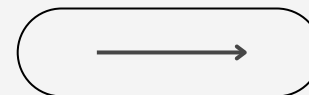




WEB SÉMANTIQUE

Présenté par:
CHAKRANE Ismail
LIROULET Inès
SENE Massamba
YAKOUBI Oumayma



PLAN

01

INTRODUCTION

Histoire du web et problématique

02

FONCTIONNEMENT DU WEB SÉMANTIQUE

Comment le web sémantique fonctionne-t-il?

03

VISIONS ET APPLICATIONS

Quelles visions du web sémantique existent et quelles en sont les applications ?

04

ÉVOLUTION

Comment le web sémantique a-t-il évolué au cours du temps ?

05

DÉFIS ET LIMITES

Quels sont les défis rencontrés par le web sémantique ?

06

LIEN AVEC LA COGNITION

Quel est le positionnement du web sémantique par rapport aux approches de la cognition ?

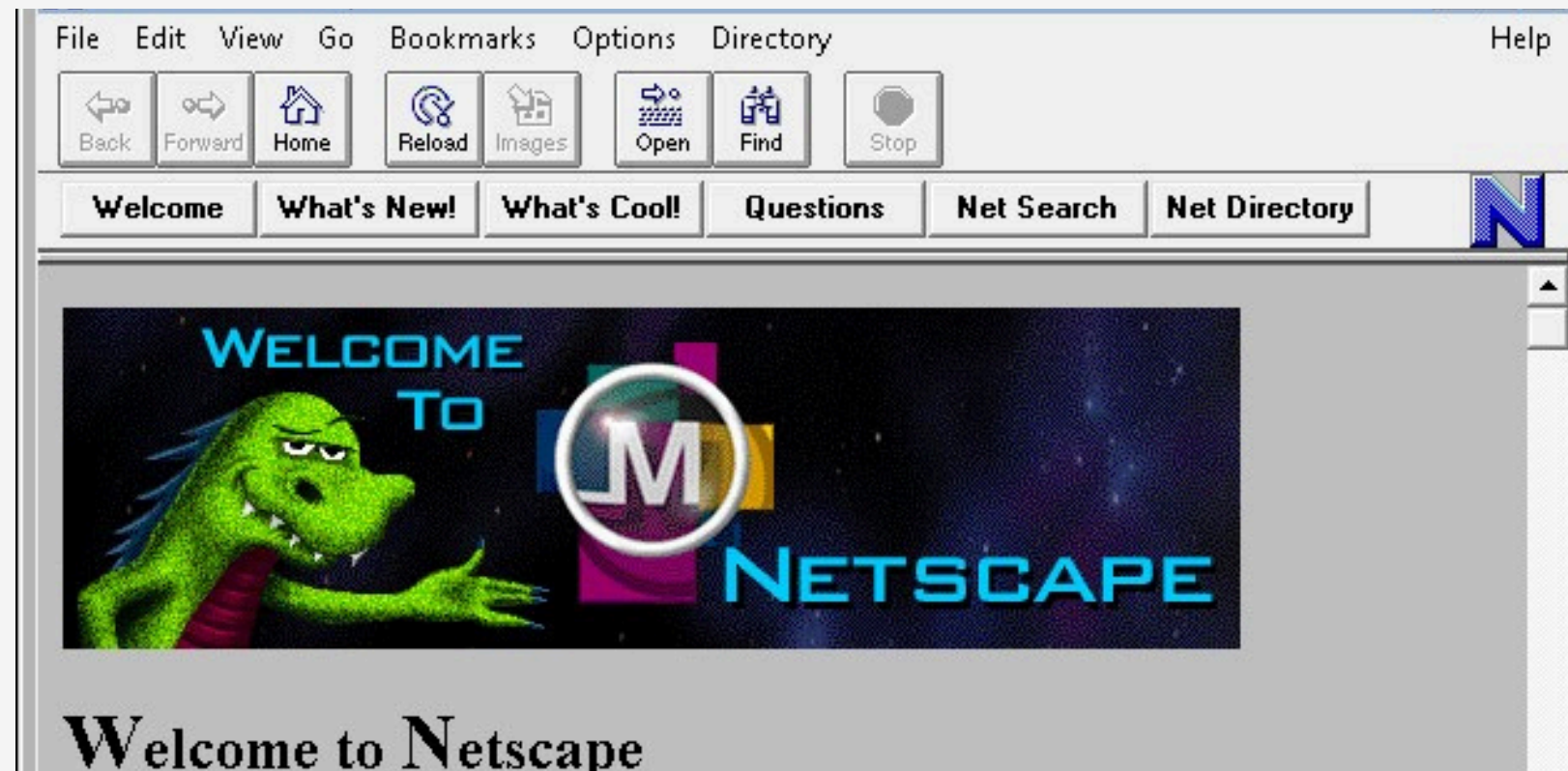
HISTOIRE DU WEB SÉMANTIQUE

1.0

WEB 1.0 OU WEB TRADITIONNEL (ANNÉES 90)

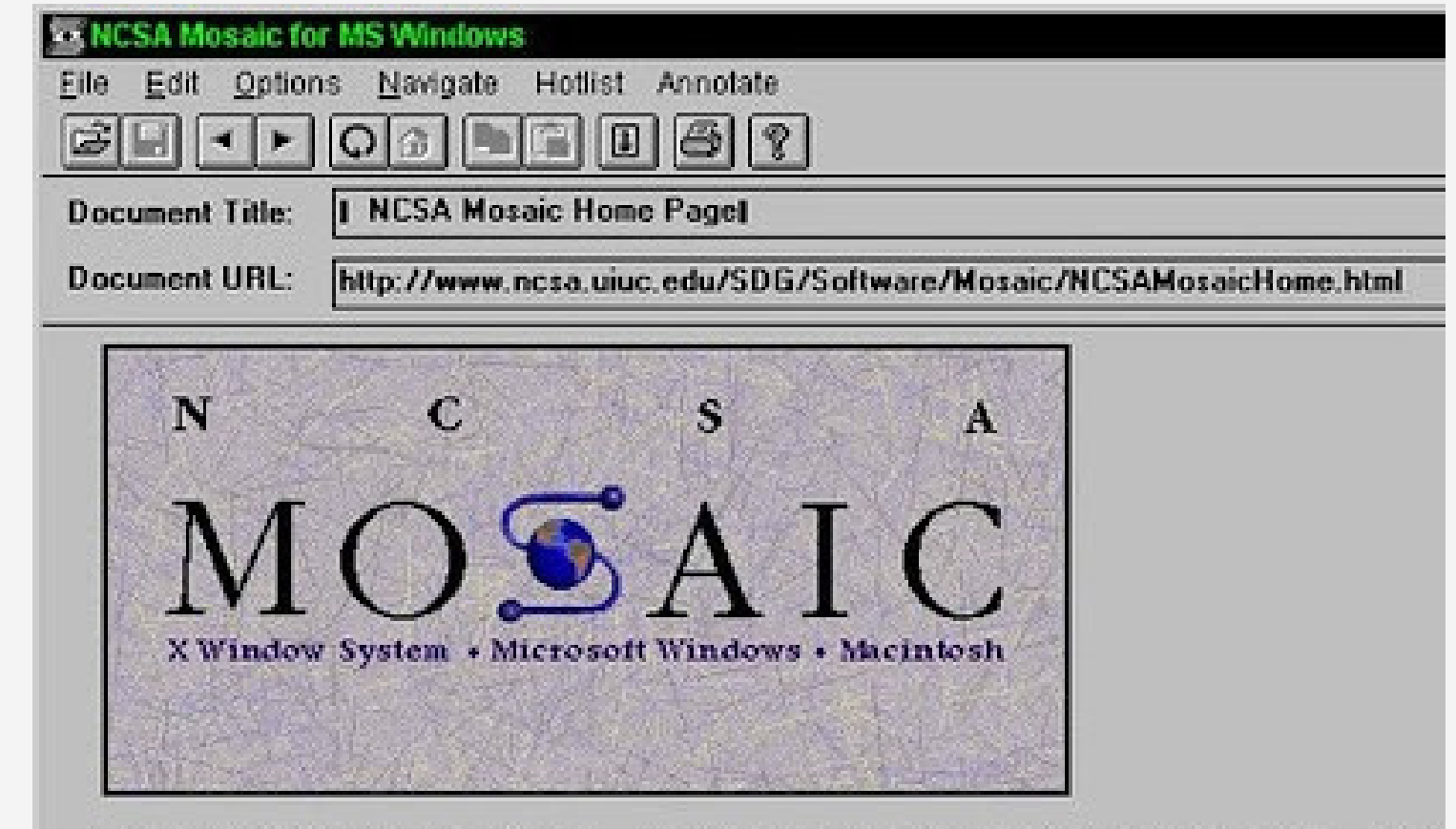
Le **World Wide Web** a été créé en **1989** par **Tim Berners Lee**, à l'origine comme un moyen de partager des informations scientifiques à travers le monde. Le début du web était caractérisé par des pages statiques avec peu d'interactivité. Les sites web étaient essentiellement constitués de pages HTML statique, conçues par des développeurs professionnels, et le contenu était unidirectionnel: les informations étaient publiées en ligne, mais les utilisateurs ne pouvaient pas y contribuer.

