

Ingénierie des connaissances

Introduction aux ontologies

**Département d'Informatique
Université Batna 2
Algérie**



Introduction

Ontologies et Web sémantique



☐ Introduction

ontologie: Définition

En informatique, une ontologie désigne un ensemble structuré de savoirs dans un domaine de connaissance particulier (une certaine vue du monde par rapport à un domaine donné).

Cette vue est souvent conçue comme un ensemble de concepts : entités, attributs, processus, leurs définitions et leurs interrelations:

‘Une conceptualisation’

- organisation hiérarchique des concepts pertinents;
- relations qui existent entre ces concepts;
- des règles et axiomes.

ontologies: But

- **modéliser** un ensemble de **connaissances** dans un domaine donné, qui peut être réel ou imaginaire.
- Destinée à des **traitements automatiques**

Les ontologies doivent être :

- **normatives**
- **cohérentes**
- **partageables**
- **Réutilisables**

ontologies: Types

- Ontologie **générique** (*Top ontology*)

Concepts de haut niveau (existant dans plusieurs domaines)
(ontologie réutilisable)

- Ontologie **de représentation**

Primitives de représentation (ex. Ontolingua, définition des éléments RDF)

- Ontologie **du domaine**

Concepts restreints à un domaine (ex. *Gene Ontology*)

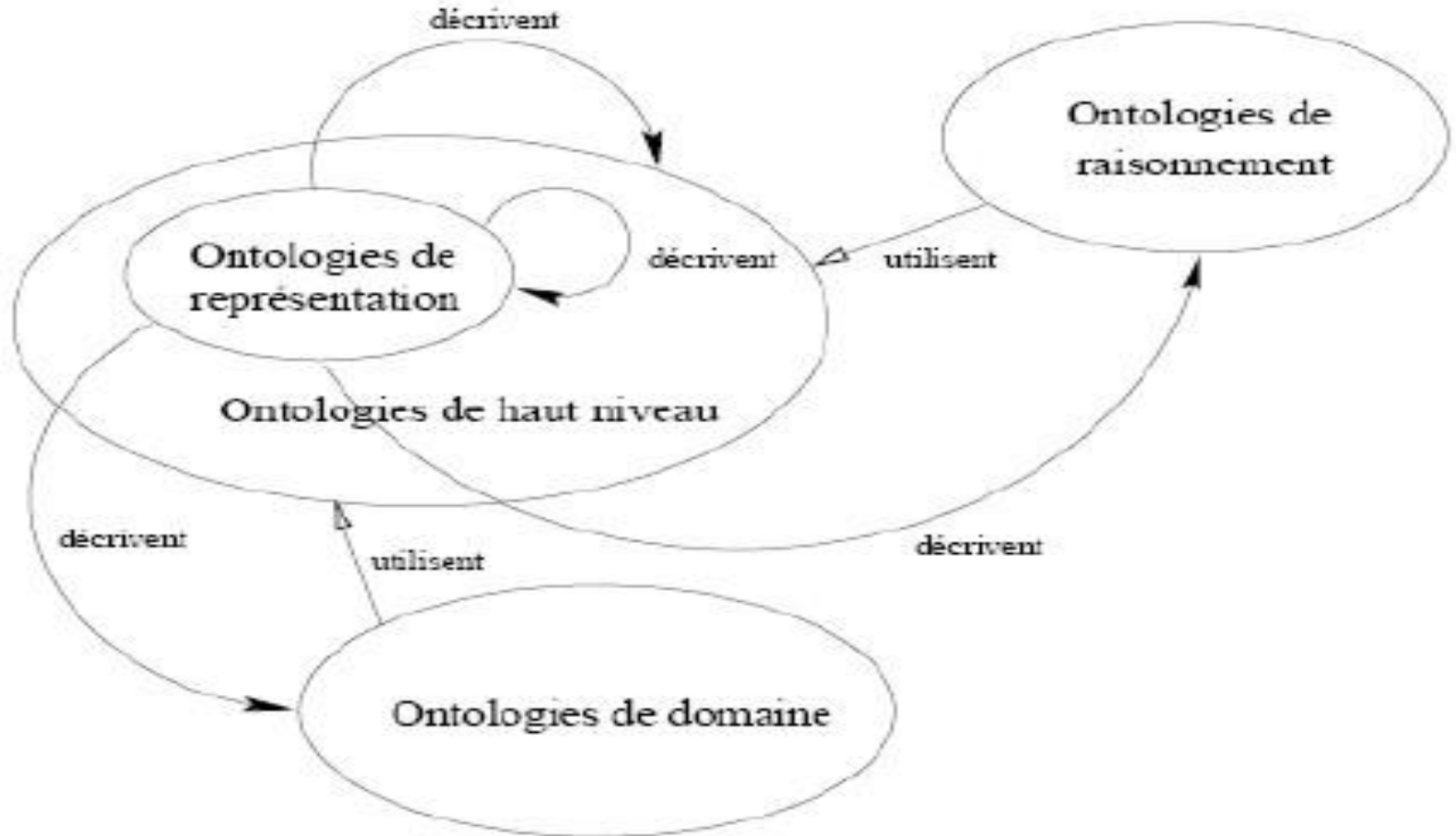
- Ontologie **d'une méthode de résolution de problème**

Concepts impliqués dans une méthode (conception, diagnostic, évaluation)

- Ontologie **d'application**

dédiée à un champ d'application précis à l'intérieur d'un domaine

ontologies: Types



[MIZOGUCHI, 1997]

Que représenter dans une ontologie?

➤ Concepts

Ensemble d'objets (individus) avec leurs propriétés:

- Décrit par un terme
- Définit en intension ou en extension
- Organisés en taxionomie

1- *Concepts primitifs* :

Humain, male, femelle

2- *Concepts définis* :

Homme : Humain & Male

Femme : Humain & Femelle

Patient : Humain(avoir Maladie)

Que représenter dans une ontologie?

➤ Individus :

Instances de concepts

Exemples :

Samir, Rima, l'ordinateur de Yasmine, ... etc.

➤ Les propriétés des concepts

Attributs simples

● Exemples :

● âge, nom, hauteur, nombre de roue, ... etc.

Que représenter dans une ontologie?

