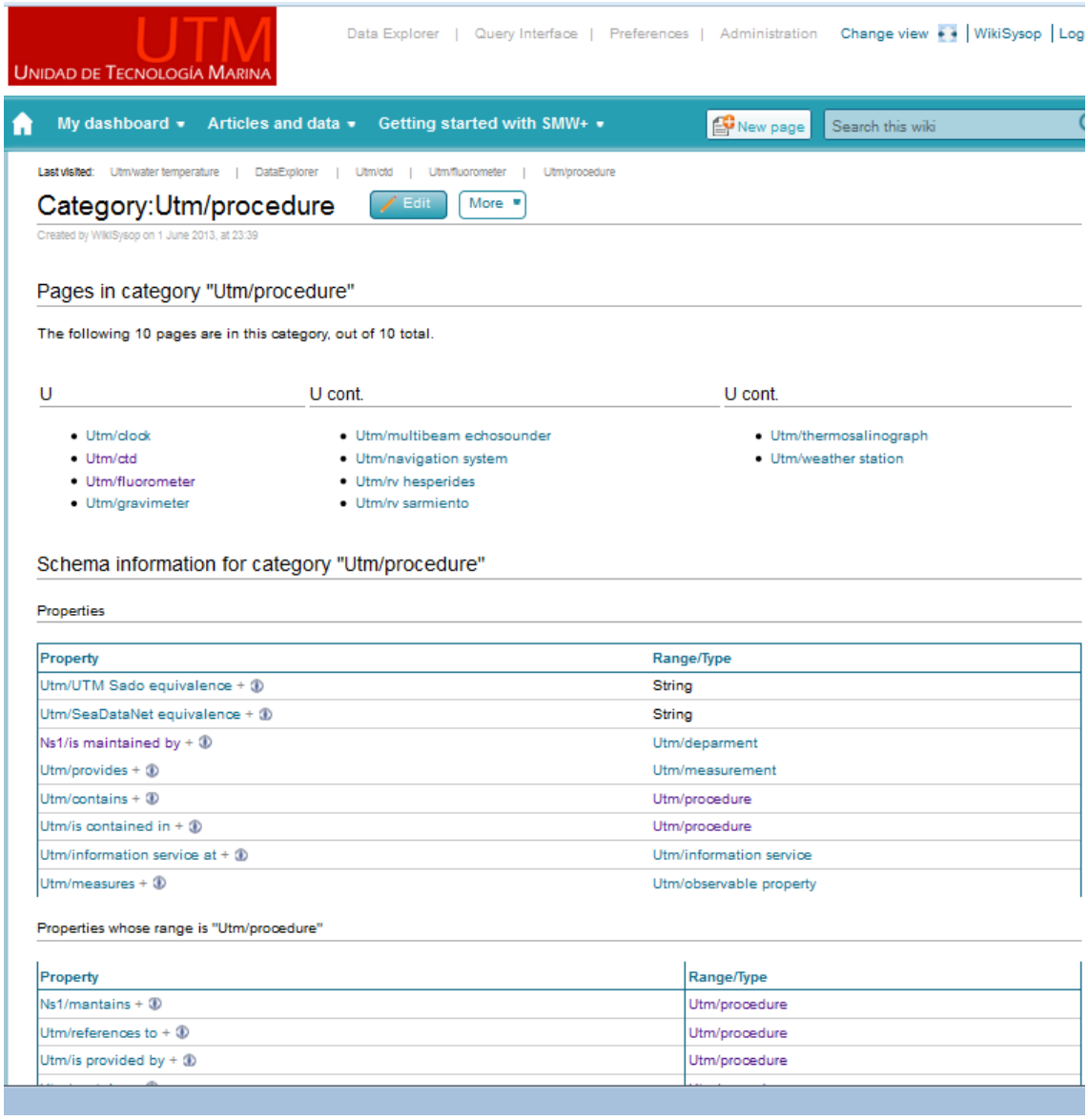


Figura 53: Procedure a SMW+

Si es selecciona una classe com Procedure es pot veure el detall de la mateixa (figura anterior, detall de "**Procedure**").