Année 1994



source : <https://blog.mozfr.org/post/2017/08/22-ans-la-folie-Netscape>

Année 1993



source : <https://www.kaliop.com/fr/il-etait-une-fois-les-navigateurs/>

L'expérience utilisateur était limitée à la lecture des pages, sans possibilité d'interaction ou personnalisation. En 1990, les premiers navigateurs web, tels que Mosaic et Netscape, ont vu le jour, permettant aux utilisateurs d'accéder aux sites. Cependant, ces sites étaient principalement des répertoires d'informations à sens unique, ce qui limitait les possibilités de communication en ligne.

HISTOIRE DU WEB SÉMANTIQUE

Février 2004



source : <https://www.europe1.fr/technologies/En-2004-Facebook-ressemblait-a-ca-643360>

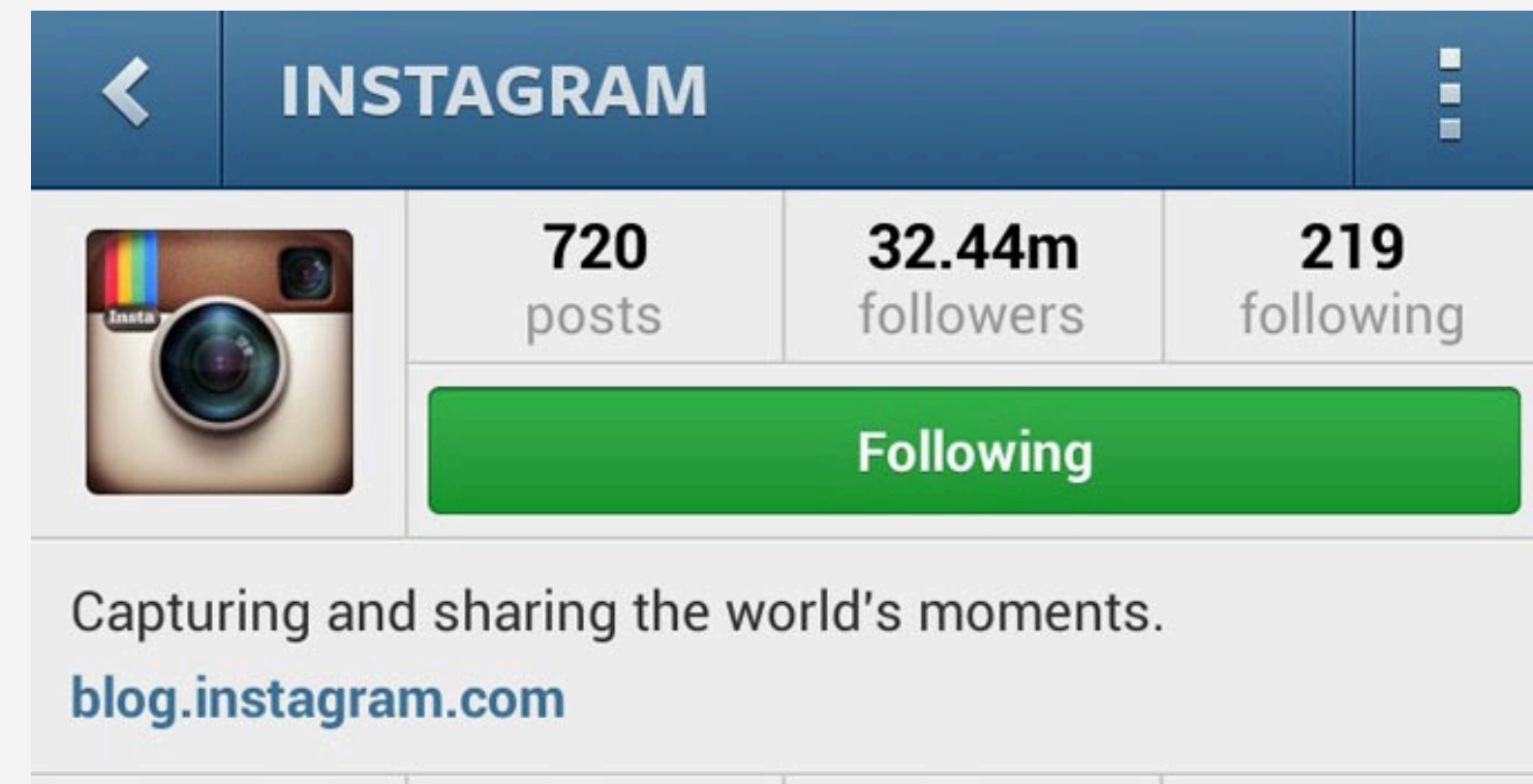
Cette évolution a transformé le web en un espace de collaboration, où les internautes pouvaient publier des commentaires, ajouter des contenus et personnaliser leur expérience en ligne. Cependant, même si le web 2.0 a amélioré l'interaction et l'engagement des utilisateurs, il présentait encore des limites en termes d'organisation et d'accessibilité de l'information.

2.0

WEB 2.0 OU WEB SOCIAL (ANNÉES 2000)

Dans les années 2000, le web a évolué vers le web social, marquant le passage d'un web passif à un web interactif. Les utilisateurs ont commencé à interagir avec le contenu et même à y contribuer activement. De nouvelles plateformes, telles que les réseaux sociaux et les sites de partage de vidéos ont permis aux utilisateurs de créer et partager du contenu.

Octobre 2010



source : <https://www.numerama.com/business/32108-instagram-en-boucle.html>

HISTOIRE - PROBLÈMATIQUE

WEB 3.0 OU WEB SÉMANTIQUE (ANNÉES 2010)

Le web 2.0 a apporté plus d'interaction, mais il reste difficile pour les utilisateurs de trouver rapidement et efficacement des informations pertinentes.

Imaginez que vous souhaitez organiser un voyage de Lyon à Paris:

- Vous saisissez "moyens de transport de Lyon à Paris" dans un moteur de recherche
- Vous obtenez une liste de liens vers différents sites web, articles de blogs, forums, etc
- Vous devez visiter plusieurs pages pour extraire des informations pertinentes, comme les horaires de train, les prix des billets d'autobus, et les options de covoiturage
- Les informations sont souvent présentées dans des formats différents. Ce qui complique la comparaison des options disponibles

A search bar with a magnifying glass icon on the left, the text "Moyens de transport de Lyon à Paris..." in the center, and a close button (X) on the right.