➤ Les relations sémantiques (rôles)

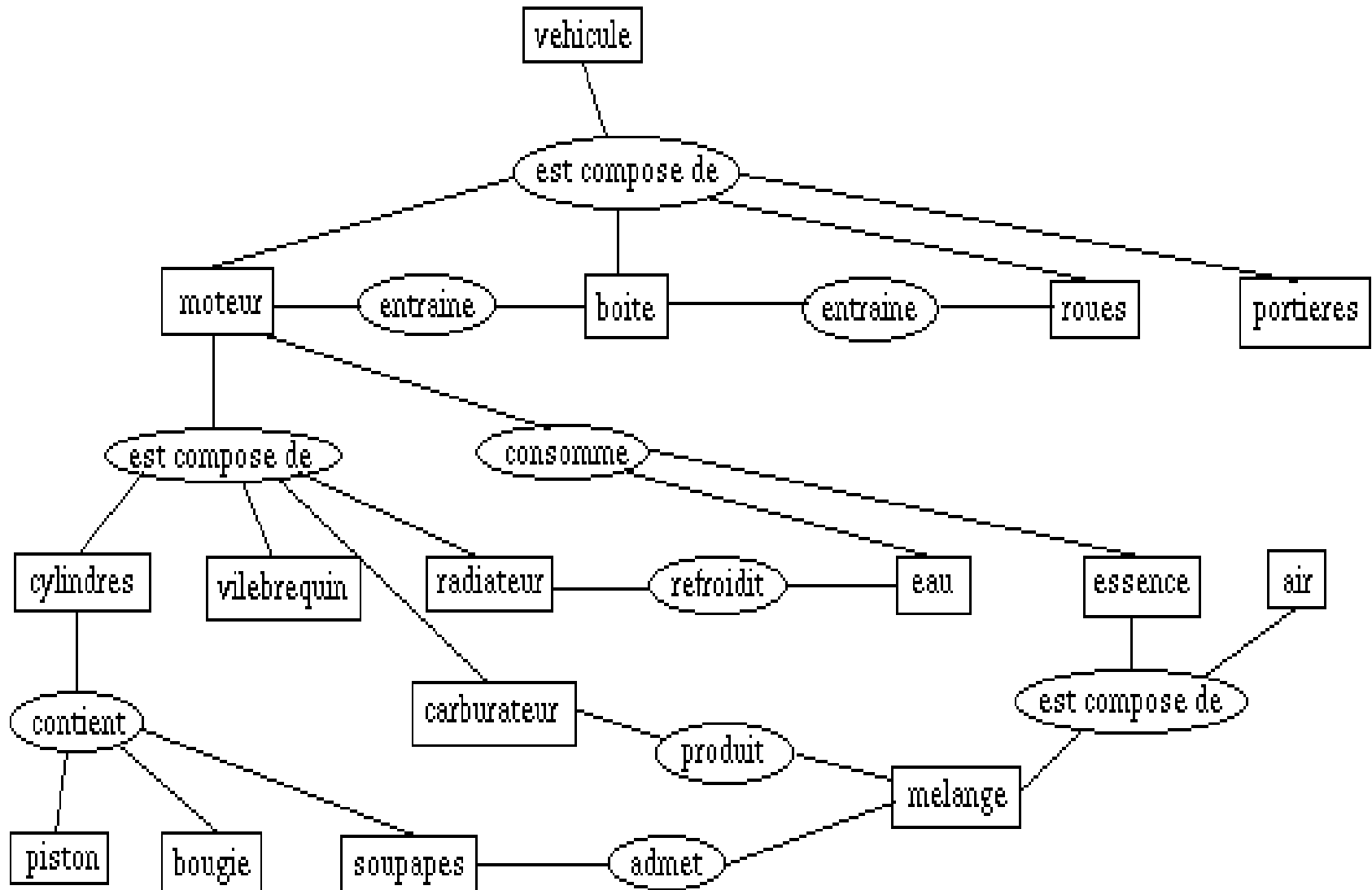
Association entre concepts

Exemples : est-un, parent-de, partie-de, contient, proche-de, dessus, dessous, ... etc.

➤ Propriétés sur les relations :

- Transitivité (est-un, partie-de)
- Symétrie : proche-de
- Inverse : dessus / dessous
-

Vehicule ontology



Représentation formelle d'une ontologie

- les langages d'échange d'ontologies sur le Web, dont la syntaxe est basée sur le langage XML.
- les langages opérationnels qui implémentent les ontologies à des fins d'inférences, pour constituer un composant d'un système d'information.

(ex. **DefOnto**: défini et implanté au LaRIA (Laboratoire de Recherche en informatique d'Amiens) de l'Université de Picardie Jules Verne)

Principes pour la construction d'ontologies

➤ Clarté

Les ambiguïtés doivent être réduites.

➤ Cohérence

La cohérence des définitions en langage naturel doit être vérifiée.

➤ Extensibilité

L'ontologie doit être construite de telle manière que l'on puisse l'étendre facilement, sans remettre en cause ce qui a déjà été fait.

➤ Indépendance

L'ontologie doit être conceptualisée indépendamment de tout langage d'implémentation.

➤ Engagement ontologique minimal

Une ontologie doit faire un minimum d'hypothèses sur le monde.


Exemples d'ontologies

- WordNet : Ontologie de la langue naturelle
- Enterprise Ontology
- Kactus : ingénierie
- UMLS : Unified Medical Language System
- EngMath, PhysSys
- O'Comma
- OMNIBUS

- Protégé ontology library (protege.stanford.edu/ontologies.html)
- DAML ontology library (www.daml.org/ontologies)
- Ontolingua ontology library (www.ksl.stanford.edu/software/ontolingua/)

Cycle de vie d'une ontologie

Le cycle de vie des ontologies est inspiré du génie logiciel, Il comprend les étapes suivantes:

- Evaluation des besoins;
- Construction; → 
- Diffusion;
- Utilisation.

- La *conceptualisation* : identification des connaissances contenues dans un corpus représentatif du domaine.
(travail mené par un expert du domaine, assisté par un ingénieur de la connaissance).

Cycle de vie d'une ontologie

L'ontologisation : formalisation, autant que possible, du modèle conceptuel obtenu à l'étape précédente.