Hi ha la possibilitat de realitzar consultes de manera molt fàcil i intuïtiva a través del **"Query Interface"**

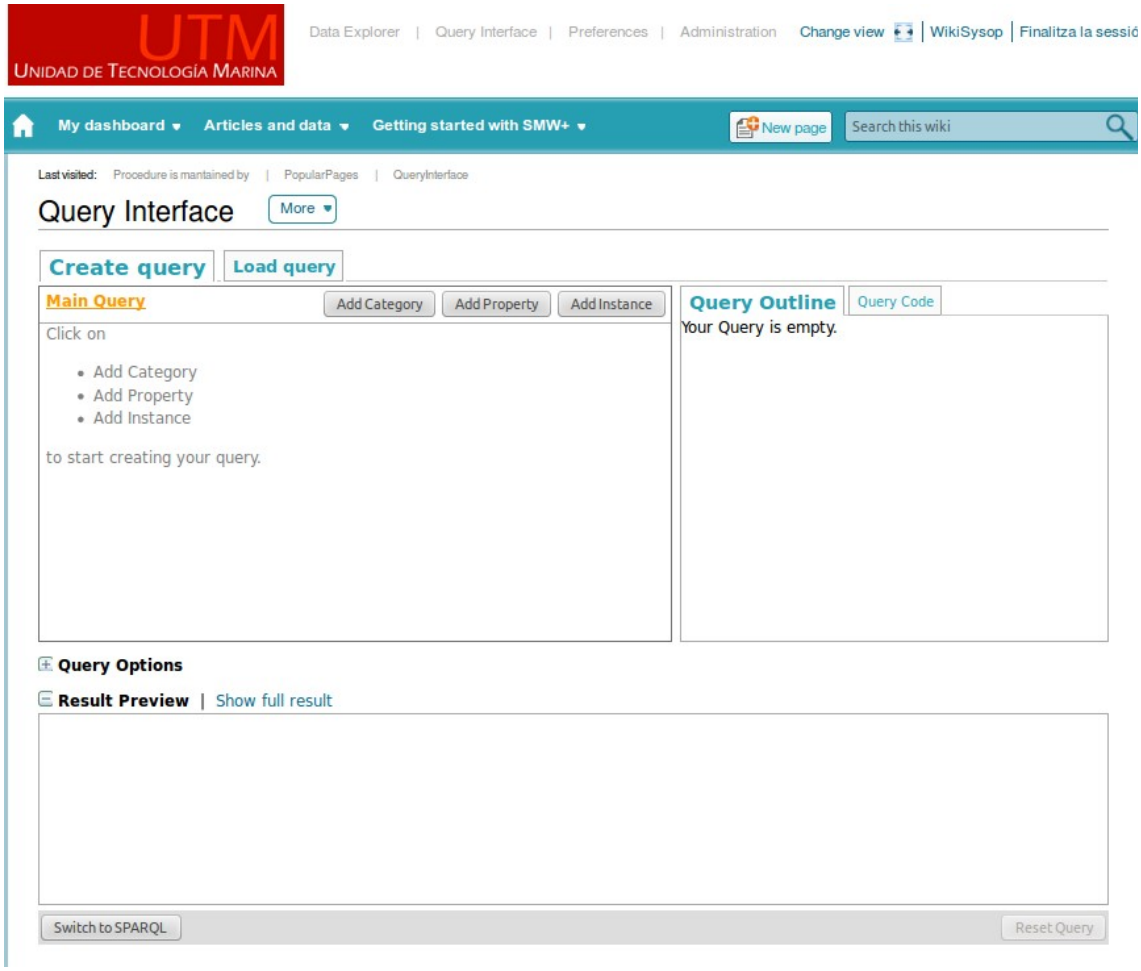


Figura 54: Query Interface

Una vegada realitzada una consulta amb l'editor aquesta es pot fer sevir per crear una pàgina amb el resultat de la cerca o emmagatzemar-la per fer-la servir posteriorment. Les consultes es poden fer directament en SPARQL que és més potent o fer-la amb l'editor i posteriorment convertir la sentència a SPARQL.

