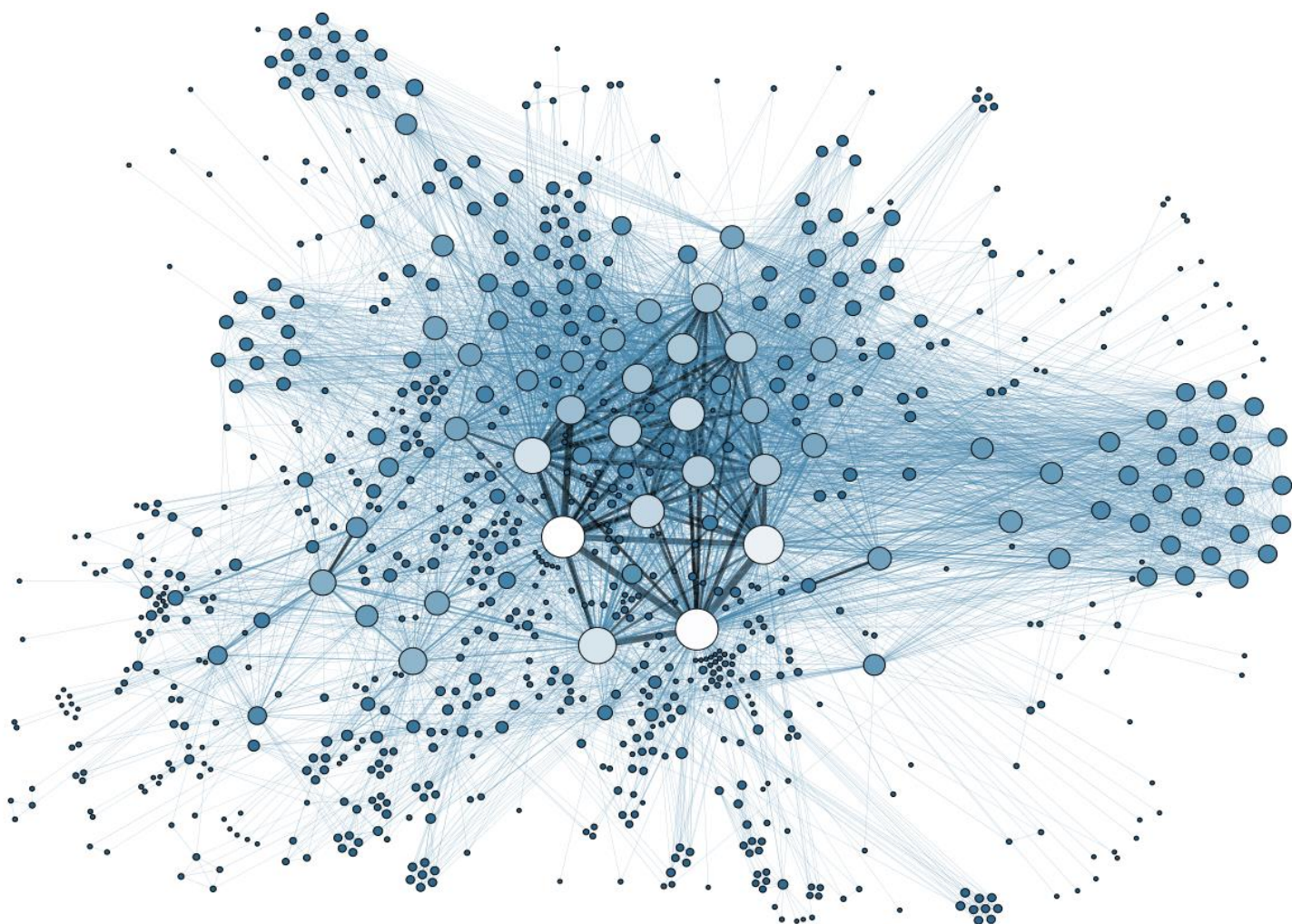


Projet 2 :

# LINKED OPEN DATA



Clarisse VERON – Pierre NEVEU – Mariana  
DEL PILAR MOLL DE ALBA

# SOMMAIRE

- 01 Les enjeux de la mise en place des Linked Open Data
- 02 Le processus de fonctionnement des datas
- 03 Un réseau riche d'informations source de limites

# 01 Les enjeux de la mise en place des Linked Open Data

## Comment définir les Linked Open Data ?

Afin de mieux comprendre que sont les Linked Open Data, il est important de définir les différentes notions présentes dans celle-ci. Tout d'abord, le terme « Linked Data » se réfère à des informations Web pouvant être lues, mais aussi être comprises par des ordinateurs. Cela peut se référer au Web sémantique, dont nous aborderons le concept plus tard. Par la suite, nous pouvons définir le terme « Open », se référant au statut légal, à la concession de licences indiquant qu'il est possible de réutiliser des données ouvertes. Or, la mise en place du Linked Open Data n'est pas arrivée du jour au lendemain. En effet, cette notion a permis une certaine transition pour passer d'un web compréhensible par l'humain et les machines, à un web manipulable pour les machines. Être manipulable dans ce contexte, c'est passer d'un web des documents à un web de données interconnectées entre elles : ceci n'est autre que le prolongement du web qui existe déjà depuis plusieurs années. C'est pourquoi le terme de « Open » apparaît ici.