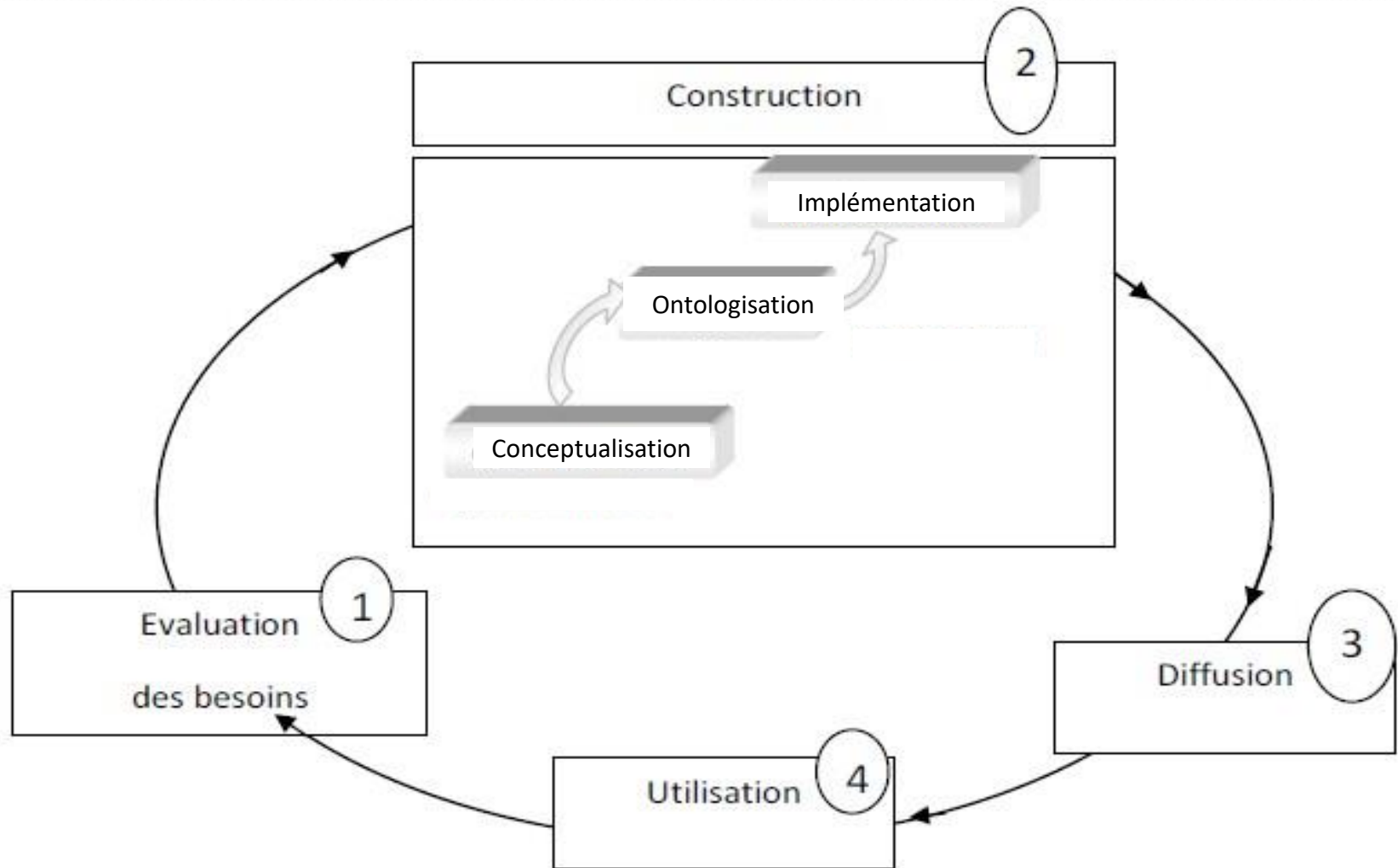
(travail mené par l'ingénieur de la connaissance assisté de l'expert du domaine).

L'implémentation : transcription de l'ontologie dans un langage opérationnel de représentation de connaissances.

(travail mené par l'ingénieur de la connaissance).

- ❖ De nombreux allers-retours sont nécessaires pour bâtir une ontologie opérationnelle adaptée aux besoins.

Cycle de vie d'une ontologie



[Furst, 2002]

Outils

- **DOE (DIFFERENTIAL ONTOLOGY EDITOR)** est un éditeur d'ontologie.
- **PROTEGE2000** : édition, contrôle, visualisation et extraction d'ontologie à partir de textes.
- **ONTOLINGUA**: une extension du langage KIF (KNOWLEDGE INTERCHANGE FORMAT):
 - Possibilité de fusionner des ontologies
 - L'ontologie est immédiatement représentée dans un formalisme.

ontologies et Web sémantique

Quels domaines utilisent les ontologies?

Utilisées dans:

- L'intelligence artificielle
- le web sémantique;
- le génie logiciel;
- L'informatique biomédicale;
-

Web sémantique

Le Web Sémantique n'est pas un Web distinct mais bien un prolongement du Web que l'on connaît, dans lequel on attribue à l'information une signification clairement définie, ce qui permet aux ordinateurs et aux humains de travailler en plus étroite collaboration. Très bientôt, ces développements permettront une nouvelle fonctionnalité car les ordinateurs seront davantage en mesure de traiter et de 'comprendre' les données que, pour l'instant, ils ne font qu'afficher.

Selon Tim Berners-Lee (T. Berners-Lee, J. Hendler et O.Lassila, Scientific American 2001)

Web

vs

Web sémantique

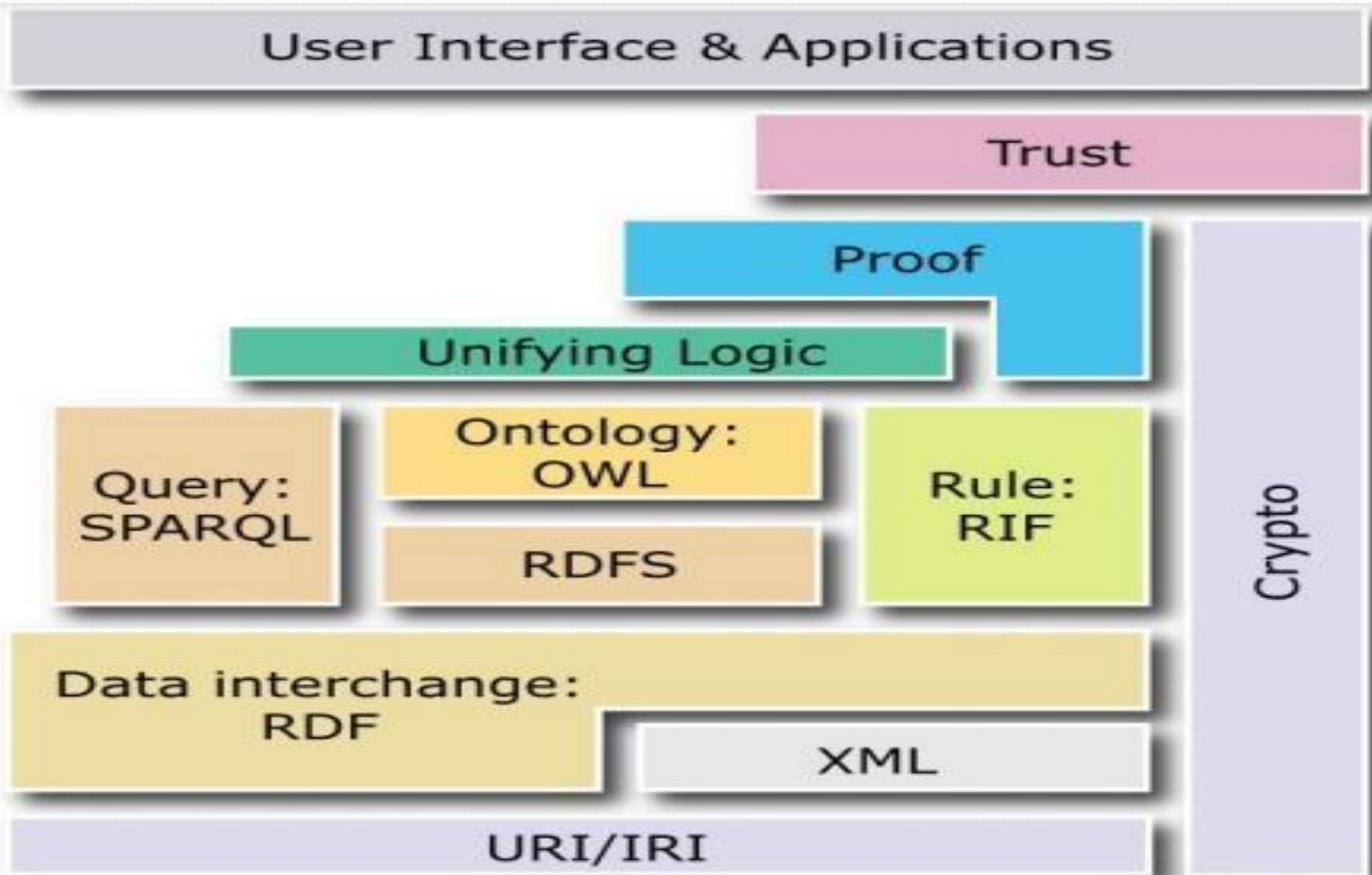
- Ensemble de **documents**;
 - Basé essentiellement sur **HTML**(langage de présentation de données);
 - Recherche par **mots clé**;
 - Utilisable par **l'humain**.
- Ensemble de **connaissances**;
 - Basé sur **XML** et **RDF(S)**;
 - Recherche par **concepts**;
 - Utilisable par **la machine**.

