

# Chapitre 1

## Notions Préliminaires

### Sommaire

---

1.1	Introduction . . . . .	9
1.2	Qu'est ce qu'un Service Web? . . . . .	10
1.3	Modèle de base des Services Web . . . . .	11
1.4	Les Services Web DaaS (Data as a Service) . . . . .	12
1.5	Les Services Web sémantiques . . . . .	13
1.6	Ontologies pour le Web sémantique . . . . .	15
1.7	RDF/RDFS . . . . .	16
1.8	SPARQL . . . . .	18
1.9	Conclusion . . . . .	19

---

### 1.1 Introduction

Dans ce chapitre, nous présentons quelques concepts de base qui nous semblent indispensables. Dans la section 1.2, nous donnons quelques définitions du concept "Service Web" et nous présentons son modèle dans la section 1.3. De la même façon, nous introduisons les concepts "Service Web DaaS" et "Services Web sémantiques" respectivement dans les sections 1.4 et 1.5. Ensuite, nous définissons le concept "ontologie" (section 1.6), les langages de description RDF/RDFS (section 1.7) et enfin, le langage d'interrogation SPARQL (section 1.8).

## 1.2 Qu'est ce qu'un Service Web ?

Le Web est de plus en plus le support privilégié des applications. Les Services Web constituent le développement ultime dans ce domaine. Un Service Web est souvent vu comme une application accessible à d'autres applications sur le Web. Cependant, il existe plusieurs définitions pour le Service Web, nous citerons quelques unes :

**Définition 1 :** Le consortium W3C<sup>1</sup> définit un Service Web comme étant :« **une application, ou un composant logiciel qui vérifie les propriétés suivantes** » [23] :

- Il est identifié par un URI<sup>2</sup> ;
- Ses interfaces et ses liens peuvent être décrits en XML ;
- Sa définition peut être découverte par d'autres Services Web ;
- Il peut interagir directement avec d'autres Services Web à travers le langage XML en utilisant des protocoles Internet standards.

**Définition 2 :** « Un Service Web est une application accessible à partir du Web. Il utilise les protocoles Internet pour communiquer et un langage standard pour décrire son interface » [12].

**Définition 3 :** « Les Services Web sont la nouvelle vague des applications Web. Ce sont des applications modulaires, auto-contenues et auto-descriptives qui peuvent être publiées, localisées et invoquées depuis le Web. Les Services Web effectuent des actions allant de simples requêtes à des processus métiers complexes. Une fois qu'un Service Web est déployé, d'autres applications (y compris des Services Web) peuvent le découvrir et l'invoquer » [14].

---

1. World Wide Web Consortium

2. Uniform Resource Identifier

## 1.3 Modèle de base des Services Web

Le modèle des Services Web repose sur une Architecture Orientée Service (SOA)<sup>3</sup>. Celle-ci fait intervenir trois catégories d'acteurs : **le fournisseur de service**, **l'annuaire des services** et **le client**. Un fournisseur de service fournit un module logiciel accessible sur le réseau et définit une description de service pour le Service Web. Ensuite, il le publie dans l'annuaire des services de telle sorte que le client peut le trouver. L'annuaire des services correspond à un registre de descriptions des Services Web offrant des facilités de publication des Services Web à l'intention des fournisseurs, ainsi que des facilités de recherche des Services Web à l'intention des clients. La description de service contient des informations telles que l'entrée du service, sa sortie et l'adresse où le service est situé. Le client interroge l'annuaire pour un certain type de service et récupère sa description. Ensuite, il utilise les informations dans la description du service pour se lier avec le fournisseur et d'invoquer le Service Web [6]. Les interactions de base entre ces trois acteurs montrées dans la figure 1.1 incluent les opérations de publication, de recherche et de liens d'opérations. Nous citons, notamment les standards émergents suivants :

- **SOAP**<sup>4</sup> : Définit un protocole de transmission de messages basé sur XML. Il permet l'échange d'informations à distance en utilisant le formalisme XML pour à la fois, définir les messages envoyés entre les applications et représenter les données échangées.
- **WSDL**<sup>5</sup> : Un standard fondé sur XML qui permet la description des Services Web.
- **UDDI**<sup>6</sup> : Fournit l'infrastructure de base pour la publication et la recherche des Services Web.