Il est aussi important d'insister sur la différence entre données ouvertes et données liées. En effet, les données ouvertes signifient que les données peuvent être publiées et rendues publiques sous une licence ouverte sans les lier à d'autres sources. En revanche, les données liées sont reliées aux URIs (Identifiant Uniforme de Ressource), notions que nous aborderons plus tard, d'autres sources de données, en utilisant des standards ouverts tels que RDF (expliqué plus tard), sans être disponibles publiquement sous une licence.

Afin de mieux comprendre les étapes qui ont conduit Tim Berners-Lee à publier les Linked Open Data, nous allons revenir dans le temps, en nous appuyant sur la théorie de Tim Berners-Lee, et ce qui l'a poussé à créer les Linked Open Data.

# LES ENJEUX DE LA MISE EN PLACE DU LINKED OPEN DATA

## **Un retour en arrière pour mieux comprendre leur apparition**

Comme évoqué précédemment, les Linked Open Data ne sont pas arrivés par hasard. En effet, Tim Berners-Lee, un brillant informaticien britannique, en fut le principal créateur. A côté de multiples autres sujets, il s'interroge sur la manière dont nous utilisons l'information en général sur internet, ce qui donna lieu au « World Wide Web », défini comme un système hypertexte public fonctionnant sur internet et permettant de consulter avec un navigateur des pages accessibles sur des sites. Le « surnom » de cette invention est d'ailleurs « la toile d'araignée », faisant référence à l'interconnexion des hyperliens qui lient les pages web entre elles.

Mais pourquoi créa-t-il un tel système ? Tim Berners-Lee, dans un TED SHOW, explique que il était dans « un état de frustration », comme un ingénieur logiciel travaillant dans un énorme laboratoire de recherche mais qui devait changer de machine à chaque nouvelle manipulation, mais aussi de format de données, ou encore de logiciel, qui étaient tous incompatibles ! Pour lui, ce n'était pas possible que tant de données soient fermées entre elles, il voulait implémenter un « grand système de documentation virtuel », afin de faciliter la vie de tous ces ingénieurs, mais aussi de la population travaillant sur des machines comme lui. Or, à l'époque, la notion de « web » était assez floue pour les individus, mais cette notion de connexions et la création du premier site web fut une révolution pour tous. Mais cette invention ne lui suffisait pas : il souhaitait rassembler, après avoir créé le premier site web, les documents et les données. C'est ce qu'il a fait : il les a intégrées dans un certain logiciel, et ainsi absolument tout le monde pouvait mettre ses données sur le web !

C'est ainsi que, de façon très simple, le terme « Linked Data » (données liées) apparut, afin de désigner la technologie et son fonctionnement à l'intérieur du web. Or, il y a certaines règles à respecter si l'on veut insérer quelque chose sur le web. En voici les trois :

- Les utilisateurs devront toujours utiliser les adresses HTTP, afin de décrire les documents au mieux. Elles seront utilisées pour des personnes, des lieux, des produits, ou encore des événements : un quelconque concept devra commencer par http.
- Dans un second temps, comme ces adresses HTTP renvoient des informations dans un format standard, il est intéressant que les utilisateurs en sachent plus sur le concept, d'où la deuxième règle qui vise à renvoyer des informations importantes.
- Pour finir, il est important de préciser que les informations émises sur le web sont reliées à d'autres informations. La troisième règle consiste donc à dire que les données sont des relations.

# LES ENJEUX DE LA MISE EN PLACE DU LINKED OPEN DATA

Prenons un exemple, afin que cela soit plus clair : Pierre est un jeune homme né à Paris, Paris étant en France. Ici, nous avons déjà une relation, et cet élément lié est nommé avec un nom commençant par http. Depuis son nom http, je pourrai accéder à cette ressource. Dans notre exemple, je pourrai accéder à cette personne Pierre, puis par la ville où il est né (Paris), puis sa région, sa ville, les monuments de cette ville, la population de la ville, ... On comprend ici encore mieux le terme utilisé de « Linked Data ». Ce qui est important d'intérioriser, c'est que plus on a de données connectées présentes sur le web, plus elles deviennent puissantes.