Idée: permettre aux machines de:

- Comprendre les données;
- Tirer de nouvelles conclusions.

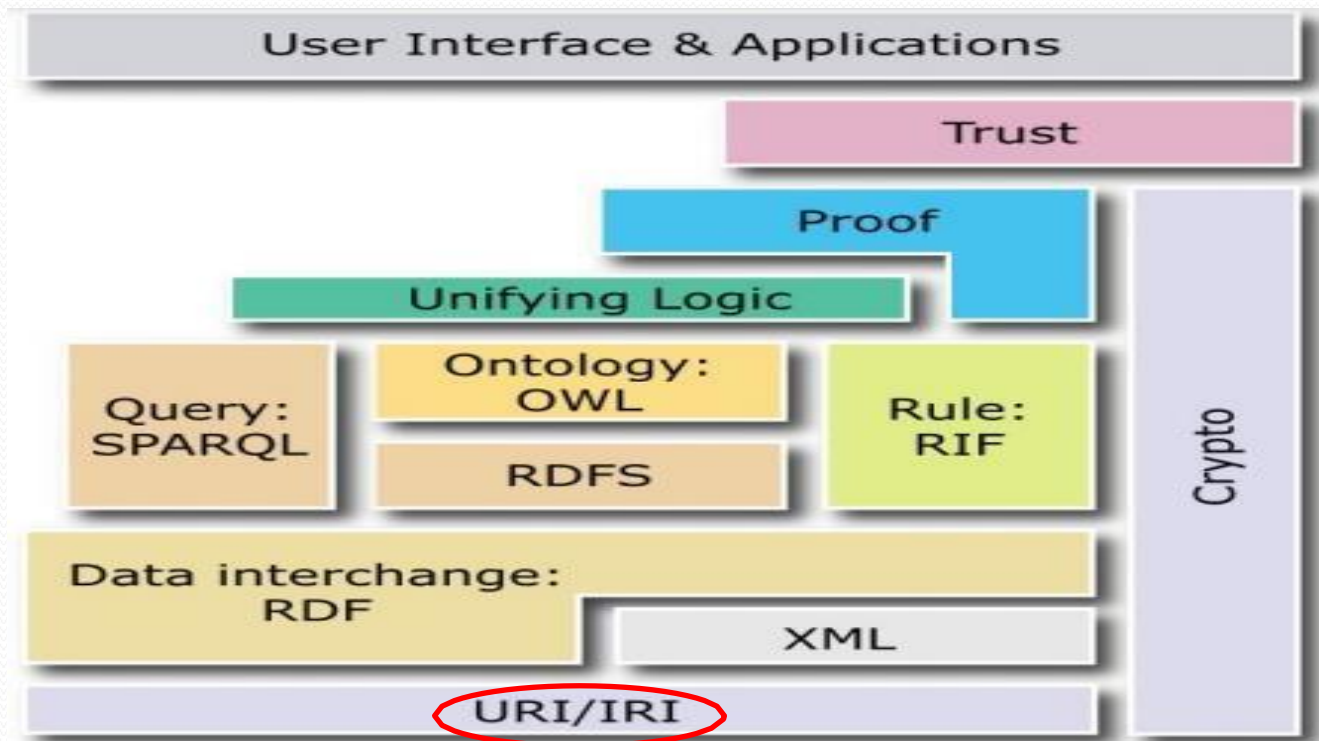
Les agents logiciels seront capables de : **rechercher**, **filtrer** et **afficher** l'information correspondant au besoin de l'utilisateur.

Couches du Web sémantique



Les différentes couches

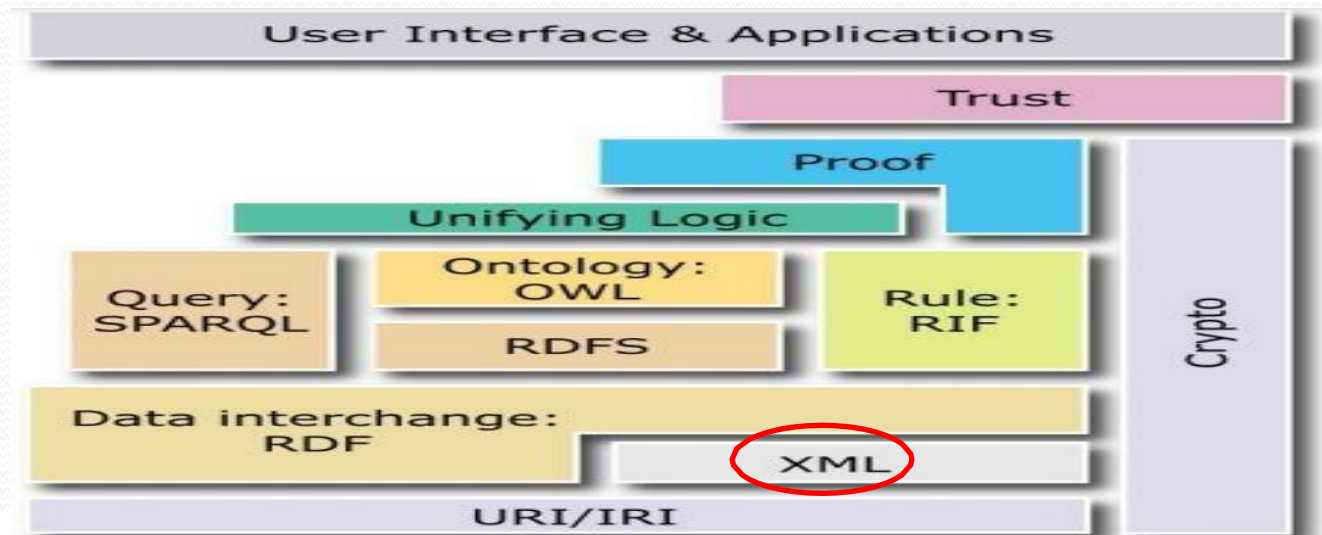
- **URI** : (Uniforme Ressource Identifier) fournit un adressage standard universel permettant d'identifier les ressources.