Sentència produïda amb "Query Interface":

```
{{#ask: [[Category:Utm/procedure]]

| ?Ns1/is maintained by

| format=table

| source=wiki

| merge=false|}}
```

UTM
UNIDAD DE TECNOLOGÍA MARINA

Data Explorer | Query Interface |

My dashboard ▾ Articles and data ▾ Getting started with SM

Last visited: Main Page | Administration | Halomenu | Procedures contained

Procedures contained In

Modifica More

Created by WikiSysop on 4 juny 2013, at 07:09

Procedure

	Utm/is contained in
Utm/Salinometer	
Utm/clock	Utm/rv sarmiento Utm/rv hesperides
Utm/ctd	
Utm/fluorometer	Utm/thermosalinograph
Utm/gravimeter	Utm/rv hesperides Utm/rv sarmiento
Utm/multibeam echosounder	Utm/rv sarmiento
Utm/navigation system	Utm/rv sarmiento Utm/rv hesperides
Utm/rv hesperides	
Utm/rv sarmiento	
Utm/thermosalinograph	Utm/rv sarmiento Utm/rv hesperides
Utm/weather station	Utm/rv hesperides Utm/rv sarmiento

Figura 55: Pàgina amb Consulta

5. Conclusions i línies de futur

Amb el present treball es mostra com es poden aprofitar les possibilitats que ofereix la web semàntica dins del marc dels vaixells d'investigació oceanogràfica. Tal i com s'ha indicat anteriorment, es tracta d'un entorn complex i heterogeni, tant per la natura dels sensors que hi intervenen, que pertanyen a disciplines molt diferents (biologia, geologia, física), com per la pròpia organització i gestió dels sensors dins de les plataformes, vaixells. Per tot això fer servir una ontologia per modelitzar i poder treballar amb tota aquesta informació és de gran utilitat.

Per altra banda també cal remarcar que tot i els grans avantatges de la web semàntica i de l'ús d'ontologies, el desconeixement sobre les mateixes és molt gran i sempre són necessàries llargues introduccions de cara a explicar les bondats d'aquesta tecnologia. Una vegada fet aquest esforç inicial la resta de persones implicades en la recollida d'informació i de requisits entén molt més el que es pretén i quins són els avantatges. Aquests avantatges es fan evidents al traslladar l'ontologia a una Wiki Semàntica, ja que permet amagar la complexitat de la descripció de l'ontologia en fitxers RDF o OWL sota pàgines web, que mostren la informació de forma ordenada i molt més amigable de cara a la majoria d'usuaris, de manera que la lògica del model segueix estant en un format estàndard, en OWL en aquest cas, que permet descriure de manera sistemàtica i ordenada l'ontologia i la visualització en HTML, llenguatge orientat a la representació. Salvant les distàncies podríem dir que hi ha una certa semblança amb el model vista controlador, MVC, en el que el controlador és l'ontologia i té la lògica de negoci i a través de la wiki i les seves pàgines web hi han les vistes.

Pel que fa a les eines de disseny d'ontologies, Protégé s'ha mostrat com un programari extremadament potent, amb moltes extensions i lliure. Totes aquestes característiques fan que sigui molt còmode de treballar-hi. Protégé compta amb una comunitat d'usuaris prou extensa que fa es pugui cercar informació a la web, de cara a solucionar problemes i dubtes i la salut del projecte en el que apareixen noves versions del producte de forma regular és molt bona. Per contra en el camp de les Wikis Semàntiques, tot i que la salut del projecte Semantic Media Wiki és també bona, cal lamentar la desaparició del SMW+ donada la gran qualitat i nombre d'extensions que havia assolit en el passat. Per altra banda, el projecte Kiwi de Wikis semàntiques, que prometia bones funcionalitats i desenvolupat amb diners públics de la Unió Europea, no permet descarregar el seu producte i no ha estat possible avaluar-lo.

En conjunt podem dir que la web semàntica és la solució adient per treballar en entorns complexos on cada vegada hi ha més informació i més variada, tal i com és el cas de la investigació oceanogràfica on molts esforços van encaminats de cara a treballar amb aquesta complexitat d'una manera ordenada i estandarditzada, tal i com seria el cas de SWE, que permet descriure perfectament els sensors i les observacions, i de la "web de les coses" en general. Per sobre de tot això es fa necessària una capa més que

permeti navegar-hi i que d'alguna manera actuï com a lligam·entre tots els elements.

En conjunt podem dir que la web semàntica és una solució ideal per al marc del present treball i que en general la seva presència cada vegada estarà més estesa per donar resposta a reptes de cerca d'informació i el seu emmagatzematge.

5.1 Línies de futur

S'apunta que podria ser de gran utilitat el desenvolupament d'una ontologia per la gestió de totes les infraestructures gestionades pel marc comú de col·laboració Eurofleets [55]. Eurofleets promou l'intercanvi de vaixells, equipament i tècnics entre els seus membres. Seria de gran utilitat per tots els organismes que hi formen part, el desenvolupament d'una ontologia que permetés modelitzar tota aquesta estructura i que fes possible accedir a aquesta informació d'una manera ordenada i sistemàtica. Tot i que es tractaria d'una tasca complexe i trans-nacional que involucraria diferents organismes, els avantatges derivats d'aquest esforç reportarien grans beneficis en termes de gestió i aprofitament dels recursos disponibles.