---

3. Service Oriented Architecture

4. Simple Object Access Protocol

5. Web Service Description Language

6. Universal Description Discovery and Integration

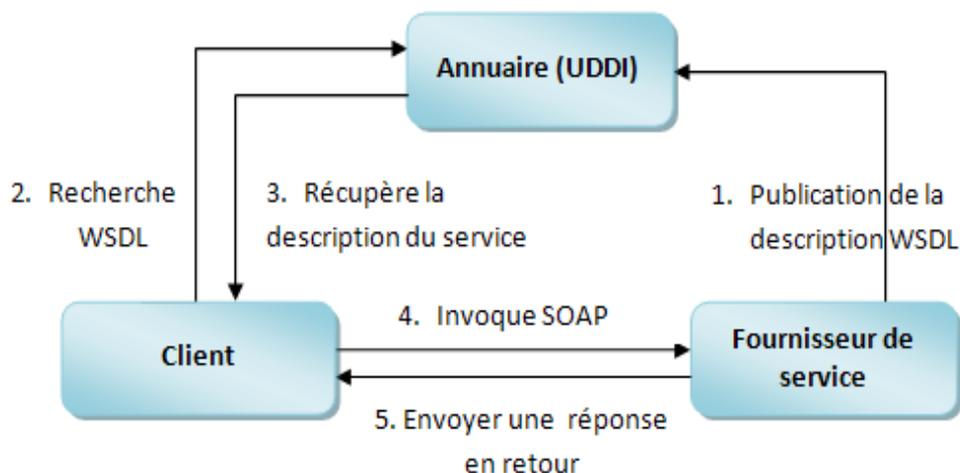


FIGURE 1.1 – Modèle du Service Web

## 1.4 Les Services Web DaaS (Data as a Service)

Au cours de la dernière décennie, les Services Web ont été largement perçus comme un moyen standardisé pour l'intégration d'applications sur le Web. Ils reposent sur une architecture orientée service permettant ainsi aux applications de différents fournisseurs d'être encapsulées comme des services, puis publiées, localisées, invoquées, composées et coordonnées de manière à couplage faible. Récemment, les Services Web ont commencé à être un support populaire pour la publication et le partage des données sur le Web.

Les entreprises modernes évoluent vers une architecture orientée service de partage de données sur le Web en mettant leurs bases de données derrière les Services Web, fournissant ainsi la méthode interopérable d'interagir avec leurs données. En outre, les données qui ne sont pas stockées dans des bases de données traditionnelles sont également mises à disposition via des Services Web. Nous appelons ce type de Services Web en tant que Services Web DaaS.

« Un Service DaaS fournit une vue simplifiée, intégrée en temps réel , une information de haute qualité sur une entité commerciale spécifique comme un client ou un produit. Il peut être fourni par le middleware ou emballé comme un composant logiciel individuel. Les informations qu'il fournit proviennent d'un ensemble diversifié de ressources d'informations » [11].

Une autre définition de Composite Software : « **Les Services Web DaaS sont une forme de Services Web optimisée pour les demandes d'intégration de données en temps réel. Ils visualisent les données pour découpler les emplacements physiques et logiques et donc, d'éviter la réplication des données inutiles. Les Services Web DaaS résument des structures de données complexes et de la syntaxe. Ils fédèrent des données disparates en composites utiles et supportent l'intégration de données à travers les applications de l'architecture orientée service** »<sup>7</sup>.

Les Services Web DaaS permettent l'accès aux sources de données des organisations. L'invocation d'un service DaaS résulte dans l'exécution d'une requête sur le schéma de la source de données. Lorsqu'un tel service est exécuté, il accepte d'un utilisateur une donnée d'entrée d'un format spécifié et il lui retourne des informations comme une sortie. Les Services Web DaaS sont maintenant utilisés dans de nombreux domaines d'application comme un moyen standard pour la publication et le partage des données. Exemples de domaines d'application comprennent, entre autres, le partage des données scientifiques (par exemple, la bioinformatique, traitement et partage de données géospatiales , etc), le partage des données médicales (eHealth), les entreprises d'intégration de données, le partage des données entre les organismes gouvernementaux (eGovernment), etc).