Q Moyens de transport de Lyon à Paris... X

Sans structure standardisée, il est difficile de comparer les différentes options de transport en un coup d'oeil. Ce qui fait que les utilisateurs passent beaucoup de temps à naviguer entre différents sites et à extraire des informations pertinentes manuellement.

HISTOIRE - PROBLÈMATIQUE

WEB 3.0 OU WEB SÉMANTIQUE (ANNÉES 2010)

Ces limitations soulèvent une question cruciale : **Comment alors permettre aux machines de comprendre et d'interpréter le sens des données sur le Web, afin d'améliorer l'interopérabilité, la recherche d'information et l'accessibilité des données, tout en favorisant une meilleure connexion entre les informations liées?**



Le web sémantique vise à améliorer l'interopérabilité, l'accessibilité et la recherche d'informations en standardisant la manière dont les données sont représentées et liées sur le web. En **structurant** les **données** de façon **compréhensible** pour les **machines**, le web sémantique facilite la connexion entre les informations et améliore les processus d'extraction et de comparaison de données, offrant ainsi une expérience plus fluide et intelligente aux utilisateurs.

HISTOIRE - PROBLÈMATIQUE

WEB 3.0 OU WEB SÉMANTIQUE (ANNÉES 2010)

Ce contraste met en lumière comment le web sémantique peut surmonter les limitations du web traditionnel en rendant les informations plus accessibles, interopérables et exploitables par les machines.

Scénario de recherche du moyens de transports avec le web sémantique

Avec le web sémantique, ce processus pourrait être beaucoup plus efficace :

- Données liées : les informations sur les transports seraient interconnectées grâce à des métadonnées et des ontologies. Un moteur de recherche pourrait comprendre que "train", "bus" et "covoiturage" sont des modes de transport.
- Résultats enrichis : En réponse à la recherche, le moteur pourrait présenter un tableau comparatif des horaires, des prix et des durées de trajet, extraits automatiquement de différentes sources, facilitant ainsi la prise de décision.

FONCTIONNEMENT DU WEB SÉMANTIQUE

01

XML (EXTENSIBLE MARKUP LANGUAGE)

02

RDF (RESOURCE DESCRIPTION FRAMEWORK)

03

ONTOLOGIES

XML (EXTENSIBLE MARKUP LANGUAGE)

- Créé par XML Working Group, formé par le World Wide Web Consortium (W3C) en 1996.
- Conçu pour stocker et transporter des données.
- Conçu pour être lisible à la fois par les humains et par les machines.
- Permet aux utilisateurs d'ajouter une structure arbitraire à leurs documents, mais ne dit rien sur la signification de ces structures.

```
<?xml version="1.0" encoding="UTF-8"?>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```

source : <https://www.w3schools.com/xml/>

RDF (RESOURCE DESCRIPTION FRAMEWORK)

- Créé par RDF Working Group, formé par le World Wide Web Consortium (W3C) en 1999.
- La signification est exprimée par le RDF, qui l'encode sous forme d'ensembles de triplets (**sujet, prédicat (verbe), objet**) d'une phrase élémentaire.
- Ces triplets peuvent être écrits en utilisant des balises XML / HTML.
- Dans le RDF, un document fait des assertions selon lesquelles des éléments particuliers (personnes, pages Web ou autres) possèdent des propriétés (comme "**est la sœur de**", "**est l'auteur de**") avec des valeurs spécifiques (une autre personne, une autre page Web).

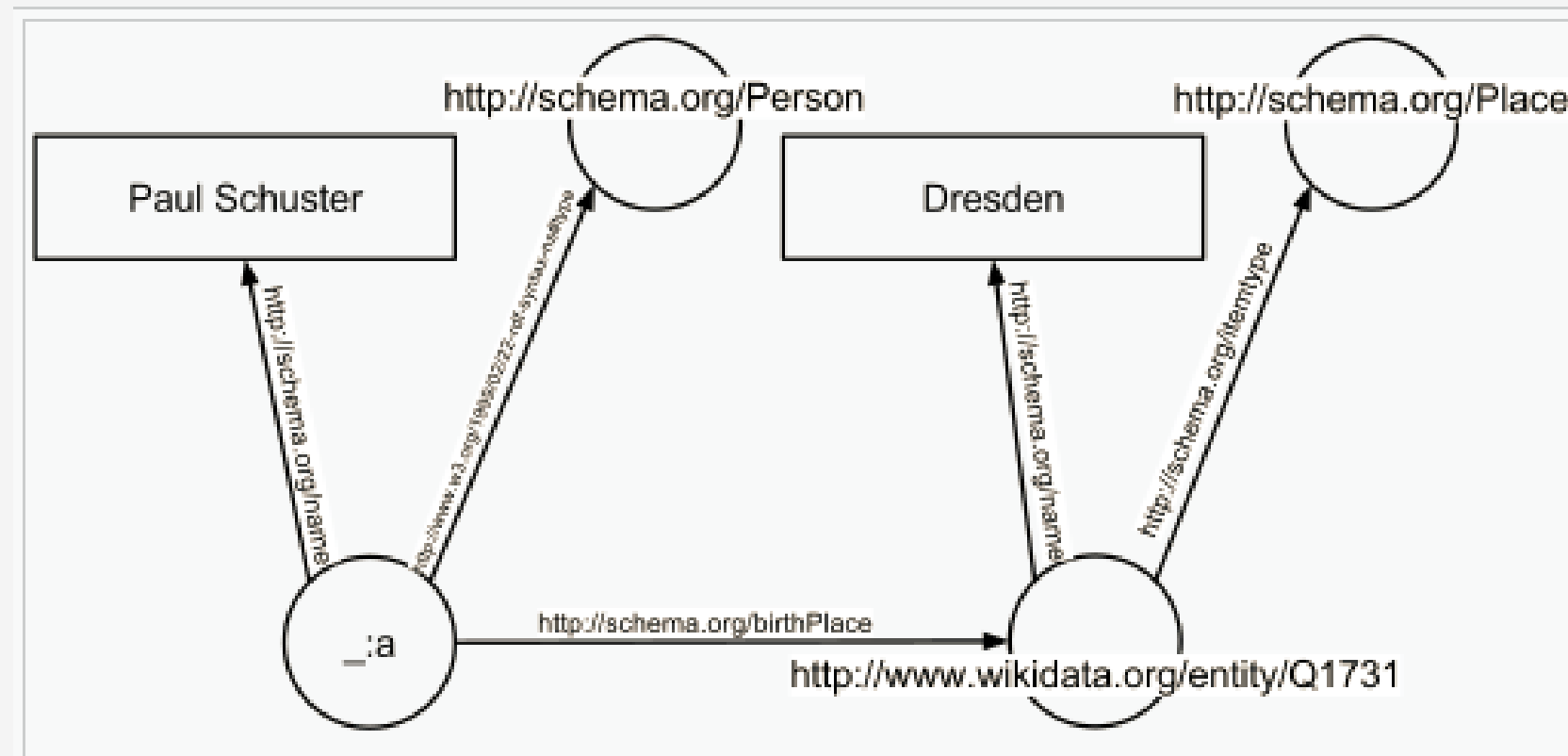
```
<div vocab="https://schema.org/" typeof="Person">
  <span property="name">Paul Schuster</span> was born in
  <span property="birthPlace" typeof="Place" href="https://www.wikidata.org/entity/Q1731">
    <span property="name">Dresden</span>.
  </span>
</div>
```

RDFa