6. Glossari

- Campanya: és el període de temps de vaixell assignat a un investigador per a realitzar la recollida de dades i mostres. A cada campanya hi correspon un investigador principal, responsable de la mateixa, una data de inici i una data de final.
- Manteniment: per tal que tant els vaixells d'investigació com els propis instruments que hi han embarcats operin correctament a banda dels períodes de campanya en els que el vaixell es troba operatiu recollint informació i mostres, hi ha uns altres períodes de temps que resulten claus per el normal desenvolupament de les campanyes oceanogràfiques, són els períodes de manteniment, també anomenats PIP o varada. Durant aquests períodes es realitzen els diferents manteniments dels sistemes embarcats, ja siguin manteniments evolutius (millores) o correctius (solució de problemes detectats). Pot ser que el manteniment es pugui realitzar al propi vaixell o calgui enviar l'instrument a terra o al fabricant. Seria per exemple el cas d'un calibratge a realitzar pel fabricant.
- Varada: de forma periòdica, per exemple en el cas del BIO Sarmiento de Gamboa es cada dos anys, cal treure tot el vaixell de l'aigua a les drassanes i realitzar un manteniment de tot el casc del vaixell.
- Investigador principal: també anomenat “Cap científic” de cara a la realització d'una campanya d'investigació oceanogràfica la persona que sol·licita el temps de vaixell i que una vegada en campanya és la responsable d’indicar quines maniobres de recollida de dades vol realitzar, demanar quins sensors seran necessaris (amb antelació a l'inici de la campanya) i quin és el marc geogràfic de treball on es vol operar.
- Cap tècnic: és l'encarregat de coordinar les necessitats tècniques de la campanya i que actua com a interlocutor entre el cap científic i la resta de tècnics embarcats.
- Sensor: com a sensor o instrument entenem tot aparell que proporciona dades, observacions o mostres. Cada sensor té tota una sèrie de característiques molt variades i complexes, des del pes i mida fins els rangs d'operació, si calen altres sensors operant de forma simultània o per contra és incompatible el seu funcionament al mateix moment. Depenent del sensor caldrà un tipus de tècnic especialitzat o un altre.

7. Bibliografia

[1] Marta Sabou, “Building Web Service Ontologies”, SIKS Dissertation Series No. 2006-4, VRIJE UNIVERSITEIT 2006.

[2] Harry Halpin, Valentina Presutti , “An Ontology of Resources for Linked Data”, LDOW 2009, April 20–24, 2009, Madrid, Spain. ACM 978-1-60558-487-4/09/04 .

[3] Neiswender, C., Isenor, A., Montgomery, E., Bermudez, L., Miller, S.P. 2011. "Vocabularies: Dictionaries, Ontologies, and More". In The MMI Guides: Navigating the World of Marine Metadata. http://marinemetadata.org/guides/vocabs. Accessed March 10, 2013. “http://marinemetadata.org/print/book/export/html/2092 ”

[4] Alex Hall, Mike Lang Jr., “Ontology Integration, SPARQL Federation and SWRL Rules ”, , May 18, 2009.

[5] Cassia Trojahn, Paulo Quaresma, Renata Vieira, “A Framework for Multilingual Ontology Mapping ”

[6] Natalya F. Noy and Deborah L. McGuinness, “Ontology Development 101: A Guide to Creating Your First Ontology”, Stanford University, Stanford, CA, 94305.

[7] Luis Bermudez, John Graybeal, Robert Arko, “A Marine Platforms Ontology: Experiences and Lessons ”.

[8] Autonomous Ocean Sampling Network (AOSN), http://www.mbari.org/aosn .

[9] Bechhofer, S., Harmelen, F.v., Hendler, J., Horrocks, I., McGuinness, D.L., Patel-Schneider, P., Stein, L.A.: Owl Web Ontology Language Referenctravertede, W3C Recommendation, (2004) .