De plus, Tim Berners-Lee mis l'accent sur la forme dont apparaissent les données. En effet, le web est constitué d'une diversité d'informations, il est donc important d'avoir le pouvoir d'en ajouter selon nos convenances. Par exemple, les données peuvent être gouvernementales, des données scientifiques, personnes, d'entreprises, événementielles, de discussion... Toutes ces données sont très importantes ! D'ailleurs, le créateur donne l'exemple des données gouvernementales : Barack OBAMA a dit dans un de ses discours que les données gouvernementales américaines seront disponibles sur le web « dans des formats accessibles », dit-il. Il a également insisté sur le fait qu'il espérait qu'elles utilisent le format « Linked Data ». Ce à quoi Tim Berners-Lee répond qu'il a bien fait d'établir ce lien de transparence, mais aussi car « ces données proviennent de tous les départements gouvernementaux. [...] C'est très utile, ça a de la valeur, je peux l'utiliser pour la société ». En effet, il fait ici référence à la société telle que l'éducation ou encore le journalisme qui peuvent utiliser ces données en les citant, et les intégrant dans leurs devoirs ou articles. Ainsi, grâce aux Linked Open Data, « on peut rendre la gestion du monde meilleure, en rendant ces données disponibles » selon le créateur des Linked Open Data. Ainsi, nous allons passer à la prochaine partie, qui va traiter des objectifs des Linked Open Data, en nous appuyant sur des exemples concrets.

# LES ENJEUX DE LA MISE EN PLACE DU LINKED OPEN DATA

## **Les principaux objectifs des Linked Open Data et quelques applications concrètes**

Après avoir abordés plus en détails les définitions concrètes et l'historique de l'apparition des Linked Open Data, nous allons aborder les principaux objectifs de ce concept. Comme nous l'avons vu précédemment avec l'exemple des données gouvernementales, promouvoir la transparence, la participation et l'Open Gouvernement sont des atouts essentiels. De plus, son apparition pourrait permettre selon Tim de favoriser la libération des données, et en particulier des données publiques, que cela soit pour l'Art, la Santé, le Gouvernement, ... Les Linked Open Data favorisent également l'innovation et la croissance économique des pays qui les utilisent. Pour finir, un des principal objectif est de rendre les données interopérables et compréhensibles par les machines, défini comme le web sémantique.

Le web sémantique est un concept visant à permettre aux différentes machines de comprendre la signification de l'information sur le web. Ainsi, on peut mettre ne place en plus du réseau des hyperliens entre les pages web classiques, un réseau de liens entre données structurées. Ce terme inventé par Tim Berners-Lee en 1989 permet aux machines d'accéder plus intelligemment aux différentes sources de données contenues sur le web, et permet à celles-ci d'effectuer des traitements plus précis pour les utilisateurs. Quant à la notion d'interopérabilité, elle désigne la capacité que possède un système, dont les interfaces sont entièrement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre, nous indique le site techno science. Il se différencie à la compatibilité, qui lui peut faire fonctionner un outil dans un environnement donné en respectant toutes les caractéristiques.

Pour mieux comprendre en quoi les Linked Open Data peuvent résoudre d'immenses challenges que l'humanité doit relever depuis des années, prenons un exemple sur le domaine de la santé. Nous savons que beaucoup de connaissances de la race humaine ont été trouvés, mais elles ne sont actuellement pas accessibles par tous, mais sont contenues dans des bases de données non partagées au monde. Le créateur pris dans son allocution au TED SHOW l'exemple de l'Alzheimer, très représentatif des grands challenges du monde d'aujourd'hui. En effet, il est nécessaire de comprendre le fonctionnement des médicaments pour y remédier, ou du moins avancer dans la recherche scientifique. Or, on a pu observer qu'il y avait énormément de données qui apparaissaient dans une telle base de données, d'autres données sur les protéines dans une autre base. C'est grâce aux Linked Data que les scientifiques peuvent désormais se poser des questions plus concrètes en ayant réunies toutes les informations nécessaires, comme : « Quelles sont les protéines utilisées dans la transduction des signaux et qui



# LES ENJEUX DE LA MISE EN PLACE DU LINKED OPEN DATA

sont en rapport avec les neurones pyramidaux ? ». En l'écrivant sur notre moteur de recherche, nous pouvons accéder à une multiplicité d'informations, mais pas la réponse à la question en elle-même, car encore aucun scientifique ne se l'est posée. Mais, en interrogeant les données du « Linked Data », il n'y aura seulement qu'une trentaine de réponses, et « chacune d'elle est une protéine qui a cette propriété », dit intelligemment Tim. Cela montre ainsi l'énorme capacité des données à créer des liens, qui peuvent énormément aider les scientifiques, en faisant de multiples liens entre des domaines et des questionnements différents.

