
Etat de l'art sur les ontologies de capteurs pour une intégration intelligente des données

Rimel Bendadouche*, Catherine Roussey*, Gil de Sousa*, Jean-Pierre Chanet*, Kun Mean Hou**

* *Irstea, UR TSCF Technologies et systèmes d'information pour les agrosystèmes
24 avenue des Landais, BP 50085, 63172 Aubière cedex, France
{rimel.bendadouche, catherine.roussey, gil.de-sousa, jean-pierre.chanet}@irstea.fr*

** *Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes
(LIMOS)
Complexe scientifique des Cézeaux, 63173 AUBIERE cedex, FRANCE
{kun-mean.hou}@isima.fr*

RÉSUMÉ. Ce document présente un état de l'art sur les ontologies de capteurs développées dans différents domaines et applications. Dans le contexte de nos travaux, nous nous intéressons aux Réseaux de Capteurs Sans Fil (RCSF) pour la surveillance des phénomènes environnementaux. Nous avons étudié les ontologies de capteurs afin de mettre en place une méthode d'intégration de données issues des RCSF en différenciant les politiques d'acquisition et de transmission des données. L'ontologie permet ainsi de définir les concepts nécessaires à la description de ces politiques. De plus, elle devra être aussi une composante d'un système d'aide à la décision capable de piloter le RCSF pour optimiser la transmission des données. Notre état de l'art fait une analyse critique des différentes ontologies existantes. Cette analyse porte sur les thèmes et sur les différents usages des ontologies dans les RCSF. Ainsi, nous mettrons en évidence qu'il n'existe pas à l'heure actuelle d'ontologie capable de différencier les données collectées et celles transmises au système d'information.

ABSTRACT. This paper is a state of the art of sensor ontologies developed in different fields and applications. In the context of our work, we are interested in Wireless Sensor Networks (WSN) for monitoring environmental phenomena. We have studied sensor ontologies to develop a WSN data integration method which differentiates the acquisition and transmission data policies. The ontology allows us to define the needed concepts to describe these policies. Furthermore, it should also be a component of a decision support system which pilots the WSN to optimize data transmission. Our state of the art is a critical analysis of existing sensor ontologies. This analysis focuses on the topics and on the various uses of ontologies in WSN. Thus, we will highlight that currently there is no ontology able to differentiate data collected from those transmitted to the information system.

MOTS-CLÉS : RCSF, Intégration de données, Contexte, Ontologies, Ontologies de capteurs.

KEYWORDS: WSN, Data integration, Context, Ontologies, Sensor ontologies.

1. Introduction

Au cours des dernières décennies et grâce à l'avancée des systèmes embarqués et des technologies sans fil, les Réseaux de Capteurs Sans Fil (RCSF), ou « Wireless Sensor Network (WSN) », sont de plus en plus utilisés dans de nombreux domaines (médical, militaire, environnemental, etc.). Dans le contexte de nos travaux, nous nous intéressons aux RCSF pour la surveillance des phénomènes environnementaux. Un RCSF est généralement constitué de plusieurs nœuds, leur nombre allant de quelques dizaines à des centaines, qui stockent, traitent et acheminent des données mesurées en utilisant les technologies sans fil (Akyildiz et al., 2002). Par nature, ces nœuds se caractérisent par des ressources limitées en énergie, puissance de calcul et capacité de stockage. Dans ce type de réseaux, l'énergie est la contrainte principale car elle conditionne les services rendus par chaque nœud. Sachant que la transmission est le service qui consomme le plus d'énergie, un nœud de RCSF doit limiter la quantité de données transmises autant que possible afin d'augmenter sa durée de vie. C'est pourquoi la politique d'acquisition des données peut être très différente de la politique de transmission. Ainsi, deux nœuds peuvent être paramétrés avec la même fréquence d'acquisition et avoir une fréquence de transmission différente en fonction de leur niveau d'énergie. Un nœud de RCSF ayant assez d'énergie va transmettre toutes les données collectées avec une fréquence identique à la fréquence d'acquisition. Au contraire, un nœud ayant un faible niveau d'énergie ne communiquera, par exemple, qu'une valeur agrégée (max, min, etc.) de la journée à une fréquence moins élevée que la fréquence d'acquisition. Ainsi, les données générées par ces deux nœuds sont de nature différente même si, dans chaque cas, il s'agit de mesures de température. Par conséquent, pour une intégration intelligente des données d'un RCSF, il est nécessaire de décrire les flux de données générés par chaque nœud du RCSF. Les ontologies sont une solution pour décrire ces flux car elles permettent, entre autres, de définir des vocabulaires de métadonnées. Par le biais des ontologies, il sera ainsi possible d'associer à une donnée transmise par un nœud, la description de son environnement d'acquisition et de transmission pour améliorer le processus d'intégration. De plus, les ontologies sont aussi des composantes de système de raisonnement. Nous pouvons donc aussi les utiliser pour développer un système intelligent capable de modifier le comportement des nœuds afin d'optimiser la durée de fonctionnement du RCSF. Les ontologies pour nous ont deux objectifs : elles définissent des métadonnées et portent l'information nécessaire au raisonnement.