## 1.5 Les Services Web sémantiques

Les Services Web sémantiques se situent à la convergence de deux domaines de recherche importants qui concernent les technologies d' Internet : le Web sémantique et les Services Web. Cette tâche de convergence est accomplie en rendant les Services Web auto-exploitable par machines, et de réaliser l'interopérabilité entre les applications via le Web en vue de rendre le Web plus dynamique.

En effet, le Web sémantique a été inventé en 1989 par Tim Berners-Lee. Ce dernier a proclamé que le Web sémantique est la prochaine évolution du Web, c'est-à-dire que l'on

---

7. <http://compositesoftware.com/solutions/soa.shtml>

va arriver à un Web intelligent où les informations sont stockées de façon compréhensible par les machines afin d'apporter à l'utilisateur ce qui cherche vraiment. Aujourd'hui, les humains sont les seuls qui ont la capacité de comprendre ce que nous trouvons et de décider en quoi cela se rapporte à ce que nous voulons chercher vraiment. Par quels moyens ? Ce sont les moteurs de recherche qui nous aident, mais ils sont capables de répondre seulement aux deux questions : Quelles sont les pages contenant un terme ? et Quelles sont les pages les plus populaires à un sujet ?.

« **Le Web sémantique est une extension du Web actuel dans lequel l'information est munie d'une signification bien définie permettant aux machines et aux personnes de travailler en coopération** » [19]. Le Web sémantique reste entièrement fondé sur le Web classique et ne le remet pas en cause. Cela reste un moyen de publier et consulter des documents, mais les documents traités par le Web sémantique contiennent non pas des textes en langage naturel mais des informations formalisées pouvant être traitées automatiquement par des agents logiciels. Un des objectifs du Web sémantique est d'affiner la recherche sur Internet. Ceci est réalisé par l'annotation du contenu du Web en ajoutant aux informations existantes une couche de métadonnées pour améliorer la compréhension pour la machine. Pour le faire, on utilise le vocabulaire conceptuel fourni par une *ontologie*.

Un Service Web sémantique est un Service Web décrit en utilisant des annotations sémantiques dans un langage du Web sémantique bien défini, qui permettent au Service Web d'avoir une interface compréhensible par les humains et les machines. Ces Services Web sémantiques s'appuient en général sur les langages du Web sémantique pour décrire leurs fonctionnalités et les données qu'ils échangent. Les motivations pour développer, ou tendre vers les Services Web sémantiques sont évidemment de faciliter les phases automatiques de découverte, sélection et composition de Services Web. En effet, si leur sémantique est connue, alors chercher et composer des Services Web pourra être fait automatiquement en donnant la sémantique cible.

## 1.6 Ontologies pour le Web sémantique

L'ontologie joue un rôle très important pour le Web sémantique parce qu'elle représente la sémantique des documents en permettant leur exploitation par les applications et les agents intelligents. Elle est très utile pour structurer et définir la signification des termes de métadonnées actuellement collectées et normalisées. Donc, à l'aide des ontologies, les applications sur le Web de demain pourront devenir intelligentes, au sens où elles pourront opérer plus précisément au niveau conceptuel humain.

Une des définitions de l'ontologie qui fait autorité est celle de Gruber [17] : « **Une ontologie est une spécification formelle et explicite d'une conceptualisation partagée** ». La *conceptualisation* représente la collection des objets, de concepts et des autres entités qui sont supposés exister dans un certain domaine d'intérêt, et les relations qui les relient. Une conceptualisation est une vue abstraite, simplifiée du monde que l'on veut représenter. *Explicite* signifie que le type des concepts utilisés et les contraintes sur leur utilisation sont explicitement définis. *Formelle* se réfère au fait que l'ontologie doit être compréhensible par les machines. *Partagée* reflète la notion de connaissance consensuelle décrite par l'ontologie, c'est-à-dire qu'elle n'est pas restreinte au point de vue de certains individus seulement, mais reflète un point de vue plus général, partagé et accepté par un groupe.

Une ontologie est une structure de donnée opérationnelle qui rend compte des concepts d'un domaine et de leurs relations. Le développement des ontologies croissant en Intelligence Artificielle vient de leur intérêt pour associer du sens à des ressources textuelles, pour localiser et gérer des connaissances dans diverses applications. Son but est donc de définir quelles primitives avec leur sémantique associée sont nécessaires pour la représentation des connaissances dans un contexte donné. Elle est représentée par un graphe orienté qui contient :

- Les nœuds (concepts) représentant le vocabulaire d'un domaine particulier.
- Les arcs représentant les relations (ou rôles) nommées entre les concepts.