Un autre exemple plus concret pour des étudiants de notre génération que le créateur met en avant est l'utilisation des réseaux sociaux : si on veut ajouter un ami sur un réseau social tel que Facebook, alors on a une relation et puis une donnée. Si on veut dire que telle photo image une telle personne, cela crée également une donnée. Ainsi, cela rend la vie des autres utilisateurs du site « plus intéressante », car ils auront des données supplémentaires sur l'information qu'ils souhaitaient rechercher.

Pour conclure sur les enjeux de la mise en place des Linked Open Data, nous pouvons dire qu'ils ont beaucoup évolué, et contrairement à ce que nous pourrions penser, ils sont partout dès lors que nous utilisons internet. Même les personnes qui produisent le moins sont elles-mêmes connectées avec tout le reste. Les Linked Open Data sont énormes, et sont présents dans tous les domaines de la vie. L'important est, celui Tim Berners-Lee, non pas « le nombre de sources des données », mais « de pouvoir toutes les relier ». C'est lorsqu'on connecte les données entre elles que cela passe au-dessus de la simple puissance du web des documents, d'où l'objectif révolutionnaire des Linked Open Data dans le monde d'aujourd'hui.

## 02 LE PROCESSUS DE FONCTIONNEMENT DES DATAS

### **L'interopérabilité, une mise en relation complexe**

Dans un monde où la quantité d'information explose, se pose la question de l'interopérabilité. L'interopérabilité, à ne pas confondre avec compatibilité, représente la capacité d'un système à fonctionner avec d'autres systèmes existants ou même futurs (et est donc d'autant plus un sujet de recherche permanente), comme nous l'avons vu précédemment dans la première partie. Le problème étant que chacun (Homme comme machine) utilise les données de manière différente, le but ultime est donc de créer un standard pour toutes les données, c'est-à-dire une façon universelle de représenter l'information, et ainsi permettre une mobilité absolue des données. Les API (Application Programming Interface) sont les outils les plus utilisés de nos jours pour convertir les données d'un système à l'autre et leur permettre d'interopérer. Nous allons voir comment les données du Web sont normalisées pour faciliter leur mobilité grâce aux Resource Data Framework.

### **Les Ressources Data Framework (RDF)**

Les RDF ont été créés par le World Wide Web consortium (W3C). Elles permettent d'exprimer de l'information sur des ressources. Ces ressources peuvent être quelconques : d'un concept abstrait à un document en passant par des personnes. Les RDF ont été prévus pour des situations dans lesquelles de l'information provenant du Web doit être traitée par des machines et fournit donc une manière commune pour représenter l'information à des machines ainsi qu'à des personnes. De ce fait, les personnes qui conçoivent les applications peuvent utiliser des outils de traitement, d'analyse de RDF et l'information peut être utilisable par des applications étrangères à celle pour laquelle elle a été originellement créée. Par exemple, les RDF peuvent être utilisés pour interconnecter des données du Web de manière massive. À l'heure actuelle, il existe de grandes sources d'Open Linked Data décentralisées telles que Dbpédia qui reprend toutes les données de Wikipédia afin d'en faire une structure ordonnée et normalisée.

- **Concrètement, comment ça marche ?**

Les RDF permettent de définir des relations entre les ressources de manière simplifiée : Ressource1 (sujet) et Ressource2(objet) seront représentées par le « triplet » <sujet><prédicat(~relation)><objet>. Le prédicat définit donc la nature de la relation qui lie



# LE PROCESSUS DE FONCTIONNEMENT DES DATAS

les deux ressources. La relation est appelée « une propriété » dans RDF et a un sens de lecture spécifique : du sujet vers l'objet. Pour mieux comprendre le concept voici un exemple :

```
<Arnaud> <est une> <personne>.  
<Arnaud> <est un ami de> <Béatrice>.  
<Arnaud> <est né> <le 14 juillet 1990>.  
<Arnaud> <s'intéresse à> <la Joconde>.  
<la Joconde> <a été créée par> <Léonard de Vinci>.  
<la vidéo 'La Joconde à Washington'> <est à propos de> <la Joconde>.
```

Dans cet exemple, on remarque que Arnaud est le sujet de plusieurs triplets et c'est souvent le cas dans les RDF. De par la multiplicité des relations entre Arnaud et d'autres ressources, on peut déduire beaucoup d'informations sur celles-ci. Ici, on peut facilement se rendre compte qu'Arnaud a probablement regardé la vidéo « La Joconde à Washington ». C'est en partie ce qui fait la puissance des RDF.