```
_:a <http://www.w3.org/1999/02/22-rdf-syntax-ns#type [archive]> <http://schema.org/Person [archive]> .
_:a <http://schema.org/name [archive]> "Paul Schuster" .
_:a <http://schema.org/birthPlace [archive]> <http://www.wikidata.org/entity/Q1731> .
<http://www.wikidata.org/entity/Q1731> <http://schema.org/itemtype [archive]> <http://schema.org/Place [archive]> .
<http://www.wikidata.org/entity/Q1731> <http://schema.org/name [archive]> "Dresde" .
```

source : https://en.wikipedia.org/wiki/Semantic_Web

RDF (RESOURCE DESCRIPTION FRAMEWORK)



source : https://en.wikipedia.org/wiki/Semantic_Web

Universal Resource Identifier (URI)
 ex : URLs (Uniform Resource Locators)

RDF (RESOURCE DESCRIPTION FRAMEWORK)

SPARQL (SPARQL Protocol and RDF Query Language)

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
SELECT DISTINCT ?nom ?image ?description
WHERE {
  ?personne rdf:type foaf:Person.
  ?personne foaf:name ?nom.
  ?image rdf:type foaf:Image.
  ?personne foaf:img ?image.
  ?image dc:description ?description
}

```

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rss="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <foaf:Person rdf:about="http://example.net/Paul_Dupont">
    <foaf:name>Paul Dupont</foaf:name>
    <foaf:img rdf:resource="http://example.net/Paul_Dupont.jpg"/>
    <foaf:knows rdf:resource="http://example.net/Pierre_Dumoulin"/>
  </foaf:Person>
  <foaf:Person rdf:about="http://example.net/Pierre_Dumoulin">
    <foaf:name>Pierre Dumoulin</foaf:name>
    <foaf:img rdf:resource="http://example.net/Pierre_Dumoulin.jpg"/>
  </foaf:Person>
  <foaf:Image rdf:about="http://example.net/Paul_Dupont.jpg">
    <dc:description>Photo d'identité de Paul Dupont</dc:description>
  </foaf:Image>
  <foaf:Image rdf:about="http://example.net/Pierre_Dumoulin.jpg">
    <dc:description>Photo d'identité de Pierre Dumoulin</dc:description>
  </foaf:Image>
</rdf:RDF>

```

source : <https://fr.wikipedia.org/wiki/SPARQL>

ONTOLOGIES

- Une ontologie est une théorie sur la nature de l'existence, sur les types de choses qui existent ; l'ontologie, en tant que discipline, étudie ces théories. **(Philosophique)**
- Une ontologie est un document ou un fichier qui définit formellement les relations entre les termes. Le type d'ontologie le plus courant pour le Web comprend une taxonomie et un ensemble de règles d'inférence. **(Intelligence Artificielle)**

OWL (WEB ONTOLOGY LANGUAGE)

- Créé en 2004 par le World Wide Web Consortium (W3C).
- Utilisé pour représenter des ontologies sur le Web.
- Il permet de définir des concepts, des relations entre ces concepts, ainsi que des propriétés de ces relations de manière formelle.

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

  <!-- Définition des classes -->
  <owl:Class rdf:ID="Animal"/>
  <owl:Class rdf:ID="Mammifere">
    <rdfs:subClassOf rdf:resource="#Animal"/>
  </owl:Class>
  <owl:Class rdf:ID="Oiseau">
    <rdfs:subClassOf rdf:resource="#Animal"/>
  </owl:Class>

  <!-- Définition des propriétés -->
  <owl:ObjectProperty rdf:ID="aDesAiles"/>
  <owl:ObjectProperty rdf:ID="peutVoler"/>

  <!-- Définition des individus -->
  <rdf:Description rdf:about="#Aigle">
    <rdf:type rdf:resource="#Oiseau"/>
    <aDesAiles rdf:resource="true"/>
    <peutVoler rdf:resource="true"/>
  </rdf:Description>

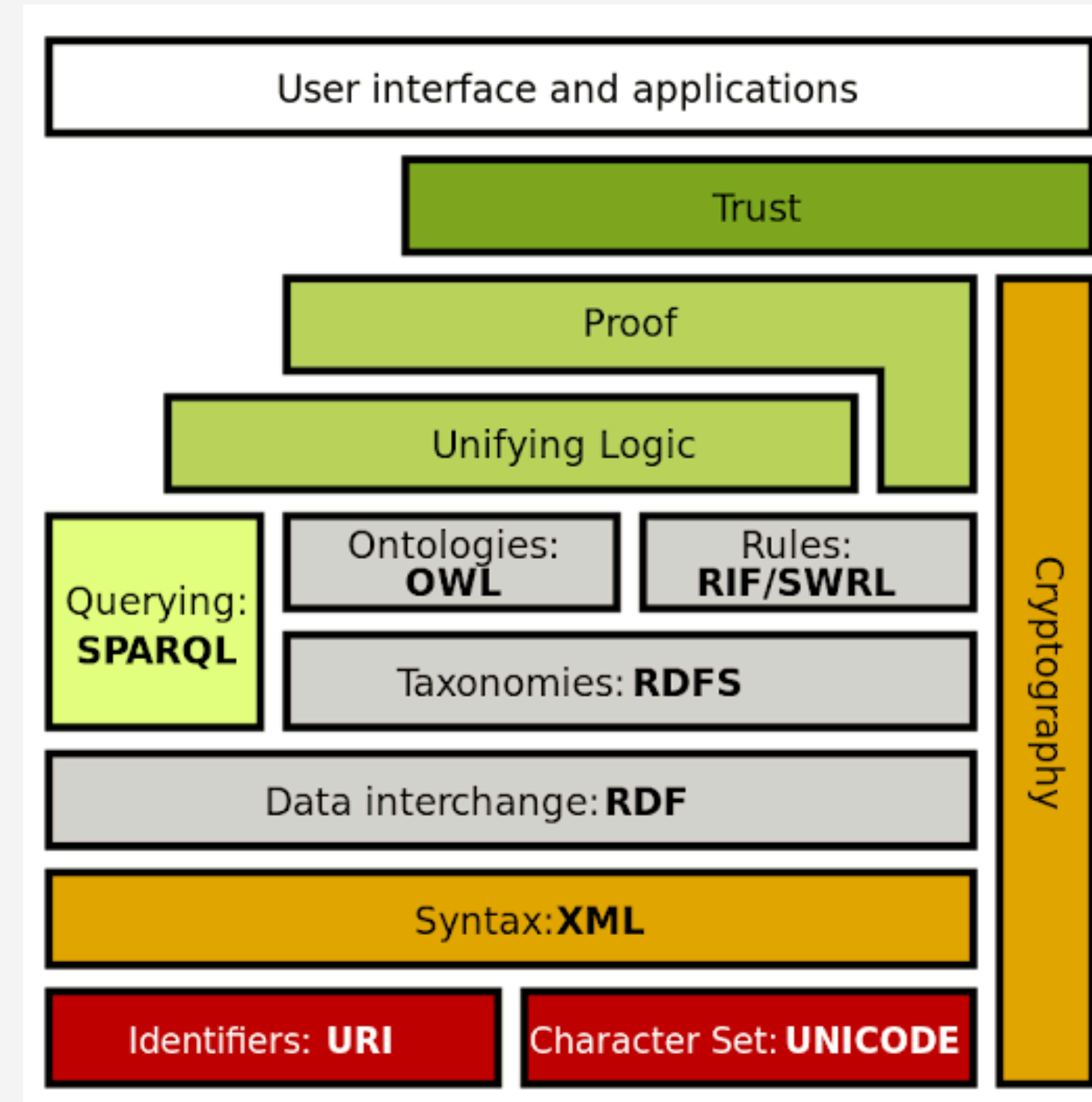
  <rdf:Description rdf:about="#ChauveSouris">
    <rdf:type rdf:resource="#Mammifere"/>
    <aDesAiles rdf:resource="true"/>
    <peutVoler rdf:resource="true"/>
  </rdf:Description>

  <rdf:Description rdf:about="#Elephant">
    <rdf:type rdf:resource="#Mammifere"/>
    <aDesAiles rdf:resource="false"/>
    <peutVoler rdf:resource="false"/>
  </rdf:Description>

</rdf:RDF>
```

AUTRES...