Le document est structuré de la façon suivante : la section 2 détaille nos besoins et nous permet de définir plusieurs notions relatives aux environnements d'acquisition et de transmission de données : *contexte, état, nœud et phénomène observé*. La section 3 définit et caractérise différents types d'ontologies. Puis, nous présentons notre état de l'art sur les ontologies de capteurs dans la section 4. Celui-ci se compose de trois parties : une description générique des ontologies de capteurs existantes, une analyse des thèmes décrits par ces ontologies, suivie d'une analyse critique sur les usages des ontologies dans les systèmes d'information alimentés par

des RCSF. Enfin, nous indiquons les ontologies qui répondent le mieux à nos besoins.

2. Description des besoins

Certaines informations sur le nœud source (ou collecteur) d'une donnée donnent une indication sur le contexte du nœud comme par exemple : son niveau d'énergie, sa politique d'acquisition, sa politique de transmission. Ceci permet notamment d'expliquer et d'étudier précisément les flux de données de manière à réaliser une intégration intelligente de ceux-ci. Aussi, nous définissons la notion de *contexte*¹ qui s'appuie sur celle d'*état*¹ :

Etat : "*L'état est une donnée qualitative, qui change dans le temps, résumant un ensemble d'informations*".

Contexte : "*Le contexte est un ensemble d'états d'entités ou d'informations décrivant un environnement dans lequel un événement se produit*".

Le contexte d'un nœud de RCSF contient les informations relatives à au moins deux événements : la prise de la mesure et la transmission de la donnée mesurée.

Les ontologies en tant que vocabulaire et schéma de métadonnées vont nous permettre d'explicitier le contexte du nœud.

Nous souhaitons que le nœud soit capable de modifier intelligemment sa politique d'acquisition et de transmission de données sous la supervision du réseau et du Système d'Information (SI). En plus du contexte du nœud, un autre contexte doit être pris en compte : le contexte du phénomène observé par le RCSF et le SI. Prenons l'exemple du phénomène environnemental de crue dans un bassin versant composé de plusieurs cours d'eau. Pour faciliter la surveillance du bassin versant, nous limitons la modélisation du phénomène de crue à une succession de quatre états :

- normal,
- en attente d'augmentation du niveau d'eau,
- augmentation du niveau d'eau,
- alerte de crue.

Ces états et leur succession sont le résultat d'une analyse des phénomènes de crues. Notons que l'état courant du phénomène observé peut être déduit de mesures effectuées par le RCSF. Par exemple, si le total des précipitations sur le bassin versant dans l'heure est supérieur à un seuil fixé, alors le phénomène de crue passe de l'état « *normal* » à celui d'« *en attente d'augmentation du niveau d'eau* ».

¹ Reformulation des définitions données par Wikipédia

Ainsi, les politiques d'acquisition et de transmission des données par un nœud de RCSF tiennent compte à la fois de l'état courant du nœud et de l'état courant du phénomène. L'ensemble de ces changements d'états peut être effectué par un système à base de règles. Ces systèmes constituent un type de systèmes experts. Le raisonnement est modélisé à l'aide de règles de décisions appliquées à un ensemble de faits. **Certains types d'ontologies définissent les entités utilisées dans les raisonnements à base de règles.**

Dans le cadre de notre étude, nous allons utiliser les ontologies pour répondre à plusieurs besoins :

– Normaliser le vocabulaire des RCSF. Ce vocabulaire devra définir les concepts concernant :

- la composition des RCSF et leur paramétrage. Les concepts concernés sont, par exemple, le nœud, le capteur, le module d'énergie, etc.
- l'acquisition de la mesure : le stimulus, la mesure, le phénomène observé, etc.
- la transmission de la donnée : les types de données, le flux de données, etc.

– Définir des formats de message à l'aide du vocabulaire normalisé pour construire les flux de données d'acquisition et les flux de données de transmission. Les messages donneront aussi des indications sur le contexte du nœud et ses changements d'états.

– Modéliser les entrées du raisonnement à base de règles. L'ontologie devra permettre de formaliser les faits utilisés par les règles. Les règles devront être capables de déduire les changements d'état des nœuds et les changements d'état du phénomène observé à partir des données transmises par le RCSF. De plus, un autre ensemble de règles devra piloter le comportement global du RCSF à partir des changements d'état du phénomène étudié en tenant compte de l'état des nœuds.

En résumé, notre ontologie doit nous permettre de poser un vocabulaire normalisé et un schéma de métadonnées utilisables par la communauté des RCSF. Ces métadonnées décriront différents contextes. L'ontologie devra aussi définir le format d'échange des données pour décrire les flux de données générés par le RCSF. Enfin, elle sera aussi le support de l'intelligence du nœud, du RCSF et du SI. Elle devra être utilisable dans un système à base de règles pour inférer de nouvelles données sur l'état du nœud et du phénomène observé et piloter le comportement du RCSF en optimisant la transmission de données.

Il existe différents types d'ontologies qui remplissent chacun des besoins différents. Pour clarifier ce que sont les ontologies et les typologies associées, nous allons dans la section suivante les définir de manière plus précise.

3. Typologies des ontologies

Le terme *ontologie* a été utilisé dans une grande variété de contextes applicatifs (système de recherche documentaire, interopérabilité des bases de données, etc.). Cependant, derrière ce terme se cachent des objets informatiques différents. Cette section a pour but d'une part, de clarifier la notion d'ontologie (3.1) et, d'autre part, de présenter différentes formes d'ontologies (3.2) et (3.3). Comme il existe plusieurs critères pour catégoriser les ontologies dans la littérature (Mizoguchi Riichiro et al., 1995), (Van Heijst et al., 1997), (Guarino, 1998) et (Lassila et al., 2001), nous avons choisi de présenter uniquement deux typologies : la première se définit à partir des objectifs de la conceptualisation, la seconde classe les ontologies en fonction du scénario de communication utilisé par le SI.