Afin d'identifier et de faire référence à chaque ressource, on utilise les International Resource Identifier (IRI). Ils peuvent apparaître aux 3 positions d'un triplet afin d'identifier le sujet, le prédicat ou l'objet . Par exemple, l'IRI pour le sujet/objet Léonard de Vinci dans DBpédia est:

[http://fr.dbpedia.org/resource/Léonard\\_de\\_Vinci](http://fr.dbpedia.org/resource/Léonard_de_Vinci)

Et l'IRI pour le prédicat « connaitre » est :

[<http://xmlns.com/foaf/0.1/knows>](http://xmlns.com/foaf/0.1/knows)

En effet, ici l'IRI de Léonard de Vinci dans DBpédia est un URL, c'est un sous-ensemble des IRI.

- **Pourquoi utiliser les RDF ?**

Plutôt que d'expliquer l'utilité par des concepts abstraits : voici différents usages pratiques parmi tant d'autres :

Les RDF permettent d'ajouter des informations lisibles sur des pages Web dans un format amélioré et d'être traitées par des application externes. La force des RDF réside aussi dans leur capacité à enrichir un jeu de données en le liant à d'autres, et ainsi élargir le panel d'informations en plus de leur donner plus de sens. En interconnectant de la sorte les jeux de données au sein d'une organisation, les requêtes SPARQL (langage de requêtes comme SQL) seront facilités. Les RDF permettent également d'agréger (regrouper) les Linked Data autour

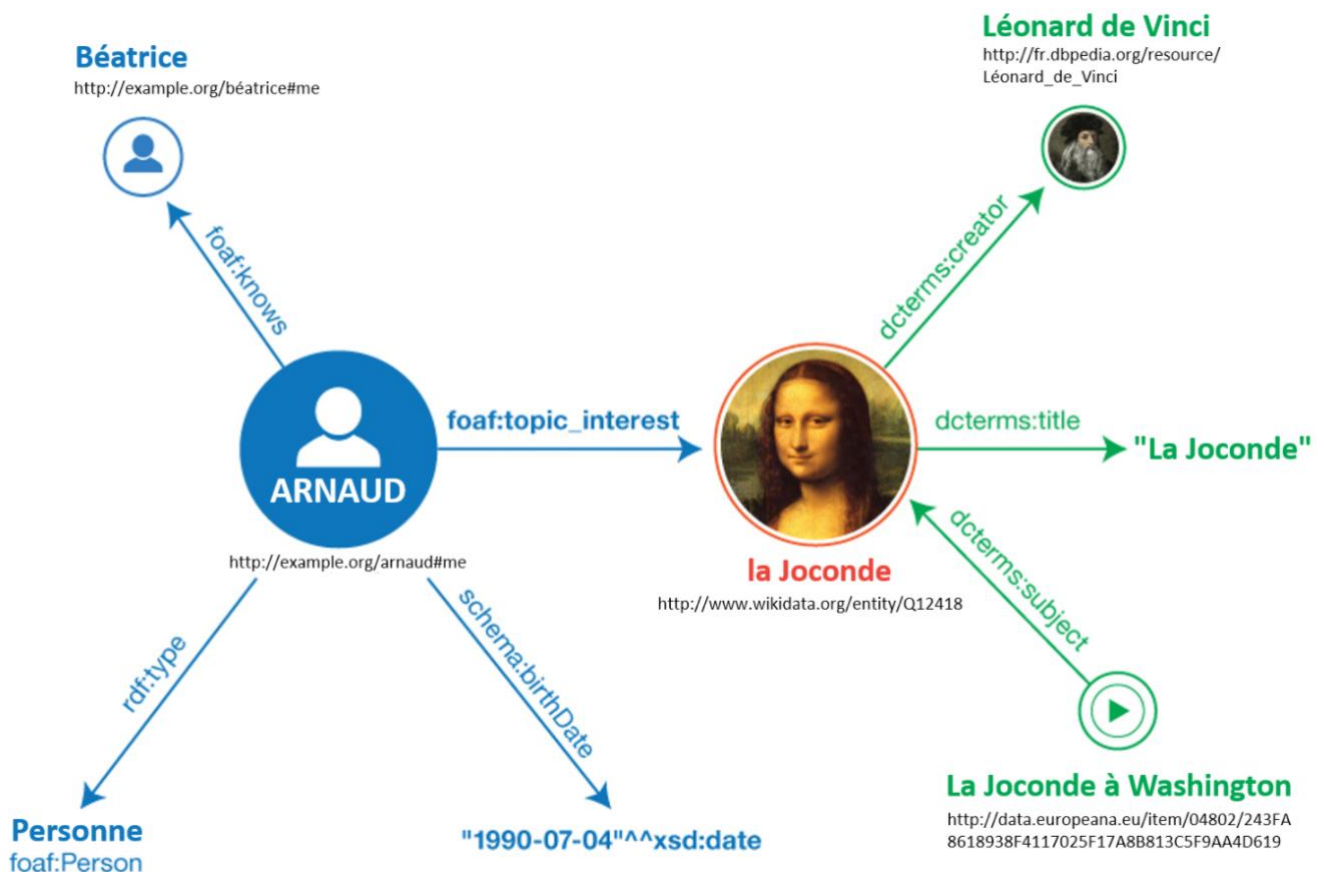
# LE PROCESSUS DE FONCTIONNEMENT DES DATAS