Les différentes couches

➤ XML (eXtensible Market Language)

Langage de balisage extensible : un métalangage qui nous permet de définir nos propres balises pour nos documents.



Exemple:

```
<livre>
  <language> es <language>
  <title> Cronicas de Bustos Domecq <title>
  <auteur> Jorge Luis Borges <auteur>
  <auteur> Adolfo Bioy Casares <auteur>
  <publisher> Editorial Losada <publisher>
  <date> 1967 <date>
  <ISBN> 0525475486 <ISBN>
</livre>
```

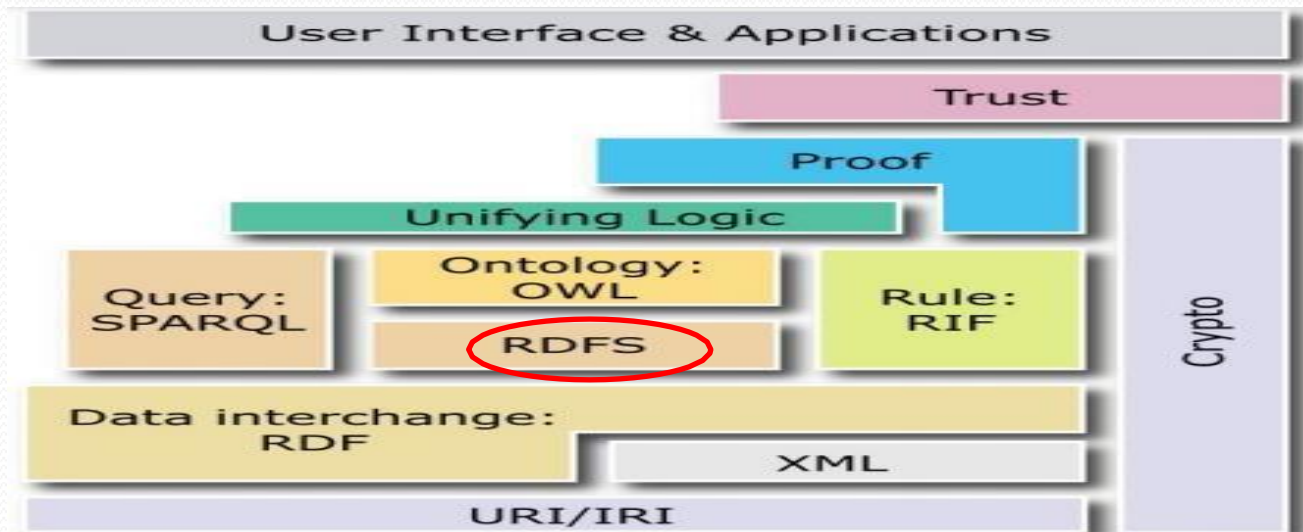
Les différentes couches

➤ RDF (Resource Description Framework)

Modèle de données pour décrire des ressources du web.

Graphe: représenté par un ensemble d'énoncés

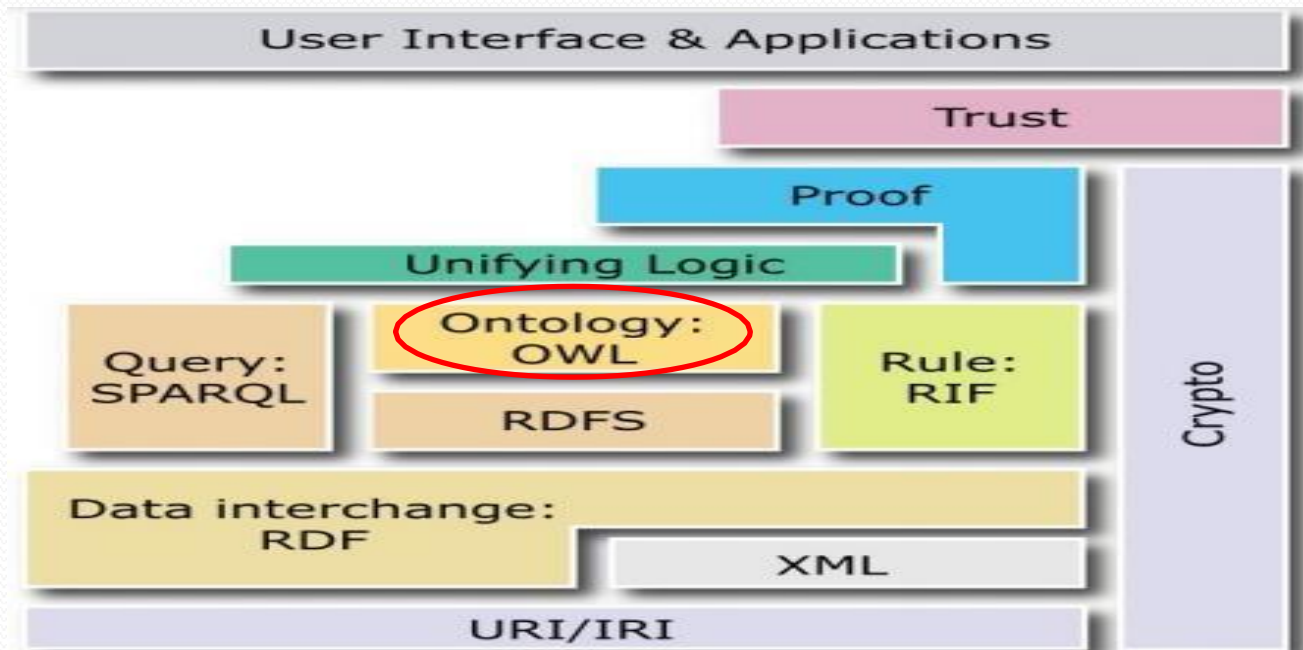
(statements): les noeuds représentent des ressources et les arcs représentent des relations entre ces ressources



Les différentes couches

➤ Ontologie :

Permet de définir une **sémantique** et un **vocabulaire** commun des données utilisées dans le web, ce qui évite tous conflits de partage et d'intégration de ces données.