3.1. Définition des ontologies

La communauté en ingénierie des connaissances a proposé plusieurs définitions du terme *ontologie*. La définition la plus connue et la plus utilisée est celle de (Gruber, 1993) : une ontologie est « *une spécification explicite d'une conceptualisation* ». (Studer et al., 1998) ont complété ensuite cette définition. Une ontologie devient « *une spécification explicite et formelle d'une conceptualisation partagée. Une « conceptualisation » référence un modèle abstrait d'un phénomène du monde, en identifiant les concepts importants de ce phénomène. « Explicite » signifie que les types de concept utilisés et les contraintes sur leur usage sont explicitement définis. « Formelle » correspond au fait que l'ontologie doit être lisible par la machine. « Partagé » reflète la notion qu'une ontologie capture des connaissances consensuelles, c'est-à-dire non propres à un individu mais acceptées par un groupe* ». C'est cette définition que nous adopterons par la suite.

Dans le cadre de notre travail, la conceptualisation correspond au modèle abstrait décrivant les entités impliquées dans la prise de mesure, le traitement des données, leur transmission jusqu'au système d'information. Les entités sont, par exemple, le RCSF, les nœuds du RCSF, les capteurs et les flux de données générés. Pour formaliser ces ontologies, il existe plusieurs langages informatiques de modélisation. Nous verrons dans la section 3.3 que le choix du langage pour formaliser les ontologies dépend du niveau d'expressivité désiré.

3.2. Types d'ontologies par objectif de la conceptualisation

Cette typologie est proposée par (Gómez-Pérez et al., 2004).

– **Ontologies de représentation des connaissances** (Knowledge Representation or Meta Ontologies) : ces ontologies définissent les primitives utilisées dans les langages de représentation de connaissances. L'exemple le plus représentatif est la

*Frame-Ontology*². Elle modélise les primitives utilisées dans les langages à base de frames comme les classes, les individus, les attributs, les facettes, la relation sous-classe de, etc.

– **Ontologies de haut niveau** (Top Level, Upper or Foundational Ontologies) : ces ontologies proposent des concepts généraux qui doivent rester cohérents et applicables quel que soit le domaine d'étude. Ces concepts devront être spécialisés pour construire une ontologie propre à un domaine particulier. Ils formeront des points d'ancrage entre concepts spécialisés afin de faciliter la comparaison de différentes conceptualisations. *SUMO*³ (Suggested Upper Merged Ontology) et *DOLCE*⁴ (Descriptive Ontology for Linguistic and Cognitive Engineering) sont des exemples d'ontologies de haut niveau. Elles proposent des classes d'entités physiques ou abstraites et les primitives permettant de les distinguer les unes des autres. Des ontologies de domaines, spécialisant une même ontologie de haut niveau, pourront être fusionnées car elles utilisent les mêmes primitives de base.

– **Ontologies génériques** (General or Common Ontologies) : ces ontologies représentent des connaissances communes à plusieurs domaines. Elles définissent des notions d'unités, de types de données ou des entités génériques comme des événements, etc. *Mereology Ontology*⁵ est un exemple d'ontologie générique. Elle détaille les différents types de relations d'appartenance possibles.

– **Ontologies de domaine générique** (Core or Generic Domain Ontologies) : ces ontologies se focalisent sur un domaine particulier dont elles proposent des concepts et des relations permettant de structurer ce domaine. Par exemple, dans le domaine médical, le meta thesaurus *UMLS*⁶ (Unified Medical Language System) contient des concepts comme les maladies et les syndromes, les substances biologiques, les fonctions physiologiques et des relations associées. A l'aide de ces réseaux d'entités génériques, UMLS est capable de relier et d'intégrer plusieurs thésaurus médicaux spécialisés.

– **Ontologies de domaine** (Domain Ontologies) : ces ontologies modélisent la conceptualisation d'un domaine d'étude en particulier (santé, biologie, réparation automobile, construction de bâtiments, etc.). L'ontologie *GENE*⁷ est un exemple d'ontologie de domaine. Elle vise à standardiser la formalisation des gènes.

– **Ontologies d'application** (Application Ontologies) : ces ontologies dépendent d'une application particulière. Elles contiennent toutes les définitions pour modéliser une connaissance requise par le fonctionnement d'une application. Par exemple, nous pourrions créer une ontologie dédiée aux agences de voyages françaises afin de les aider à rechercher des voyages en fonction de leur destination.