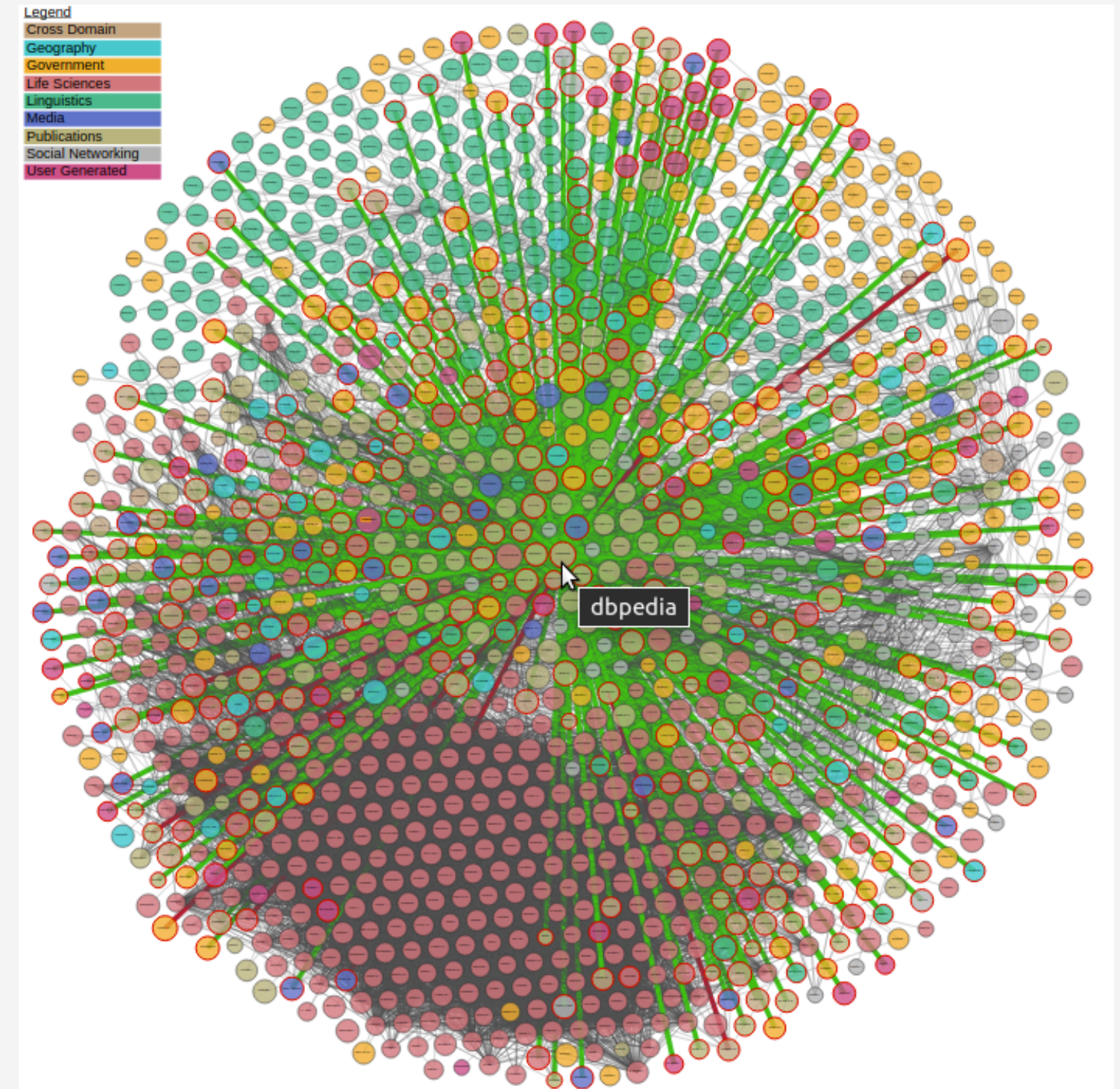
- **Unifying Logic** : Permet d'intégrer divers systèmes logiques sous une même structure, facilitant l'harmonisation et la cohérence des inférences sur des ensembles de données hétérogènes.
- **Proof** : Le mécanisme de preuve permet de valider et de vérifier la cohérence des inférences logiques.
- **Trust** : Le modèle de confiance évalue la fiabilité des données et des sources dans le Web sémantique.
- **RIF/SWRL** : règles définies par RIF (Rule Interchange Format) et SWRL (Semantic Web Rule Language) permettent d'exprimer des inférences et des relations complexes sur des données du Web sémantique, en ajoutant des capacités logiques avancées aux ontologies pour déduire de nouvelles informations à partir de faits existants.