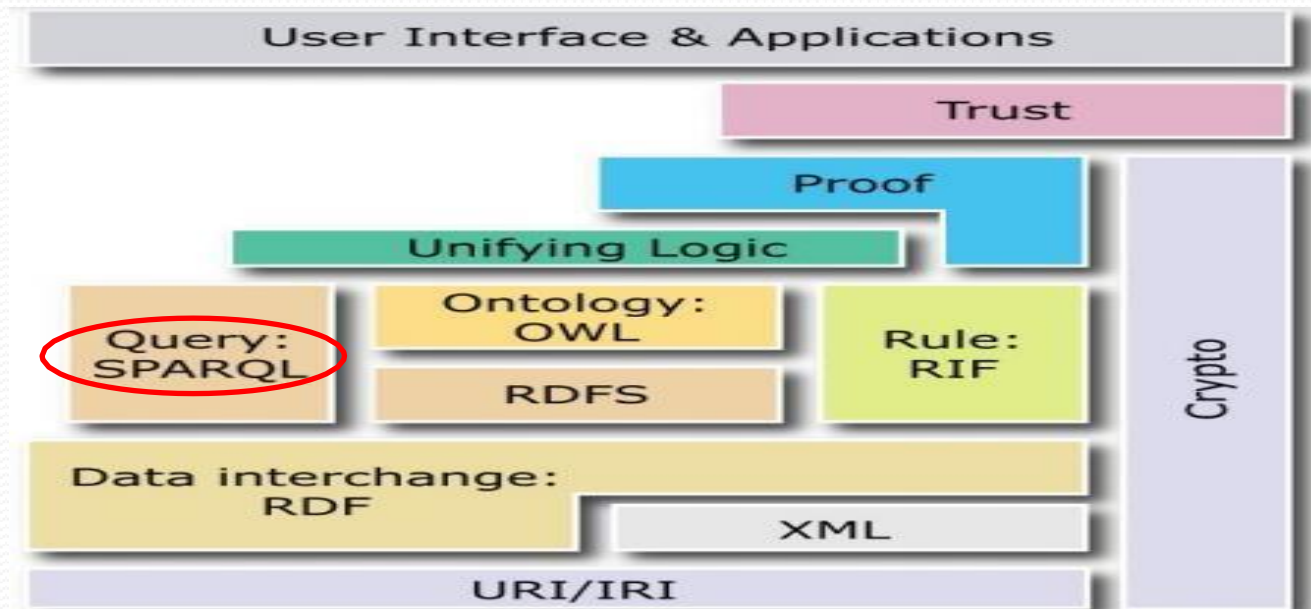


Les différentes couches

➤ SPARQL:

Comment accéder à une BD en RDF?

Nécessité d'un langage de requête

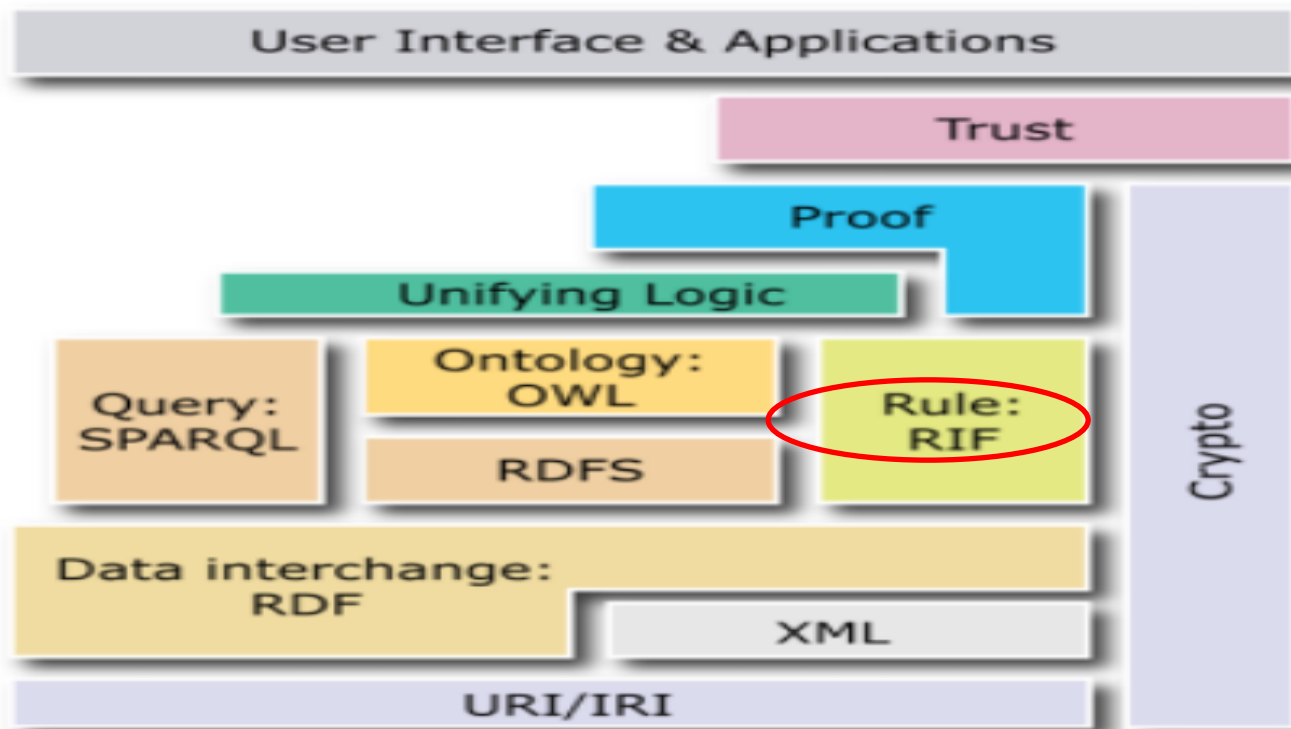


W3C propose le langage **SPARQL**: un langage de requête basé graphes

Les différentes couches

➤ RIF:

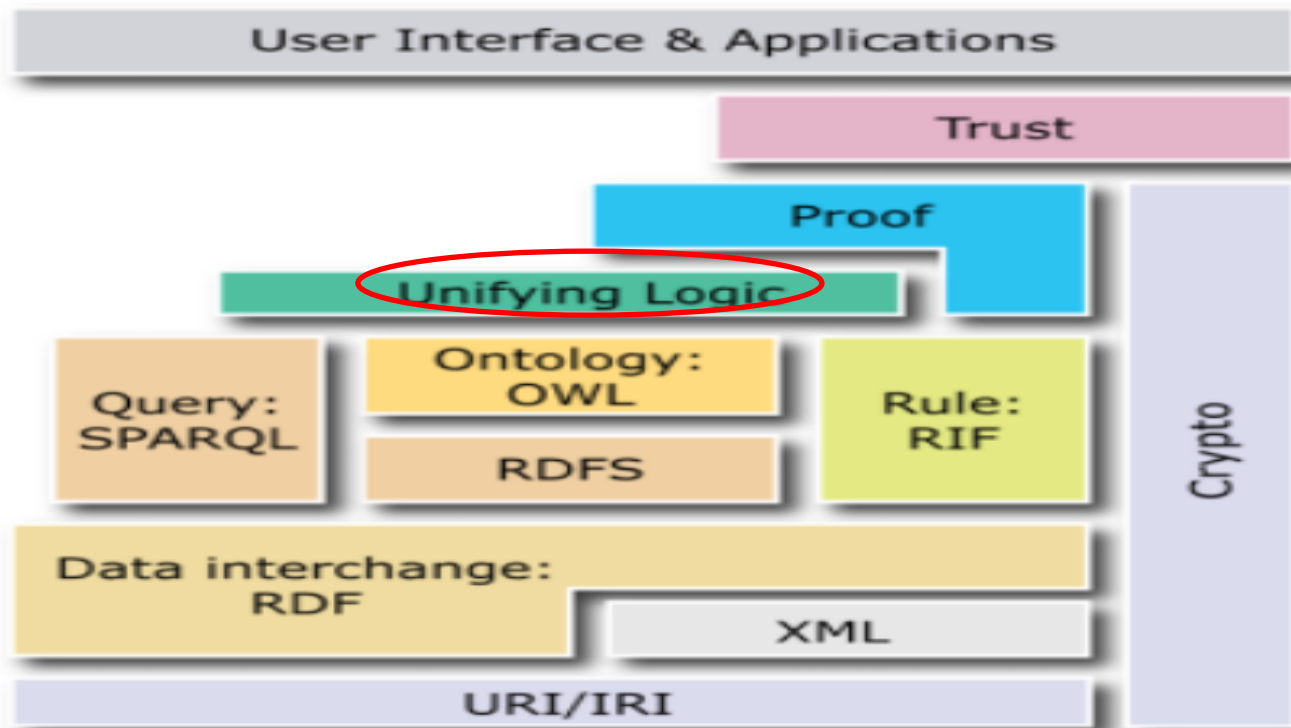
un langage de règles qui a pour objectif d'appliquer des conditions des échanges de données.



Les différentes couches

➤ Unifying Logic:

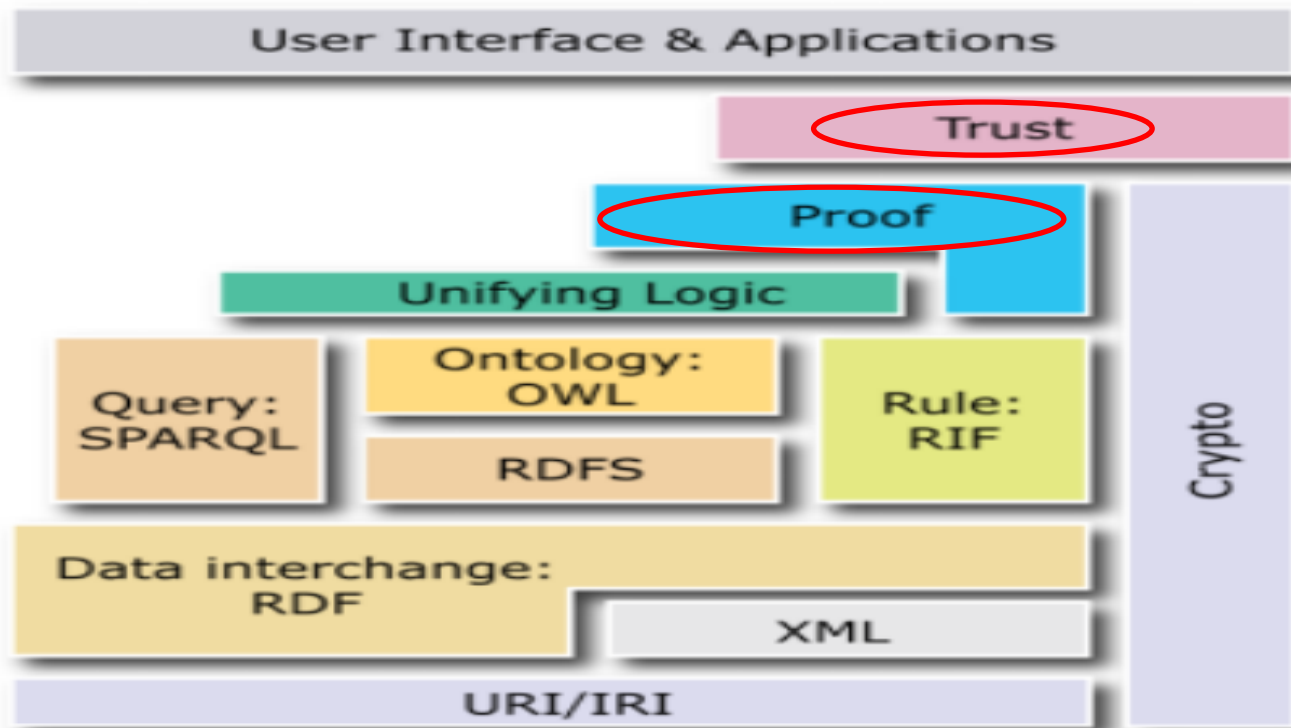
Permet l'interopérabilité des informations entre les systèmes.



Les différentes couches

➤ Proof/ Trust:

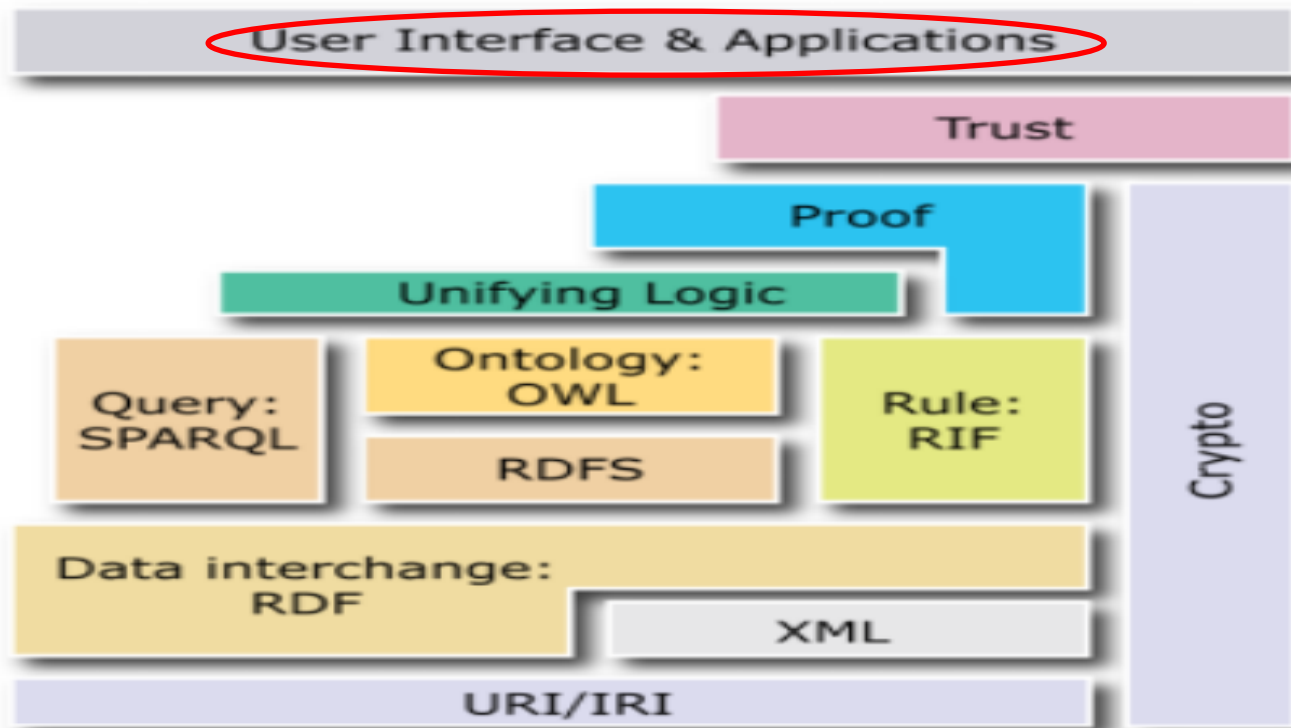
S'assurer de la véracité et de la fiabilité de sources fournissant des données.