de thèmes spécifiés. De par leur simplicité, elles permettent de fournir un moyen conforme aux normes pour l'échange des données entre les bases de données.

- **Comment fonctionnent les langages RDF ?**

Nous avons présenté RDF dans les grandes lignes. Sans rentrer dans les détails, essayons de comprendre maintenant comment fonctionnent les langages RDF.

« N-Triples » est un langage RDF qui fournit la syntaxe de base pour écrire les triplets RDF. Ce langage met à disposition un moyen simple, ligne à ligne en texte brute, de sérialiser des graphes RDF. La sérialisation dans notre cas, représente l'action d'explicitier les 3 composantes d'un triplet plutôt que d'exploiter celui-ci sous forme compacte et implicite. Voyons un exemple :



# LE PROCESSUS DE FONCTIONNEMENT DES DATAS

Cette figure peut être représenté en N-triples de la manière suivante :

```
01 <http://example.org/arnaud#me> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> .
02 <http://example.org/arnaud#me> <http://xmlns.com/foaf/0.1/knows> <http://example.org/béatrice#me> .
03 <http://example.org/arnaud#me> <http://schema.org/birthDate> "1990-07-14"^^<http://www.w3.org/2001/XMLSchema#date> .
04 <http://example.org/arnaud#me> <http://xmlns.com/foaf/0.1/topic_interest> <http://www.wikidata.org/entity/Q12418> .
05 <http://www.wikidata.org/entity/Q12418> <http://purl.org/dc/terms/title> "La Joconde" .
06 <http://www.wikidata.org/entity/Q12418> <http://purl.org/dc/terms/creator> <http://fr.dbpedia.org/resource/Léonard_de_Vinci> .
07 <http://data.europeana.eu/item/04802/243FA8618938F4117025F17A8B813C5F9AA4D619> <http://purl.org/dc/terms/subject> <http://www.wikidata.org/entity/Q12418> .
```

Dans ces lignes de codes, on observe par ligne et dans l'ordre un sujet, un prédicat et un objet. La ligne 02 par exemple, signifie que Arnaud connaît Béatrice. Les IRI sont encadrés par <> et chaque ligne se finit par un point qui signifie la fin d'un triplet. On peut remarquer à la ligne 03 « 1990-07-14 »^^, c'est une valeur appelé littéral. Les littéraux sont des valeurs de base associé à un type de donnée (propre aux RDF) non représentées par des IRI. En pratique, on utilise des centaines de ressources et donc on est obligé d'utiliser des outils de traitements de texte ligne à ligne.

Cependant, il faut admettre que N-triples n'est pas très lisible, c'est pourquoi il existe une extension appelée Turtle qui rend le langage largement plus agréable :

```
01 BASE <http://example.org/>
02 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
03 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
04 PREFIX schema: <http://schema.org/>
05 PREFIX dcterms: <http://purl.org/dc/terms/>
06 PREFIX wd: <http://www.wikidata.org/entity/>
07
08 <arnaud#me>
09   a foaf:Person ;
10   foaf:knows <béatrice#me> ;
11   schema:birthDate "1990-07-14"^^xsd:date ;
12   foaf:topic_interest wd:Q12418 .
13
14 wd:Q12418
15   dcterms:title "La Joconde" ;
16   dcterms:creator <http://fr.dbpedia.org/resource/Léonard_de_Vinci> .
17
18 <http://data.europeana.eu/item/04802/243FA8618938F4117025F17A8B813C5F9AA4D619>
19   dcterms:subject wd:Q12418 .
```

N-Triples connaît d'autres variations/extensions plus ou moins utiliser et utiles dans certaines situations.

Il existe évidemment de multiples façons de travailler avec les RDF et c'est aussi possible sur Python ! Grâce à la librairie rdflib, les RDF peuvent tout aussi bien être manipulées par les adeptes de Python. Essayons de manipuler par nous-même.