La figure suivante montre un exemple d'ontologie pour les images pneumologiques :

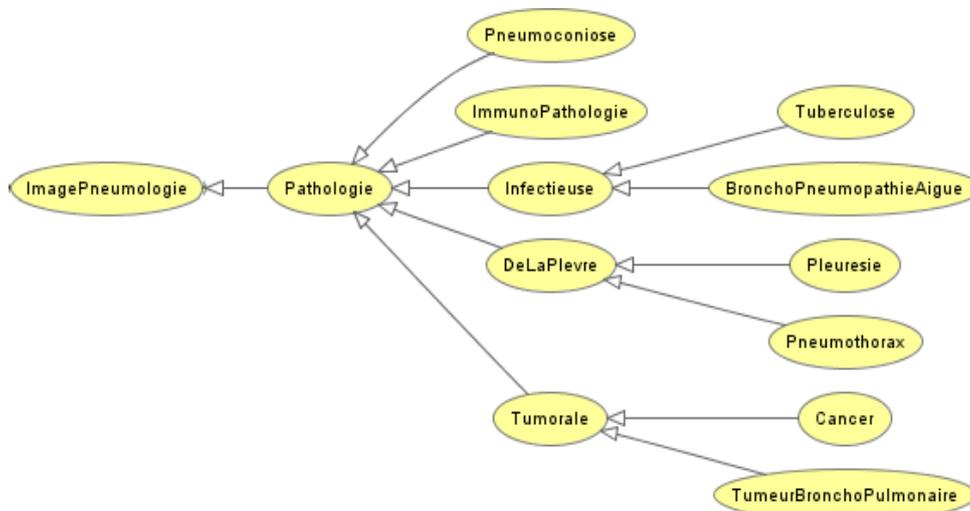


FIGURE 1.2 – Exemple d’une partie d’ontologie des images médicales  
(images pneumologiques) [1]

A partir de cette structure, la sémantique de chaque mot est déduite par les relations que ce mot possède dans l’ontologie, ce qui permet de restreindre les interprétations possibles. Cependant pour être exploitable par une machine, une ontologie doit respecter certaines règles :

- Etre définie par une syntaxe formelle et une sémantique non ambiguë.
- Permettre la déduction de nouvelles connaissances qui sont présentées implicitement.

De plus, pour assurer sa pérennité, une ontologie doit posséder un niveau d’abstraction permettant son extension [15].

## 1.7 RDF/RDFS

1. « **RDF (Resource Description Framework) est un modèle standard pour l’échange de données sur le Web. Il permet de décrire de façon formelle les ressources Web et leurs métadonnées** » [21]. En utilisant RDF, les ressources sont décrites par un ensemble de déclarations RDF sous forme de triplets  $\langle \text{Sujet}, \text{Propriété}, \text{Objet} \rangle$  où :

- *Sujet* : représente la ressource à décrire. Il est identifié par une URI.
- *Propriété ou prédicat* : il s'agit d'une propriété utilisée pour caractériser et décrire une ressource. Une propriété est une liaison étiquetée et orientée du sujet vers l'objet. Elle est identifiée par une URI.
- *Objet* : la valeur de la propriété pouvant être un littéral ou bien une autre ressource (identifiée par une URI).

Cet ensemble de triplets RDF peut être représenté par un graphe orienté étiqueté (Figure 1.3) où les éléments apparaissant comme sujet ou objet sont des nœuds et les propriétés sont des arcs.