² <http://www-ksl.stanford.edu/knowledge-sharing/ontologies/html/frame-ontology/index.html>

³ <http://www.ontologyportal.org/>

⁴ <http://www.loa.istc.cnr.it/DOLCE.html>

⁵ <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/borst/kaw96doc.html>

⁶ <http://www.nlm.nih.gov/research/umls/>

⁷ <http://www.geneontology.org/>

3.3. Types d'ontologies par scénario de communication

Les ontologies ont pour but de faciliter la communication entre différents agents humains ou logiciels. Pour permettre à ces agents de communiquer correctement, il est nécessaire de résoudre les problèmes d'hétérogénéité des messages à plusieurs niveaux : lexical, structurel, syntaxique et sémantique. Pour expliquer les différentes solutions d'interopérabilité, (Roussey et al., 2010) proposent plusieurs scénarios de communication. La Figure 1 issue de ces travaux présente les types d'ontologies utilisés pour chacun des scénarios :

– Les **ontologies terminologiques** sont orientées principalement pour favoriser la communication entre humains. Elles contiennent des définitions en langage naturel des concepts et leur associent des termes pour les nommer sans ambiguïté. Des liens entre concepts sont aussi possibles pour aider à identifier les concepts connexes. Ces ontologies ont pour objectif de normaliser le vocabulaire d'un domaine et de définir chaque terme de ce vocabulaire. Elles constituent un réseau terminologique sur lequel il n'est pas possible d'effectuer des raisonnements logiques. Le langage informatique préconisé par le W3C (World Wide Web Consortium) pour formaliser ce type d'ontologies est SKOS (Simple Knowledge Organization System). Un autre langage possible est le schéma XML pour stocker les Topic Maps XTM (XML Topic Maps). Un exemple d'ontologie terminologique très utilisé en recherche d'information est la base lexicale Wordnet (WordNet 2012) contenant la description de tous les sens des termes de la langue anglaise.

– Les **ontologies de données** ont pour objectif de proposer une description structurelle et syntaxique des concepts du domaine et de leurs propriétés. Elles ont vocation à définir des formats d'échange ou des schémas de données pour faciliter la communication entre logiciels. Dans ces ontologies, un concept se définit comme un agrégat de données auquel il est possible d'associer des contraintes d'intégrité sur les valeurs de ces données. Les langages informatiques les plus utilisés pour construire ce type d'ontologies sont OWL (Web Ontology Language) ou UML (Unified Modeling Language). Un exemple est le format d'échange de données sur l'eau SANDRE (Service d'Administration National des Données et Référentiels sur l'Eau) utilisé en France.

– Les **ontologies logiques** contiennent des descriptions logiques des concepts et des relations. Elles sont utilisées dans des systèmes d'information travaillant sur plusieurs sources de données hétérogènes. Dans ce cadre, leur but est de permettre l'intégration des données ou la médiation entre ces sources. Ces ontologies contiennent forcément des formules logiques pour pouvoir être utilisées par des moteurs d'inférences. Ces moteurs permettent de valider le modèle abstrait sous-jacent à la conceptualisation, de détecter la classe d'appartenance d'une instance ou de reconnaître les instances d'une classe et de générer de nouvelles connaissances à partir de règles. Pour formaliser ce type d'ontologies, on peut utiliser différents langages logiques, comme les langages basés sur les logiques de description OWL DL ou KIF (Knowledge Interchange Format).

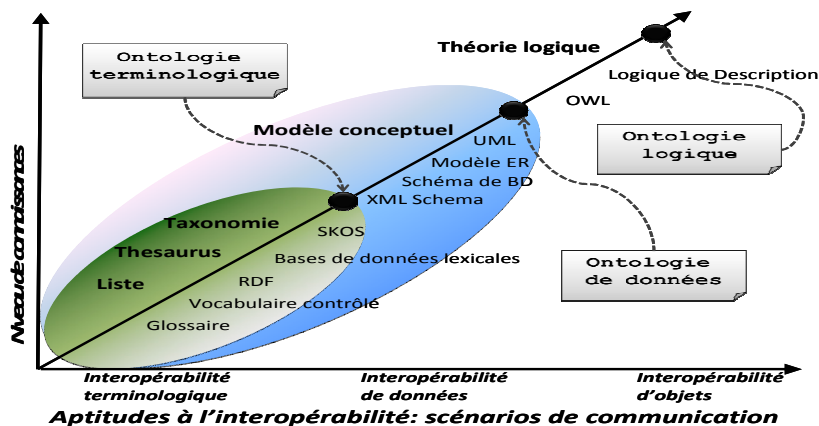


Figure 1 : Niveaux d'interopérabilité et types d'ontologies (Roussey et al., 2010)

4. État de l'art sur les ontologies de capteurs

Dans cette section, nous allons présenter un état de l'art des ontologies de capteurs. La section 4.1 présente de manière générale les ontologies de capteurs existantes. La section 4.2 propose une analyse critique des thèmes abordés par ces ontologies. Cette étude va nous permettre de choisir l'ontologie de capteur la plus adaptée à nos besoins (cf. section 2). La section 4.3 détaille les usages des ontologies de capteurs dans les SI alimentés par des RCSF.

4.1. Présentation générale des ontologies de capteurs existantes

Dans cette section, nous allons introduire les ontologies de capteurs qui vont être analysées dans les sections (4.2) et (4.3).

L'ontologie **SSN**⁸ (Semantic Sensor Network) a été créée par le groupe de travail W3C SSN Incubator Group. Elle a été utilisée dans plusieurs projets comme, par exemple, le projet « SmartProducts⁹ » et le projet « météorologie en agriculture¹⁰ ». Le projet « SmartProducts » met en place des couteaux équipés d'un accéléromètre, capteur de mouvement, qui reconnaît les actions de l'utilisateur. Les observations sont transmises à un dispositif nommé « aide de cuisine » qui guide l'utilisateur à cuisiner ou à suivre une recette.