Les différentes couches

➤ User Interface:

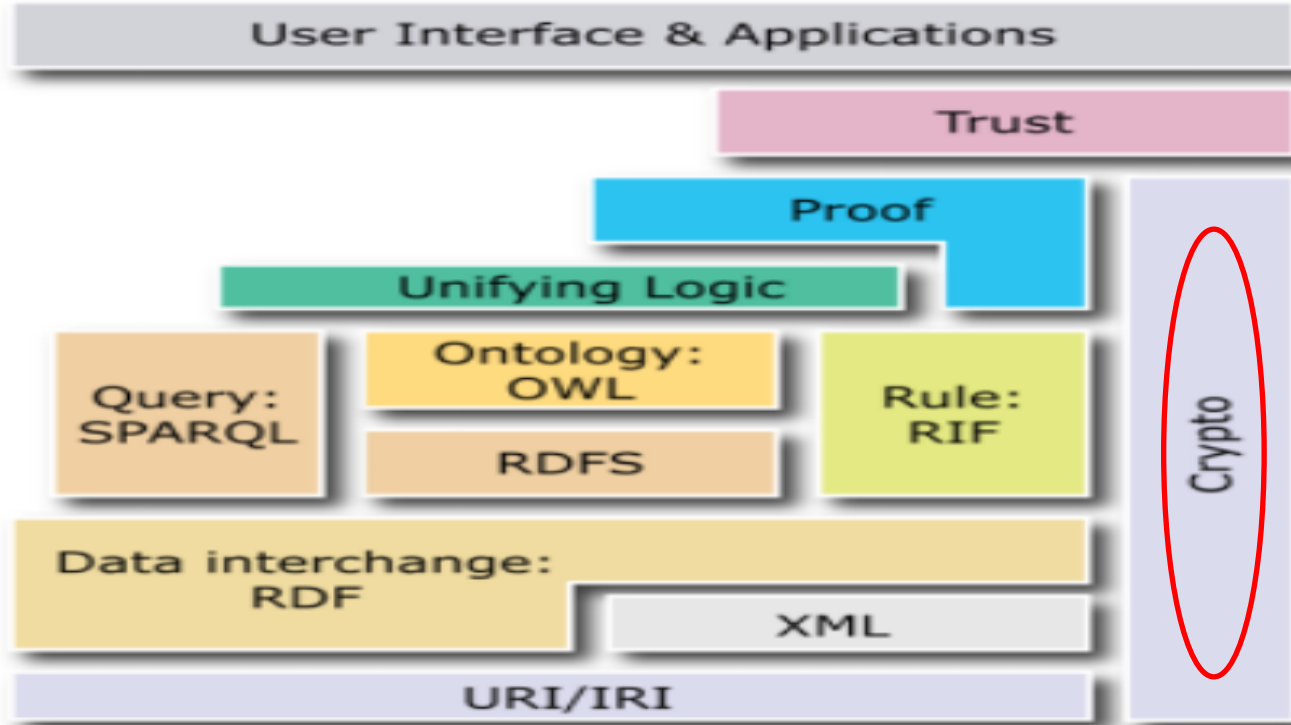
Permet l'utilisation des applications utilisant le Web Sémantique.



Les différentes couches

➤ Crypto:

Veiller à la sécurité et à l'intégrité des données.



XML: Exemples

```
<?xml version="1.0" encoding="UTF-8"?>
<personne>
  <nom>Dupond</nom>
  <prenom>Jean</prenom>
  <naissance>
    <lieu>
      <ville>Paris</ville>
      <pays>France</pays>
    </lieu>
    <date>
      <jour>14</jour>
      <mois>7</mois>
      <annee>1789</annee>
    </date>
  </naissance>
</personne>
```

racine

XML: Evolutions...

- XML (1998) a donné un cadre à la **structuration des connaissances**
- Création de nouveaux langages web destinés non plus à un rendu graphique à l'écran pour un utilisateur humain, mais à **un réel partage et à une manipulation des savoirs**
- RDF (1999): un langage XML permettant de **décrire des métadonnées** et facilitant **leur traitement**

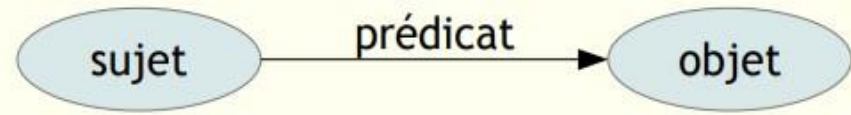
RDF: Objectifs

- permettre aux informations d'être **manipulées en dehors de l'environnement particulier** dans lequel elles ont été créées, éventuellement à l'échelle d'Internet. (**Interopérabilité** des savoirs).
- **optimiser la coopération** entre applications, en permettant de combiner les données de plusieurs applications, pour générer de nouvelles informations.
- faciliter le **traitement automatique** de l'information du Web **par des agents logiciels**.

RDF, RDF/XML

RDF

La structure fondamentale de toute expression en RDF est **une collection de triplets**,



chacun composé d'un **sujet**, un **prédicat** et un **objet**.

L'ensemble est appelé un **graphe RDF**

Chaque triplet représente l'existence d'une relation entre les choses symbolisées par les noeuds qui sont joints.

RDF/XML

RDF/XML (Recommandation du W3C de 1998, et mise à jour en 2004 [W3C 2004a]) propose **une syntaxe de sérialisation** de RDF. RDF/XML s'appuie sur XML par sa syntaxe.

OWL: Plus riche...

OWL (Web Ontology Language)

En plus d'apporter à l'utilisateur la capacité de décrire des classes (ie. avec des constructeurs) et des propriétés, OWL intègre **des outils de comparaison des propriétés et des classes** : identité, équivalence, contraire, cardinalité, symétrie, transitivité, disjonction, etc.