source : <https://lincsproject.ca/fr/docs/terms/semantic-web>

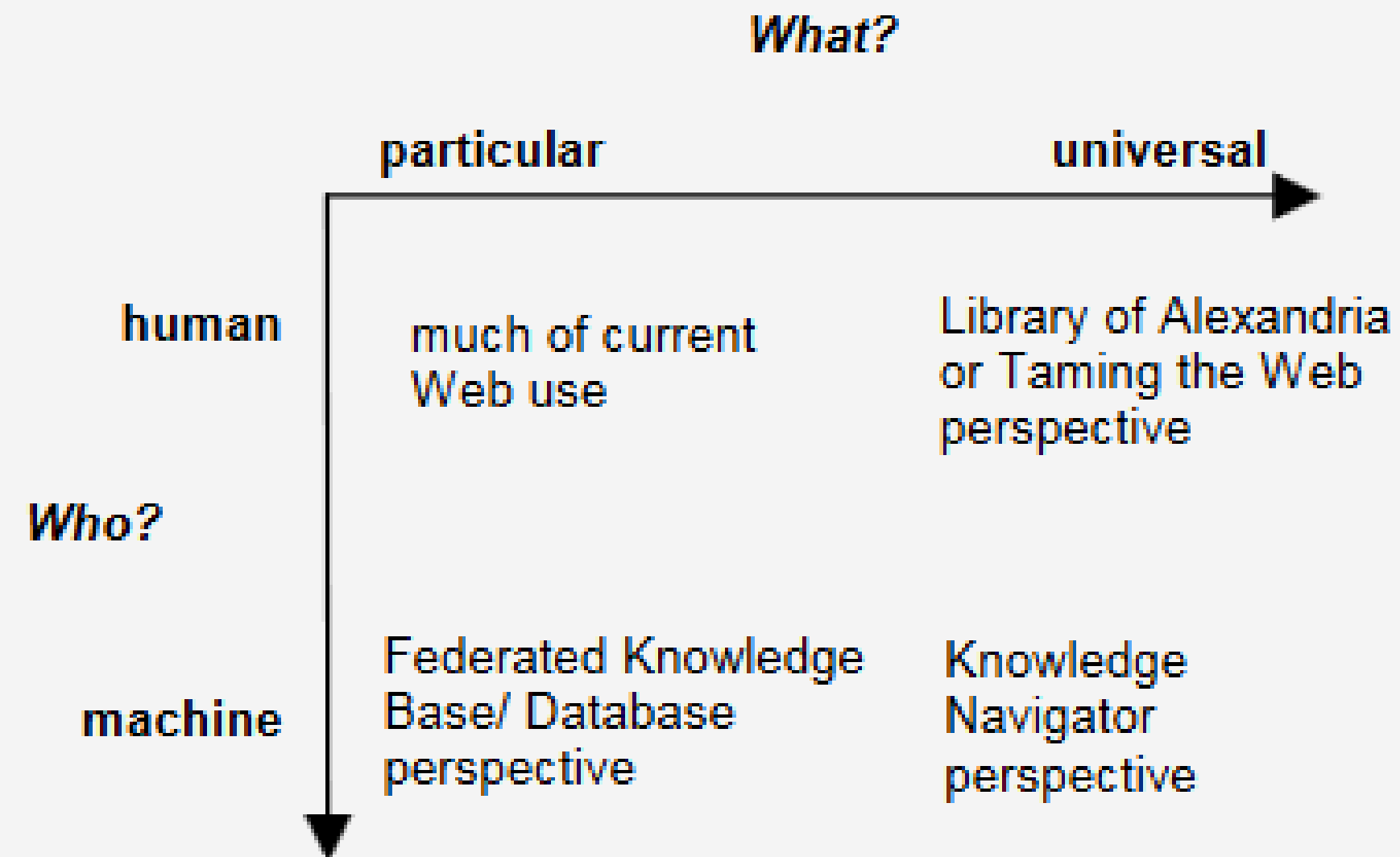
LINKED DATA

Les **données liées** (*Linked Data*) sont un concept créé par **Tim Berners-Lee** en **2006**, visant à **structurer** les données sur le Web de manière interconnectée. Elles permettent de lier des informations provenant de différentes sources en utilisant des formats standardisés comme RDF et des URI (Uniform Resource Identifiers). L'objectif est de faciliter l'accès, le partage et l'intégration des données à travers différents systèmes, rendant les informations plus exploitables par les machines et les humains.



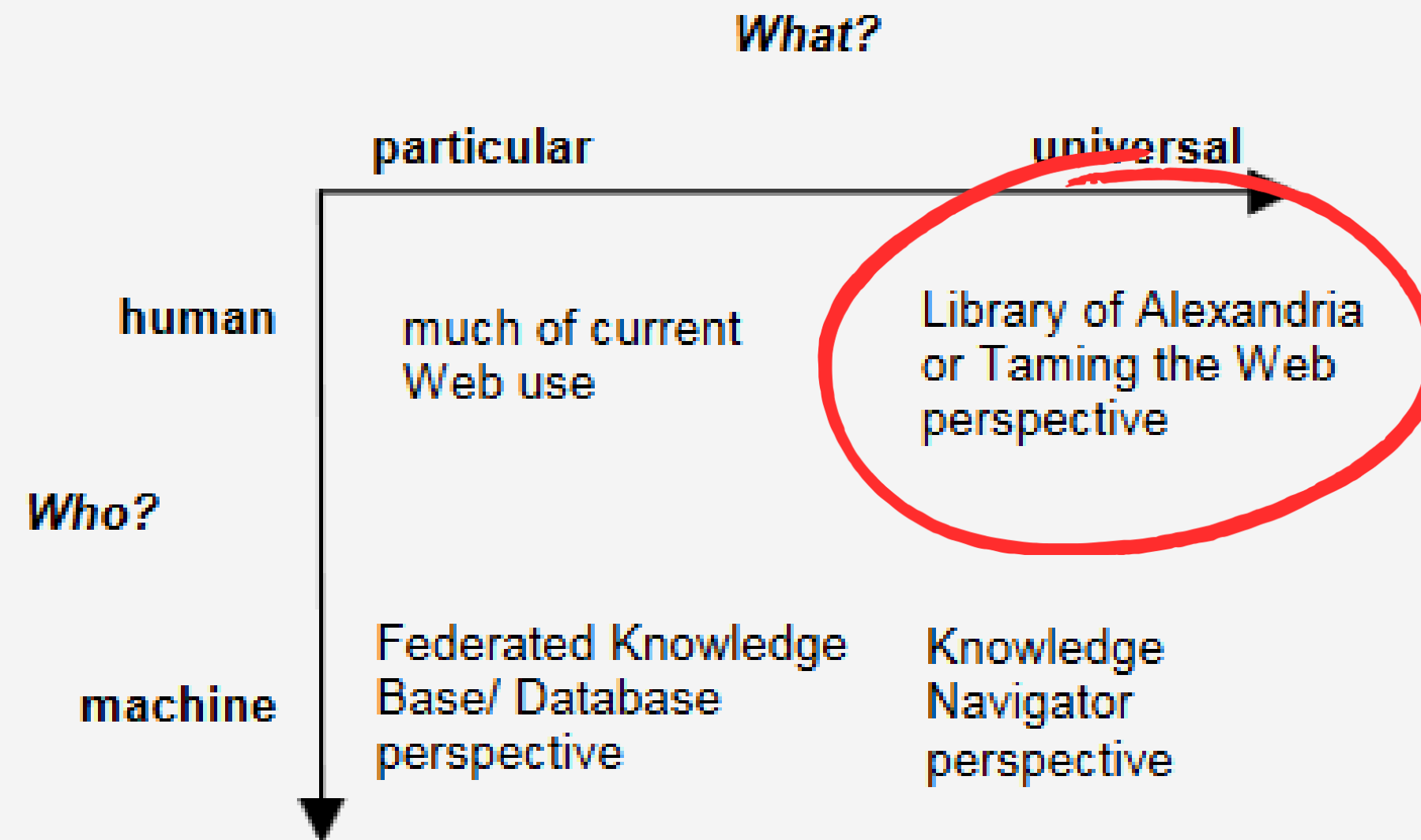
source : <https://lod-cloud.net/>

VISIONS DU WEB SÉMANTIQUE



source : "Which semantic web?" (2003), Catherine C. Marshall and Frank M. Shipman