⁸ <http://www.w3.org/2005/Incubator/ssn/XGR-ssn-20110628/>.

⁹ <http://www.smartproducts-project.eu/>

¹⁰ <http://www.csiro.au/science/Wireless-sensors-in-agriculture>

L'ontologie **CESN** (Calder et al., 2010) a été développée dans le cadre du projet « CESN Semantic Data Reasoner ». Les auteurs de ce projet ont réalisé un système de raisonnement à base de règles qui sont en mesure de valider les données issues des réseaux de capteurs.

L'ontologie **CSIRO** (Compton et al., 2009) a été développée par le centre CSIRO ICT. Elle a été prévue pour faciliter l'intégration, la recherche et la classification de données issues de capteurs. A notre connaissance, elle n'a jamais été utilisée car elle a été intégrée dans l'ontologie **SSN** construite peu après.

L'ontologie **Stimuli-centered** (Stasch et al., 2009) a été conçue par l'Université de Muenster et le « National Institute for Space Research (INPE) ». Le but de cette ontologie est la conception d'un patron de conception, réutilisé dans l'ontologie **SSN**. C'est pourquoi on remarque l'absence de projet associé à cette ontologie.

Sensei O&M (Barnaghi et al., 2009) a été créée par l'Université de Surrey. Elle a été développée dans le cadre du projet européen « SENSEI ». Le but du projet est l'intégration de RCSF hétérogènes dans une plate-forme commune rendant les services et applications accessibles via une interface universelle.

L'ontologie **OOSTethys** (Bermudez et al., 2009) a été développée et utilisée par une communauté de développeurs logiciels et de scientifiques du domaine maritime. Ceux-ci développent des outils « open source » pour l'intégration des systèmes d'observation océanographique.

L'ontologie **MMI** (Marine Metadata Interoperability) est une composante océanographique. Elle a été développée dans le cadre du projet « MMI ». Le but de ce projet est de fournir l'orientation, le vocabulaire et les services sémantiques pour la communauté des marins.

L'ontologie **SWAMO**¹¹ a été développée par la NASA dans le cadre du projet « NASA ROSES AIST 2005 ». L'objectif du projet SWAMO est de passer d'un contrôle terrestre centralisé des vaisseaux spatiaux à un autre basé sur un ensemble d'agents intelligents collaboratifs et distribués.

SEEK « Extensible Observation Ontology » (Madin et al., 2007) est une ontologie développée par l'Université de Californie. Elle a été conçue pour le projet « SEEK » et utilisée par le projet « SPIRE ». L'objectif du projet SEEK est de faciliter non seulement l'acquisition de données et leur archivage mais, également, l'intégration, la transformation, l'analyse et la synthèse des données écologiques et de biodiversité qui, pour certaines, étaient auparavant intraitables.

SDO « Sensor Data Ontology » (Eid et al., 2007) a été développée par l'Université d'Ottawa. Elle n'est pas disponible, seul son schéma a été publié en 2007. A notre connaissance, elle n'a jamais été utilisée dans un projet.

¹¹ <http://marinemetadata.org/references/ontswamo>.

L'ontologie **SeRes O&M** (Probst, 2006) a été développée par le projet « Semantic Reference Systems ». Le but de ce projet est de permettre aux fournisseurs d'informations géospatiales d'annoter leurs sources de données pour permettre aux demandeurs d'information de retrouver les sources correspondant à leur besoin.

OntoSensor (Russomanno et al., 2005) a été créée par l'Université de Memphis pour la conception d'une base de connaissance répertoriant des capteurs. Cette base de connaissance permet de retrouver un capteur à partir d'une requête spécifiant les services offerts.

4.2. Analyse des thèmes développés dans les ontologies de capteurs

Dans le cadre de notre étude, l'ontologie qui répond à nos besoins doit décrire au mieux les différents aspects du RCSF et de ses composants (Sensor), du processus de mesure (Observation) et des flux de données générés par les nœuds du RCSF (Data). Dans la quête de l'ontologie la plus appropriée pour satisfaire nos besoins, nous allons analyser ces douze ontologies de capteurs suivant trois aspects. Cette analyse sera le sujet de cette section. Le Tableau 1 synthétise les différents thèmes ci-dessous portés par les ontologies :

Sensor : un des thèmes qu'une ontologie peut décrire est **le réseau de Capteurs**, système électronique. Les ontologies qui décrivent ce thème mettent l'accent sur la description physique du réseau et des capteurs. Elles permettent de représenter les capteurs, leurs différents types, leurs composants, les procédures et les traitements. Elles permettent également de décrire le déploiement et la configuration des réseaux.

Observation : un deuxième thème qu'une ontologie peut décrire est **l'Observation d'un phénomène**. Une situation où un agent physique, le capteur, réalise une mesure suite à la détection d'un stimulus. Cette mesure permet de caractériser un phénomène observé, par exemple la vitesse du vent. Les ontologies qui décrivent ce thème tendent à décrire le capteur et les mesures impliquées. Elles permettent de représenter le processus d'acquisition de la mesure ainsi que les caractéristiques de celui-ci comme la précision et la fréquence. Ces ontologies permettent également de décrire les domaines de valeurs et les types de valeurs des mesures (*field of view*).