# LE PROCESSUS DE FONCTIONNEMENT DES DATAS

**Démonstration : manipulation des RDF avec la librairie rdflib dans  
Python**

Cf poly et documents de la démonstration

## 03 UN RESEAU RICHE D'INFORMATIONS SOURCE DE LIMITES

### **Création d'un réseau d'information riche, grâce aux Linked Data et au Web Sémantique**

Dans cette dernière partie, nous allons aborder le réseau des Link Open Data. Les Open Data font référence à des données qui peuvent être réutilisables, et les Linked Data sont des données compréhensibles par les machines et les individus, étant caractérisées par un lien étant auto-descriptif, comme précisé dans la première partie. Ainsi, les Linked Open Data sont des données connectées avec d'autres données de différentes sources. Grâce à leur nature, les données peuvent être comprises par les machines permettant de créer des liens logiques entre les données de différentes source. Ceci permet d'exploiter les données comme jamais cela n'avait été fait, puisque toutes les données accessibles vont pouvoir être utilisées et mises en commun entre elles.

Les données sont publiées sur le web, qui relie l'ensemble de données par des liens hypertextes sur internet. Le web est défini comme la notion communément employée pour parler du World Wide Web, comme précédemment dit crée par Tim Berners-Lee, traduit en français par la toile d'araignée. De plus, grâce à la création du web sémantique, qui permet de créer une interaction non seulement entre les données mais aussi entre les données et les utilisateurs.

Le web sémantique est un processus de réponse aux recherches des utilisateurs, permettant d'affiner les réponses. Grâce aux moteurs de recherche, l'utilisateur va être capable de faire une recherche ciblée. La réponse de cette recherche va être plus efficace grâce à la sémantique existant dans les documents du web. Les données vont être identifiées et analysées par rapport à l'information qu'elles ont.

Ainsi, le nombre de visiteur dans une page ou de certaines données va générer du trafic, qui va générer des statistiques plus élevées du web, augmentant ainsi les indicateurs de performance. Ceci permet d'évaluer la fiabilité des données qui vont être connectées à d'autres données ayant aussi une grande fiabilité. Ceci aboutit en la création d'un réseau d'information très riche et sélectif, incluant de multiples données triées par rapport à "l'effectivité /qualité" d'informations qu'elles possèdent.

# UN RESEAU RICHE D'INFORMATIONS SOURCE DE LIMITES

## Contraintes des Linked Open Data et leur futur

Le lien entre les Linked Open Data dans le web sémantique se fait à partir de leur sémantique attribuée. Or, nous pouvons nous questionner sur cette sémantique, soit la précision des ordinateurs lors qu'ils analysent l'information des données et la classifient. Ce concept est appelle "linkability" des data.

De plus, une grande partie de la fiabilité des données est basée sur les recherches des utilisateurs. Nous pouvons aussi douter de ceci, puisque le fait que des individus visitent beaucoup une page n'influe pas nécessairement sur la qualité de l'information des données.

Les Linked Open Data créent un environnement de données très hétérogène, liant des données de différentes sources, ceci pouvant tendre vers une hybridation et une contamination de données. Nous pouvons retrouver des données ayant des informations riches, liées à d'autres ayant une moindre qualité. Autrement dit, il peut y avoir des problèmes de qualité des données reliées à un liant, qui peut être inconsistant ou inexacte. De plus, comme la quantité de données publiés ne cesse d'augmenter, ce problème de lien va être de plus en plus commun. Il va y avoir une abondance de données sur des sujets précis, et faire des liens efficaces va devenir de plus en plus contraignant. Ce problème est aussi relié au fait que les informations des données changent extrêmement rapidement, et nous pouvons retrouver des bases données ayant des informations caduques. Certaines données peuvent aussi représenter le monde, ce qui à la fois est en constante évolution. Ainsi, beaucoup de données peuvent se retrouver obsolètes. Ceci représente la possible inexactitude des Linked Open Data, et même de bases de données ayant des informations fausses.

Par rapport à l'investigation, il y a eu aussi un changement radical. Les informations recueillis dans les bases universitaires ( bibliothèques ) de chaque université vont pouvoir être utilisées pour d'autres projets de recherche dans des groupes de recherche partout dans le monde. Or, ceci ouvre un nouveau débat sur la propriété intellectuelle : comment pouvons-on utiliser les informations publiées sans commettre de la fraude. Ainsi nous avons pu observer comment ces dernières années des politiques contre le plagiat ont été mises en place pour lutter contre ce danger. Dans la même optique, nous avons vu ces dernières années l'apparition des NFT's ( non fungible tokens ) qui se caractérisent par un certificat numérique de propriété qui précise clairement qui possède le jeton. Or, ceci reste toujours un enjeux à surmonter pour les Linked Open Data.