VISIONS DU WEB SÉMANTIQUE

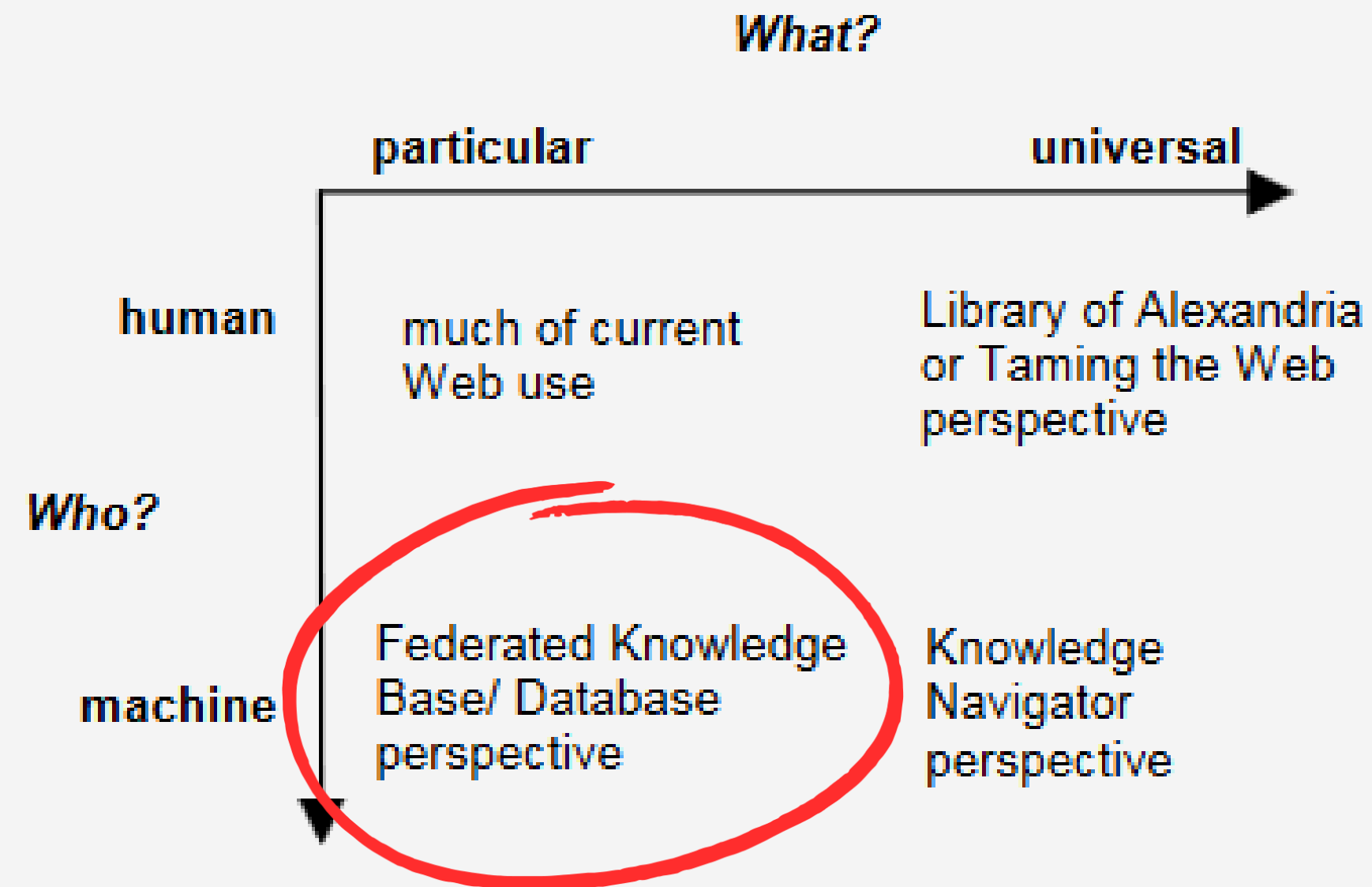


“DOMPTER LE WEB” OU LA BIBLIOTHÈQUE D’ALEXANDRIE

Une bibliothèque universelle accessible aux humains pour la recherche et l’usage d’information

- Né d’une certaine anxiété à propos du désordre du Web
- Envie d’ordonner le Web, de catégoriser l’information présente et et lui donner un sens bien défini pour permettre d’optimiser la recherche et la manipulation d’informations en ligne
- Cette vision est devenue moins pertinente avec l’apparition des moteurs de recherche
- De nos jours, les applications de cette vision se veulent plus être des extensions du Web actuel qu’un autre Web à part entière, notamment à cause de la difficulté d’imposer de nouveaux standards Web plus complexes à des utilisateurs qui ne respectent déjà pas trop les standards Web actuel

VISIONS DU WEB SÉMANTIQUE

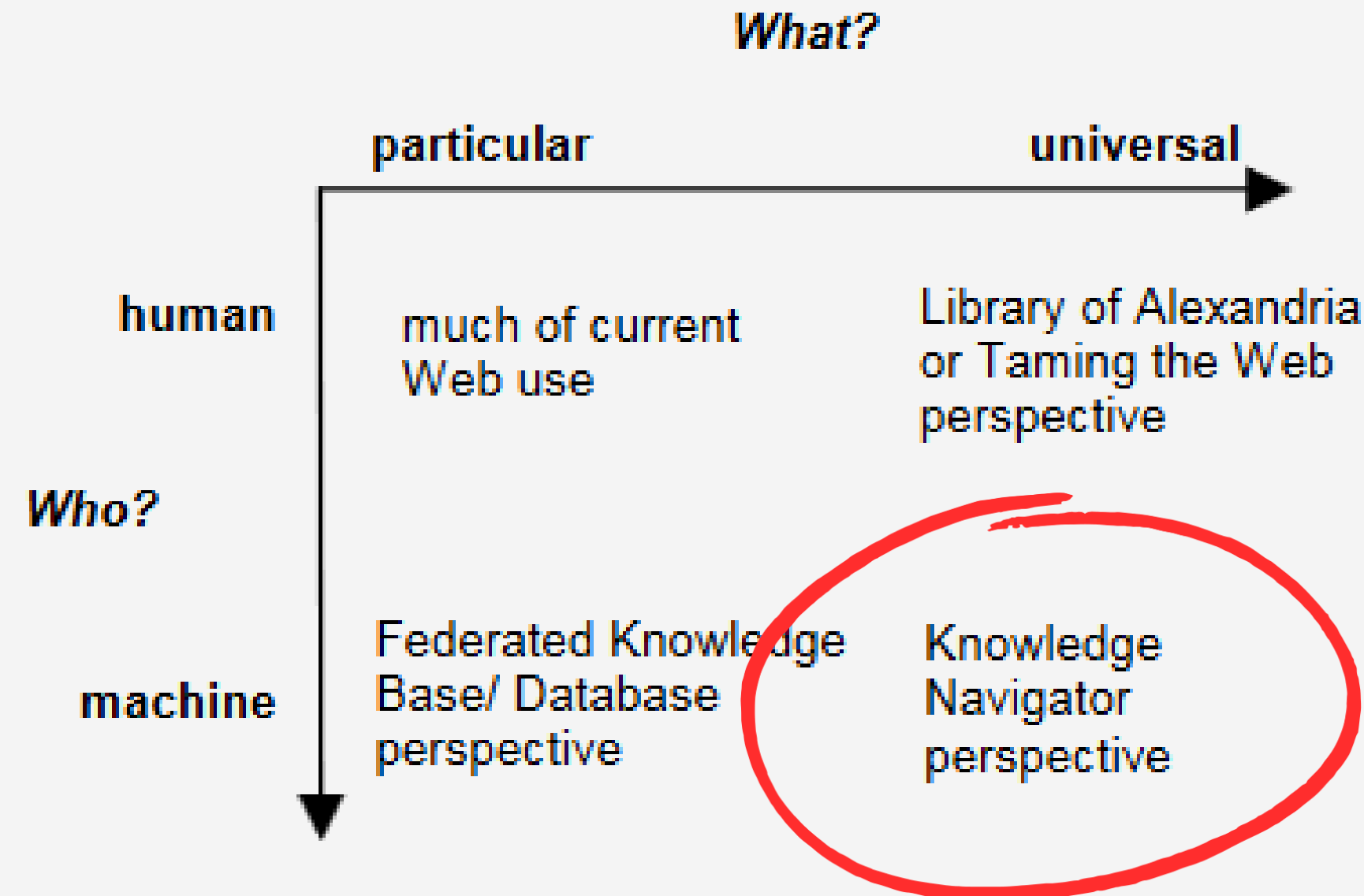


BASE FÉDÉRÉE DE CONNAISSANCE ET DE DONNÉES

Une infrastructure permettant la communication entre les bases de données sur le Web, une méthode pour fédérer les bases de connaissances afin d'accomplir des tâches les impliquant

- Inspiré par le domaine de l'IA, et particulièrement des représentations formelles
- But : permettre à des agents d'effectuer des tâches sur différentes bases de données sans se soucier de l'authentification, des protocoles de communication, des permissions d'accès, du langage de représentation des données --> tout cela sera géré par le Web Sémantique
- Comme la première vision, les applications actuelles (plus nombreuses) se voient plus comme des extensions du Web plutôt qu'un nouveau Web, notamment à cause du problème de standardisation
- En effet, la démocratisation de cette vision nécessiterait un réel accord sur le format d'échange de l'information stockée sur le Web

VISIONS DU WEB SÉMANTIQUE



NAVIGATEUR DE CONNAISSANCES OU LA VISION DE BERNERS-LEE

Une infrastructure permettant à des agents d'effectuer des tâches complexes pour les humains