Data : un troisième thème qu'une ontologie peut décrire est **le flux de données**. Une donnée, résultat obtenu par une observation, est acquise, traitée puis transmise. Les ontologies qui intègrent ce thème tendent à décrire les données générées par le réseau de capteurs. Elles doivent permettre de décrire les données (et leur domaine de valeur) acquises et communiquées, les politiques d'acquisition et de communication du nœud.

Dans le Tableau 1, la présence d'une étoile indique la capacité de l'ontologie à décrire le thème indiqué, son absence le contraire.

Ontologies	Sensor						Observation						Data						
	Sensor	Sensor type	Components	Deployment	Configuration	Action & process	Observation	Response model	Measurement	Accuracy	Frequency	Field of view/sensing	Data	Data stream	Data Stream acquisition	politic	Data stream communication	politic	Domain
SSN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
CESN	*			*			*						*						*
CSIRO	*	*	*	*	*	*	*	*	*	*	*	*	*		*				*
Sensei O&M							*						*						
OOSTethys			*			*	*						*						*
MMI	*		*	*	*	*			*	*					*				*
SWAMO	*		*			*						*							*
SEEK							*	*					*						
SDO	*	*			*	*	*	*	*	*	*	*	*		*				*
SeReS O&M					*	*	*	*					*						
OntoSensor	*	*	*		*	*	*	*	*	*	*	*	*	*	*				*

Tableau 1 : Thèmes des ontologies de capteurs

Les ontologies réseaux de capteurs (sensor ontologies) : *Ontosensor*, *CSIRO*, *SSN*, *SWAMO* et *SDO* sont des exemples de ce type. L'ontologie *OntoSensor* fournit une brève description des capteurs et de leurs capacités. L'ontologie *SWAMO* décrit la plate-forme du réseau. L'ontologie *CESN* décrit les composants des capteurs. Cependant, comme nous pouvons le voir dans le Tableau 1, *CSIRO*, *SSN* sont les seules ontologies qui permettent de représenter toutes les facettes d'un RCSF. Avec un développement modulaire, elles nous permettent de modéliser les dispositifs du capteur (et leurs capacités), les procédures, le système, etc. *CSIRO*, *SSN* sont les ontologies les plus appropriées pour représenter le thème Sensor.

Les ontologies observation d'un phénomène (observation ontologies) : l'ontologie *SEEK* est une ontologie logique, qui appartient à cette catégorie. Cette ontologie fournit des schémas génériques d'observation et de mesure. L'ontologie *stimulus-centred* généralise le processus de mesure et introduit la notion de stimulus. Le stimulus est le déclencheur d'une opération de mesure, donc d'une observation. Le stimulus est un concept qui reflète ce qui se passe dans le monde réel. Basé sur le patron de conception *stimulation-Sensor-observation*, les ontologies *CSIRO* et *SSN* décrivent les observations, les propriétés observées, les capteurs et la méthode de détection utilisée. *CSIRO*, *SSN*, *SDO* et *Ontosensor* sont les ontologies les plus appropriées pour représenter le thème Observation. L'ontologie *SSN*, comme la plupart des ontologies d'observation met, l'accent sur le processus d'acquisition des données, mesures d'un phénomène étudié. Cependant, ces ontologies ne se soucient

pas de la communication des données acquises (communication des flux de données).

Les ontologies flux de données (data ontologies) : l'ontologie *SDO* est un exemple de ce genre d'ontologies. *SDO* donne juste une simple classification des données de capteurs. Dans l'ontologie *SSN*, les données générées par un réseau de capteurs sont stockées dans la classe *SensorOutPut* et leurs valeurs dans la classe *observationValue*. Cependant, aucune des ontologies précédemment citées ne caractérise les flux de données générés par le RCSF et leurs caractéristiques (politique d'acquisition et de communication), d'où l'intérêt de décrire les données générées par le RCSF.

Nous sommes à présent en mesure de répondre à notre question de départ, à savoir trouver l'ontologie qui décrit le mieux les trois thèmes précédemment cités. En effet, après la première partie de notre état de l'art des ontologies de capteurs, nous pouvons constater qu'aucune de ces ontologies de capteurs ne décrit ces trois aspects qui sont Sensor, Observation et Data. Cependant, *SSN* qui intègre plusieurs ontologies dont *CSIRO* modélise le mieux les deux premiers aspects.

4.3. Analyse des usages des ontologies de capteurs

Le Tableau 2 présente un état de l'art des ontologies de capteurs en fonction de leur utilisation. Les ontologies étudiées sont généralement développées dans le cadre d'un projet. Pour comparer ces différentes ontologies, nous nous sommes appuyés sur les critères suivants :

- Le langage d'interrogation. Dans la littérature, il existe plusieurs langages d'interrogation d'ontologies comme, par exemple, SPARQL et RDQL. SPARQL est le langage le plus utilisé par les ontologies de capteurs étudiées pour retrouver un capteur et les informations sur son environnement en fonction de différents critères exprimés dans une requête. La colonne « query langage » du Tableau 2 indique pour chaque ontologie son langage d'interrogation.

- L'usage de l'ontologie. Ces ontologies sont développées pour : mettre en place un vocabulaire (1) ; définir un format d'échange pour faciliter le transfert des données (2) ; interroger des données décrivant ou issues des capteurs (3) ; intégrer des données pour accéder uniformément à des sources de données hétérogènes (4) ; produire de nouvelles connaissances (5) qui sont le résultat de l'application de règles d'inférence sur des données de bas niveau (les données produites par les capteurs) ; détecter les incohérences (6). La colonne « use of ontology » précise des usages possibles pour chacune des ontologies étudiées : de 1 à 6.