Depuis 2014, de multiples institutions publiques comme les universités et les bibliothèques nationales ( Allemagne , France , Grande Bretagne ) ont publié leur données. Les Linked Open Data sont considérées comme le futur des données



bibliographiques. Ils ont pour objectif de créer des interfaces où les utilisateurs peuvent accéder de “façon” facile aux données. Le British National Bibliography a publié en 2011 des Linked Open Data, or cela a été publié dans une application qui présente de grandes limites, qui ne sont pas bien adaptés aux recherches humaines. Par exemple, l'accès aux données en fonction du titre ou de l'auteur de l'ouvrage n'aboutit pas à une recherche assez précise.

L'hétérogénéité des propres bases de données suppose aussi un enjeu à surmonter. La grande partie des bases de données se trouvent sur le LOD Cloud, WikiData etc, qui sont des plateformes hétérogènes. Un objectif serait de créer des liens entre ses bases de données et unifier efficacement leurs données.

La recherche actuelle sur les linked open data est basé sur le triage des données pour créer des liens intéressants.

# ANNEXE

## BIBLIOGRAPHIE:

<https://perso.liris.cnrs.fr/pierre-antoine.champin/enseignement/tiw5/rdf-lod.pdf>

<https://fr.slideshare.net/antoinecourtin/brve-introduction-au-linked-open-data-appliqu-aux-institutions-culturelles>

<https://recherchemid.wordpress.com/2013/11/11/vous-avez-dit-linked-open-data/>

[https://www.researchgate.net/figure/Overview-of-the-RDF-data-slicing-approach-1-dataset-dumps-are-accessed-on-the-Web-2\\_fig2\\_240615179](https://www.researchgate.net/figure/Overview-of-the-RDF-data-slicing-approach-1-dataset-dumps-are-accessed-on-the-Web-2_fig2_240615179)

<https://www.youtube.com/watch?v=uju4wT9uBIA>

Archives Européana sur les link open data en fonction du secteur associé :

<https://www.europeana.eu/fr/search?page=1&view=grid&query=linked%20open%20data>

Ted show du créateur des linked open data, Tim Berners-Lee:

[https://www.youtube.com/watch?v=OM6XIIcm\\_qo](https://www.youtube.com/watch?v=OM6XIIcm_qo)

Description des bases de connaissances :

<https://www.atlassian.com/fr/itsm/knowledge-management/what-is-a-knowledge-base>

Web sémantique et moteurs de recherche et application du web sémantique :

<https://www.imdeo.com/web-3-0-le-web-semantique/>

Définition intéropérabilité :

<https://www.techno-science.net/glossaire-definition/Interoperabilite.html>

RDF :

[https://fr.wikipedia.org/wiki/Triplet\\_RDF](https://fr.wikipedia.org/wiki/Triplet_RDF)

[https://data.europa.eu/sites/default/files/d2.1.2\\_training\\_module\\_1.2\\_introduction\\_to\\_linked\\_data\\_fr\\_edp.pdf](https://data.europa.eu/sites/default/files/d2.1.2_training_module_1.2_introduction_to_linked_data_fr_edp.pdf)

Les cinq datas : comment mieux publier des données :

<https://5stardata.info/fr/>

Bibliographie 3<sup>ème</sup> partie :

<https://tel.archives-ouvertes.fr/tel-01742957v2/document>

[https://e-archivo.uc3m.es/bitstream/handle/10016/19674/avila\\_aplicacion\\_TFM\\_2014.pdf?sequence=1&isAllowed=y](https://e-archivo.uc3m.es/bitstream/handle/10016/19674/avila_aplicacion_TFM_2014.pdf?sequence=1&isAllowed=y)

Beyond the pillar of Hercules : Linked data and cultural Heritage :

<file:///Users/marianamolldealba/Downloads/>

<https://www.societybyte.swiss/2017/02/01/linked-data-present-future/>

[https://joinup.ec.europa.eu/sites/default/files/document/2015-05/d2.1.2 training module 1.2 introduction to linked data v1.00 fr.pdf](https://joinup.ec.europa.eu/sites/default/files/document/2015-05/d2.1.2_training_module_1.2_introduction_to_linked_data_v1.00_fr.pdf)

[https://doc.rero.ch/record/209598/files/M7-2014 memoire HUGI-PRONGUE.pdf](https://doc.rero.ch/record/209598/files/M7-2014_memoire_HUGI-PRONGUE.pdf)

<https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1384&context=charleston>