- Né d'un désir de tout simplement se débarrasser de la charge de la recherche d'information et de la complexité des tâches de la vie de tous les jours en laissant cette charge à une machine (= agent)
- Sorte de base de données mondialisée (une sorte de réseau d'ontologies) permettant l'acquisition, la représentation et l'utilisation de la connaissance
- Créer un Web complètement lisible par les machines, en parallèle du Web lisible par les humains
- Cette vision est ambitieuse : il faudrait arriver à mettre au point une représentation de la connaissance globale --> comment représenter le sens de chaque concept selon le contexte, et sans conflits ?
- Malgré l'étude des systèmes de représentation pendant 50 ans dans le domaine de l'IA, les seuls systèmes existants sont limités à un domaine spécifique et mises au point avec un problème à résoudre en tête

APPLICATIONS DU WEB SÉMANTIQUE

- **“Dompter le Web”** : moteurs de recherche enrichis, catalogage de bibliothèques en ligne...
 - ↳ exemples spécifiques d’applications : Schema.org, ScholOnto
- **Base fédérée** : e-commerce, ontologies spécifiques à un domaine, partage de données médicales...
 - ↳ exemples spécifiques d’applications : SI-Designer, DBpedia
- **Navigateur de connaissances** : assistants personnels, recommandation contextuelles...
 - ↳ exemples spécifiques d’applications : Siri/Cortana/etc., Zotero

EVOLUTION DU WEB SÉMANTIQUE

1

CONCEPTUALISATION ET PREMIÈRES ANNÉES (2001-2005)

- Vision initiale de Tim-Berners Lee dans la revue The Scientific American en 2001
- Popularisation du terme par Tim O'Reilly lors de la première conférence sur le Web 2.0 en 2004
- Which semantic web? par Marshall et Shipman sur les différentes approches et attentes du web sémantique
- Débat sur la complexité des ontologies par Shirky dans son essai "Ontology is overrated"

2

ÉMERGENCE DU WEB DES DONNÉES (2006-2010)

- Principes du Linked Data proposés par Tim Berners-Lee en 2006
- Lancement de projets concrets comme DBpedia qui extrait des données structurées de Wikipedia en 2007
- Développement de la gestion de l'incertitude
- Création de SPARQL (2008), un langage de requête pour RDF, facilitant l'interrogation des données liées

EVOLUTION DU WEB SÉMANTIQUE

3

MATURITÉ ET INTÉGRATION (2010-2015)

- Introduction de JavaScript Object Notation for Linked Data (JSON-LD) pour améliorer l'échange d'information entre les documents JSON et RDF
- Développement de schema.org en 2011, comme un effort collaboratif pour créer des vocabulaires partagés
- Développement des Knowledge Graphs
- Application dans des domaines spécifiques

4

L'ÈRE DU BIG DATA ET DE L'IA (2015-PRÉSENT)

- Développement de la recherche sémantique avec une plus grande intégration dans les moteurs de recherche
- Utilisation croissante des graphes de connaissances dans les assistants virtuels (ex: Siri, Alexa) pour améliorer la compréhension du contexte
- Utilisation des principes du web sémantique dans les LLMs

AXES DE RECHERCHE

Utilisation du web sémantique dans le data mining

La gestion efficace des big data sémantiques

Le développement de méthodes de raisonnement et d'inférence à grande échelle

L'intégration du web sémantique avec l'Internet des Objets (IoT)

L'amélioration des techniques d'extraction et d'annotation sémantique

La visualisation des données sémantiques complexes

LIMITES DU WEB SÉMANTIQUE

Limites théoriques

Complexité de la création
des ontologies

Incertitude de l'information

Limites pratiques

Scalabilité et performance

Qualité et fiabilité des données

Confidentialité et sécurité

Standardisation et adoption

Mise en place et compétitivité

LIENS AVEC LA COGNITION

Approche	Lien avec le web sémantique
<p>Cognitivism repose sur l'idée que la cognition est fondée sur des représentations mentales symbolique et des règles de traitement de l'information</p>	<p>Le web sémantique s'inscrit clairement dans cette approche, car il repose sur des représentations symboliques formelles pour structurer l'information et permettre aux machines d'inférer des connaissances à partir de données structurées.</p>
<p>Connexionnisme se concentre sur les réseaux de neurones et les modèles d'apprentissage distribué</p>	<p>Le web sémantique ne s'aligne pas directement avec cette approche, car il repose davantage sur la représentation explicite des connaissances plutôt que sur l'apprentissage à partir de données non structurées</p>
<p>Énactivisme met l'accent sur l'idée que la cognition émerge de l'interaction active avec l'environnement</p>	<p>Le web sémantique n'est pas directement aligné avec l'enactivisme, car il se concentre sur la structuration et l'organisation des données plutôt que sur l'interaction avec le monde réel.</p>
<p>Approche dynamique de la cognition considère les processus cognitifs comme des systèmes en constante évolution</p>	<p>Le web sémantique s'inscrit dans cette approche, car il intègre des éléments dynamiques, tels que les flux de données ou les graphes de connaissances qui évoluent avec le temps</p>
<p>Approche systèmes complexes étudie les systèmes composés de nombreuses unités interagissant de manière non linéaire</p>	<p>Le web sémantique peut être vu sous cet angle si l'on considère le web comme un système complexe où les données sont distribuées, interconnectées, et où le sens émerge de l'interaction entre différents ensembles de données</p>

Références

Uncertainty Reasoning for the World Wide Web. www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331.

Elanceur - L'Ontologie Est Surfaite.

web.archive.org/web/20080628190832/http://www.elanceur.org/Articles/OntologieSurfaite.html.

Shanker, Surabhi, and Saloni Bhushan. "Semantic Web-A Dream of High -Tech World." ResearchGate, May 2016, www.researchgate.net/publication/342151930_Semantic_Web-A_Dream_of_High_-Tech_World.

Marshall, C. C., & Shipman, F. M. (2003). Which semantic web? In Proceedings of the fourteenth ACM conference on Hypertext and hypermedia (Vol. 1, pp. 57–66). HT03: The Fourteenth ACM Conference on Hypertext and Hypermedia. ACM. <https://doi.org/10.1145/900051.900063>

The Semantic Web: Scientific American.

web.archive.org/web/20081114135540/http://www.sciam.com/article.cfm?id=the-semantic-web&print=true.

Scherp, Ansgar, et al. "Semantic Web: Past, Present, and Future." drops.dagstuhl.de, 2024, doi:10.4230/TGDK.2.1.3.

Lawton, George. "Semantic Web." CIO, 26 Jan. 2023, www.techtarget.com/searchcio/definition/Semantic-Web.



“Je rêve d'un Web dans lequel les ordinateurs deviennent capables d'analyser toutes les données du Web : le contenu, les liens, et les transactions entre personnes et ordinateurs. Un “Web Sémantique”, qui devrait rendre cela possible, n'a pas encore émergé, mais quand il le fera, les mécanismes journaliers du commerce, de l'administration et de nos vies quotidiennes seront traités par des machines dialoguant avec d'autres machines. Les “agents intelligents” qu'on nous vante depuis longtemps se concrétiseraient enfin.”

- Tim Berners-Lee, Weaving the Web



Est-ce que vous pensez que la vision originelle de Berners-Lee est réalisable ?



Est-il envisageable de développer un raisonnement symbolique assez puissant pour concurrencer les approches d'apprentissage profond dans le domaine de la compréhension du langage ?



Dans quelle mesure le raisonnement symbolique peut-il être intégré aux réseaux de neurones pour améliorer la compréhension contextuelle des systèmes d'IA ?



SYNTHÈSE