- Les techniques utilisées pour implémenter des raisonnements. Dans le Tableau 2, la colonne « inferences » définit les objectifs des inférences et la colonne suivante le type de moteur d'inférence utilisé.

Ontologies	Metrics	Knowledge representation languages used	Query language	Ontology type	Use of ontology	Inferences	Inferences tools
SSN	Class count (117), individual count (0), Object property count (142) and Data property count (6)	OWL DL	SPARQL	Domain ontology	Describe sensor and observation (1,2, 3, 5, 6)	Interrogate high-level events from low-level events	No information
CESN	Class count (36), individual count (18), Object property count (6) and Data property count (12)	OWL DL	No information	Domain ontology	Validated the raw data from sensors to pass raw data to a qualitative higher level event detection (1,5, 6)	For making inferences about data and anomalies in measurements.	DL reasoning and logic programming rules
CSIRO	Class count (71), individual count (0), Object property count (61) and Data property count (11)	OWL DL	SPARQL	Domain ontology	Sensor description and composition sensor (1, 3, 5, 6)	Allow the composition of sensors	DL reasoning
Stimuli-centered	No information	Algebraic specification	Not yet	Top-level ontology	Describe sensors and their observation (1)	No inferences	No information
Sensei O&M	Class count (71), individual count (104), Object property count (71) and Data property count (42)	OWL DL	SPARQL	Domain ontology	To annotate sensor observation and measurement data; describe sensor data (1,2, 3)	No information	No information
OOSTethys	No information	OWL/UML	No information	Domain ontology	Introduce a conceptual model for observation système (1, 4)	No information	No information
MMI	Class count (71), individual count (0), Object property count (61) and Data property count (11)	OWL DL	No information	Domain ontology	Marine vocabulary, description and developed 13 Uses case (1)	No inferences	None
SWAMO	Unavailable	No information	SPARQL	Domain ontology	Enable dynamic, composeable interoperability of sensor web products and services (1, 2, 3, 4, 6)	To check the consistency of the ontologies	DL reasoning
SEEK	No information	OWL DL	No information	Domain ontology	Discovery and integration of ecological data (1, 4)	No inferences	None
SDO	Unavailable (paper only)	OWL/KIF/OWL/LOOM	RDQL	Domain ontology	To search relevant sensor data over distributed and heterogeneous sensor networks (1,3, 4, 5)	To make inferences about data emanating from sensor networks	LP rules
SeReS O&M	Class count (204), individual count (0), Object property count (209) and Data property count ()	OWL DL	None	Top-level ontology	Normalise a high level vocabulary (1)	No information	No information
OntoSensor	Class count (298), individual count (119), Object property count (125) and Data property count (108)	OWL DL	Protégé2000 Query plugin	Domain ontology	Sensor description and service discovery ; knowledge base in which we display its services (1,2,3)	Not yet (next features)	DP and PL

Tableau 2 : Utilisation des ontologies de capteurs

L'objectif de cette partie est de découvrir quelles sont les ontologies utilisables par un système intelligent capable de raisonner.

Les ontologies *SEEK* et *MMI* sont des ontologies terminologiques. Ces ontologies ne permettent pas de raisonner. Les ontologies *SeReS O&M* et *Stimuli-centered ontology* sont des ontologies de haut niveau. Les ontologies de ce type n'ont pas pour but d'être utilisées directement dans un raisonnement.

L'ontologie *CSIRO* permet la composition des capteurs : par exemple, pouvoir composer un capteur de température (thermomètre) et un capteur de mesure de la vitesse du vent (anémomètre) pour obtenir un capteur virtuel plus sophistiqué permettant de calculer le « wind chill temperature » (la température de refroidissement éolien). Pour réaliser cette composition, *CSIRO* utilise un système de raisonnement basé sur les logiques de description.

Le raisonnement est uniquement cité en perspective dans l'ontologie *OntoSensor*.

Les auteurs de *SWAMO* utilisent les raisonneurs de logiques de description pour la vérification de la consistance de cette ontologie.

L'ontologie *SDO* utilise des règles de la logique de programmation pour faire des inférences sur les données provenant d'un réseau de capteurs. Les règles permettent de valider les données issues des capteurs. L'ontologie *CESN* exploite des règles de la logique du premier ordre pour détecter des anomalies sur des événements. Elle permet ainsi de valider les données brutes collectées par le réseau de capteurs.

L'ontologie *SSN* a été utilisée dans plusieurs projets. Le groupe W3C a donné un exemple applicatif sur un phénomène météorologique. Dans cet exemple, un système permet d'interroger des données de haut niveau. Ces dernières sont le résultat d'un raisonnement effectué sur les données de bas niveau issues du RCSF.

En résumé, les ontologies étudiées dans cet état de l'art utilisent plusieurs techniques pour effectuer des raisonnements sur les éléments décrits avec l'ontologie. Les techniques utilisées sont généralement basées sur des moteurs d'inférences des logiques formelles. Les logiques les plus utilisées sont les logiques de description. Elles sont donc les plus adéquates pour exprimer nos besoins en termes de raisonnement et de génération de nouvelles connaissances.