[10] Prótége, http://Protégé.stanford.edu/

[11] World Wide Web, http://es.wikipedia.org/wiki/World_Wide_Web

[12] Semantic Web, http://en.wikipedia.org/wiki/Semantic_Web

[13] "XML and Semantic Web W3C Standards Timeline" (http://www.dblab.ntua.gr/~bikakis/XML%20and%20Semantic%20Web %20W3C20Standards%20Timeline-History.pdf) . 2012-02-04.

[14] "W3C Semantic Web Activity" (http://www.w3.org/2001/sw/) . World Wide Web Consortium (W3C). November 7, 2011. http://www.w3.org/2001/sw/. Retrieved November 26, 2011.

[15] Extensible Markup Language, http://en.wikipedia.org/wiki/XML

[16] http://www.opengeospatial.org/standards/sensorml, Linked Sensor Data, Harshal Patni, Cory Henson, Amit P. Sheth Ohio Center of Excellence in Knowledge enabled Computing (Kno.e.sis) Wright State University, Dayton.

[18] Open Geospatial Consortium, OGC, http://www.opengeospatial.org/

[19] Sensor Web Enablement, SWE. http://www.opengeospatial.org/projects/groups/sensorwebdwg

[20] SensorML, http://www.opengeospatial.org/standards/sensorml

[21] O&M, http://www.opengeospatial.org/standards/om

[22] Protégé, http://Protégé.stanford.edu/

[23] Climate and Forecast, http://cf-pcmdi.llnl.gov/

[24] SeaDataNet, http://www.seadatanet.org/

[25] Linked Sensor Data, Oscar Corcho, Jean-Paul Calbimonte, Raúl García-Castro and Freddy Priyatna, 4th Future Internet Symposium FIS 2011 Vienna, Austria.

[26] Linked Data, http://linkeddata.org/

[27] Standard Generalized Markup Language, http://es.wikipedia.org/wiki/SGML

[28] JavaScript Object Notation, http://es.wikipedia.org/wiki/JSON

[29] YAML Ain't Another Markup Language, http://es.wikipedia.org/wiki/YAML

[30] Resource Description Framework, http://www.w3.org/RDF/

[31] OWL (Web Ontology Language), http://www.w3.org/2004/OWL/

[32] Wiki semàntica, http://es.wikipedia.org/wiki/Wiki_sem%C3%A1ntica

[33] OntoWiki, http://ontowiki.net/Projects/OntoWiki

[34] Virtuoso Universal Server, http://virtuoso.openlinksw.com/

[35] Semantic Media Wiki, http://semantic-mediawiki.org/

[36] MediaWiki, http://www.mediawiki.org/wiki/Download

[37] Semantic MediaWiki Plus, http://semantic-mediawiki.org/wiki/Semantic_MediaWiki_Plus

[38] OpeNDAP, http://en.wikipedia.org/wiki/OPeNDAP

[39] Web Map Service, http://www.opengeospatial.org/standards/wms

[40] Web Feature Service, http://www.opengeospatial.org/standards/wfs

[41] Sensor Observation Service, http://www.opengeospatial.org/standards/sos

[42] Catalogue Service, http://www.opengeospatial.org/standards/cat

[43] Web Coverage Service, http://www.opengeospatial.org/standards/wcs

[44] SeaDataNet Metadata formats , http://www.seadatanet.org/Standards-Software/Metadata-formats

[45] CF Metadata, http://cf-pcmdi.llnl.gov/

[46] UoM Ontology Standard,http://ontolog.cim3.net/cgi-bin/wiki.pl?UoM_Ontology_Standard

[47] “Semantically-Enabled Sensor Plug & Play for the Sensor Web”, Arne Brring, Patrick Mau, Krzysztof Janowicz, Daniel Nust and Christian Malewski; Sensors 2011, 11, 7568-7605; doi:10.3390/s110807568.

[48] Vocabulari SeaDataNet, http://seadatanet.maris2.nl/v_bodc_vocab/welcome.aspx

[49] “Semantic Interoperability and SeaDataNet”, Roy Lowry British Oceanographic Data Centre, ICAN Workshop, Copenhagen, 07-08 July 2008

[50] NeOn Toolkit, http://neon-toolkit.org/wiki/[Main_Page

[51] “Climate and Forecast”, http://cf-pcmdi.llnl.gov/

[52] IOOS, http://sdf.ndbc.noaa.gov/sos

[53] DataWiki, http://diqa-pm.com/en/DataWiki

[54] OntoLTS, http://semantic-mediawiki.org/wiki/OntoLTS

[55] Eurofleets, http://www.eurofleets.eu/np4/home.html