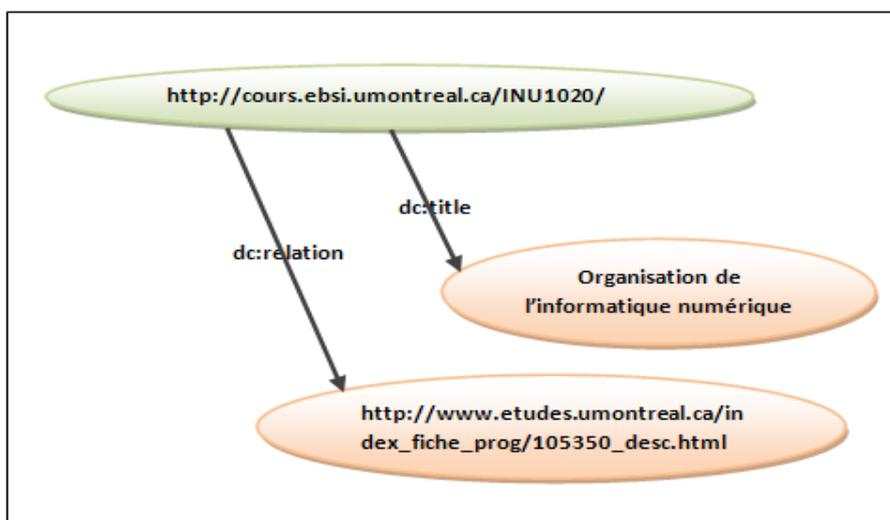


FIGURE 1.3 – Exemple de graphe RDF.

Les nœuds peuvent se présenter sous la forme :

- **Valeur littérale** : Ex : Michel Tremblay.
- **URI**.
- **Vide** : permet de désigner une ressource par ses propriétés sans expliciter cette ressource.

Les arcs peuvent se présenter sous la forme :

- **URI** :
  - **Forme longue** : `http://purl.org/dc/elements/1.1/title`
  - **Forme raccourcie** : faisant appel à un préfixe `dc:title`

---

A noter : un préfixe c'est un raccourci de l'URI complet de la ressource. Ecrire `dc:title` équivaut à écrire l'URI de la propriété dans sa forme longue.

## 2. RDFS (Resource Description Framework Schema)

Il est nécessaire, pour donner une sémantique ou un sens aux informations stockées sous forme de triplets RDF, de se donner un vocabulaire (classes et propriétés).

« RDFS est un langage extensible de représentation des connaissances qui permet de décrire précisément par vocabulaires les ressources d'un domaine donné et les relations entre elles. RDFS fournit des éléments de base pour la définition d'ontologies ou vocabulaires destinés à structurer des ressources RDF » [20].

Il permet d'organiser les classes en une hiérarchie de classes en utilisant les relations de subsomption entre classes "subClassOf" et d'organiser les propriétés en une hiérarchie de propriétés en utilisant des relations de subsomption entre propriétés "subPropertyOf". De plus, RDFS offre le moyen de spécifier le typage des propriétés en indiquant leur "domaine" et leur "co-domaine" (range).

## 1.8 SPARQL

« SPARQL est un langage de requêtes pour l'interrogation de métadonnées et l'extraction des données sous forme d'un graphe RDF ou plus exactement un langage d'interrogation de triplets RDF » [22].

SPARQL est adapté à la structure spécifique des graphes RDF et s'appuie sur les triplets qui les constituent. Il définit la syntaxe et la sémantique nécessaire à l'expression de requêtes sur une base de données de type RDF. Il est différent du classique SQL (langage de requête qui est adapté aux bases de données relationnelles) mais s'en inspire clairement dans sa syntaxe et ses fonctionnalités. Il a aussi quelques traits de ressemblances mineures avec Prolog. SPARQL permet d'exprimer des requêtes interrogatives ou constructives :

- Une requête SELECT de type interrogative permet d'extraire du graphe RDF un sous-graphe correspondant à un ensemble de ressources vérifiant les conditions dé-

finies dans une clause WHERE.

- Une requête CONSTRUCT de type constructive, engendre un nouveau graphe qui complète le graphe interrogé.

La structure d'une requête SPARQL est très similaire à celle employée dans le langage SQL :

```
SELECT ?v1?v2...?vn
FROM <description.rdf>
WHERE {
(sujet1 |vi) (predicat1 | vj) (objet1 | vk) ... (sujetx |va) (predicaty | vb) (objetz | vc)
}
```

## 1.9 Conclusion

Les Services Web représentent aujourd'hui la technologie la plus adaptée pour assurer le développement des systèmes distribués sur Internet. Un des concepts intéressants qu'offre la technologie de Service Web et qui suscite beaucoup d'intérêt est la possibilité de créer un nouveau Service Web à valeur ajoutée par composition de plusieurs Services Web existants.