5. Conclusion et perspectives

Définies comme une « *spécification explicite d'une conceptualisation* », les ontologies sont, entre autres, utilisées pour fournir un vocabulaire dans un domaine donné, un format d'échange, une médiation des données et générer de nouvelles connaissances. Dans le cadre de la problématique de l'intégration de données issues de réseaux de capteurs sans fil (RCSF), nous nous sommes donc intéressés aux ontologies de capteurs. De cette étude, a résulté l'état de l'art présenté dans cet article. Dans la première partie de celui-ci, nous avons constaté que l'ontologie *SSN* est l'ontologie qui décrit le mieux la plupart des thèmes qu'une ontologie de capteur devrait décrire. Toutefois, quelques manques demeurent dans l'ontologie *SSN*. En

effet, cette ontologie ne décrit pas les flux de données générés par le RCSF et leurs politiques de communication, d'où l'intérêt de l'enrichir. Pour satisfaire nos besoins en termes de raisonnement et d'utilisation des ontologies de capteurs, les logiques de description semblent être les plus adéquates.

En guise de perspectives, nous allons dans un premier temps enrichir l'ontologie SSN en intégrant les notions manquantes : RCSF, nœuds du RCSF, flux de données, etc. Et, dans un second temps, nous allons mettre en œuvre cette ontologie dans une plate-forme réelle de RCSF afin d'intégrer les données issues de ces réseaux tout en optimisant l'énergie consommée.

Remerciements

Les auteurs tiennent à remercier les relecteurs pour leurs remarques pertinentes qui nous ont permis d'améliorer l'article et le « Conseil Régional d'Auvergne » pour son soutien financier.

Références

- Akyildiz I.F. et al., « Wireless sensor networks: a survey ». *Computer Networks*, 2002, p.393 - 422.
- Barnaghi P. et al., « Sense and Sens'ability: Semantic Data Modelling for SensorNetworks ». *In Proceedings of the ICT Mobile Summit 2009*, 2009, p.1-9.
- Bermudez L. et al., « Ocean observing systems demystified ». *In OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*, 2009, p.1 -7.
- Calder M., Morris OA. et Peri F., « Machine reasoning about anomalous sensor data ». *Advances in environmental information management*, 2010, p. 9-18.
- Compton M. et al., « Reasoning about Sensors and Compositions ». *Proc. Semantic Sensor Networks.*, 2009, p.33-48.
- Dictionnaire Hydrographique International, 2012.
http://www.loria.fr/projets/MLIS/DHYDRO/outils/site_edition/byproducts/Tout_Martif_en_fr.html.
- Eid M., Liscano R. et El Saddik A., « A Universal Ontology for Sensor Networks Data ». *Computational Intelligence for Measurement Systems and Applications, 2007. CIMSA 2007. IEEE International Conference on*, 2007, p.59 - 62.
- FORUM 2012. <http://www2.lirmm.fr/~bella/FORUM/>.
- Gómez-Pérez A., Fernández-López M. et Corcho O., *Ontological Engineering: with examples from the areas of KnowledgeManagement, e-Commerce and the Semantic Web*, 2004.
- Gruber T.R., « Toward Principles for the Design of OntologiesUsed for Knowledge Sharing ». *International Workshop on Formal Ontology, Padova, Italy*, 1993.
- Guarino N., « Formal Ontology and Information Systems ». *Amended version of a paper appeared in N. Guarino (ed.), Formal Ontology in Information Systems. Proceedings of FOIS'98*, 1998, p.3-15.

- Van Heijst G. et SCHREIBER A. T.H., « Using explicit ontologies in KBS development ». *Int. J. Human – Computer Studies* 45, 1997, p.183 – 292.
- Lassila O. et McGuinness D., The Role of Frame-Based Representation on the Semantic Web. *Knowledge Systems Laboratory Report KSL-01-02*, 2001.
- Madin O. et al., « An ontology for describing and synthesizing ecological observation data ». *A Special Feature from the 5th International Conference on Ecological Informatics ISEI5*, 2007, p.279 - 296.
- Marine Metadata Interoperability, MMI Device Ontology: A Community Development Project.
- Mizoguchi R., Vanwelkenhuysen J. et Ikeda M., « Task Ontology for Reuse of Problem Solving Knowledge ». *Knowledge Building & Knowledge Sharing (KB&KS'95) (2nd International Conference on Very Large-Scale Knowledge Bases) Enschede, The Netherlands*, 1995, p.46-59.
- Probst F., « An Ontological Analysis of Observations and Measurements ». *In Proc. of the 4th. International Conference on Geographic Information Science (GIScience)*, 2006.
- Roussey C. et al., « Ontologies in Agriculture ». *AGENG 2010 Conference.*, 2010.
- Russomanno D.J., Kothari C. et Thomas O., « Sensor ontologies: from shallow to deep models ». *System Theory, 2005. SSTS '05. Proceedings of the Thirty-Seventh Southeastern Symposium on*, 2005, p.107 - 112.
- Stasch C., Janowicz, K. et Broring A., « A Stimulus-Centric Algebraic Approach to Sensors and Observations ». *In CIAO '09: Proceedings of the 1st Workshop on Context, Information and Ontologies*, 2009, p.169.
- Studer R., Benjamins V.R. et Fensel D., « Knowledge engineering: principles and methods ». *IEEE Transaction on Data & knowledge engineering* 25 (1-2), 1998..
- WordNet 2012. <http://www.wordnet-online.com/>.