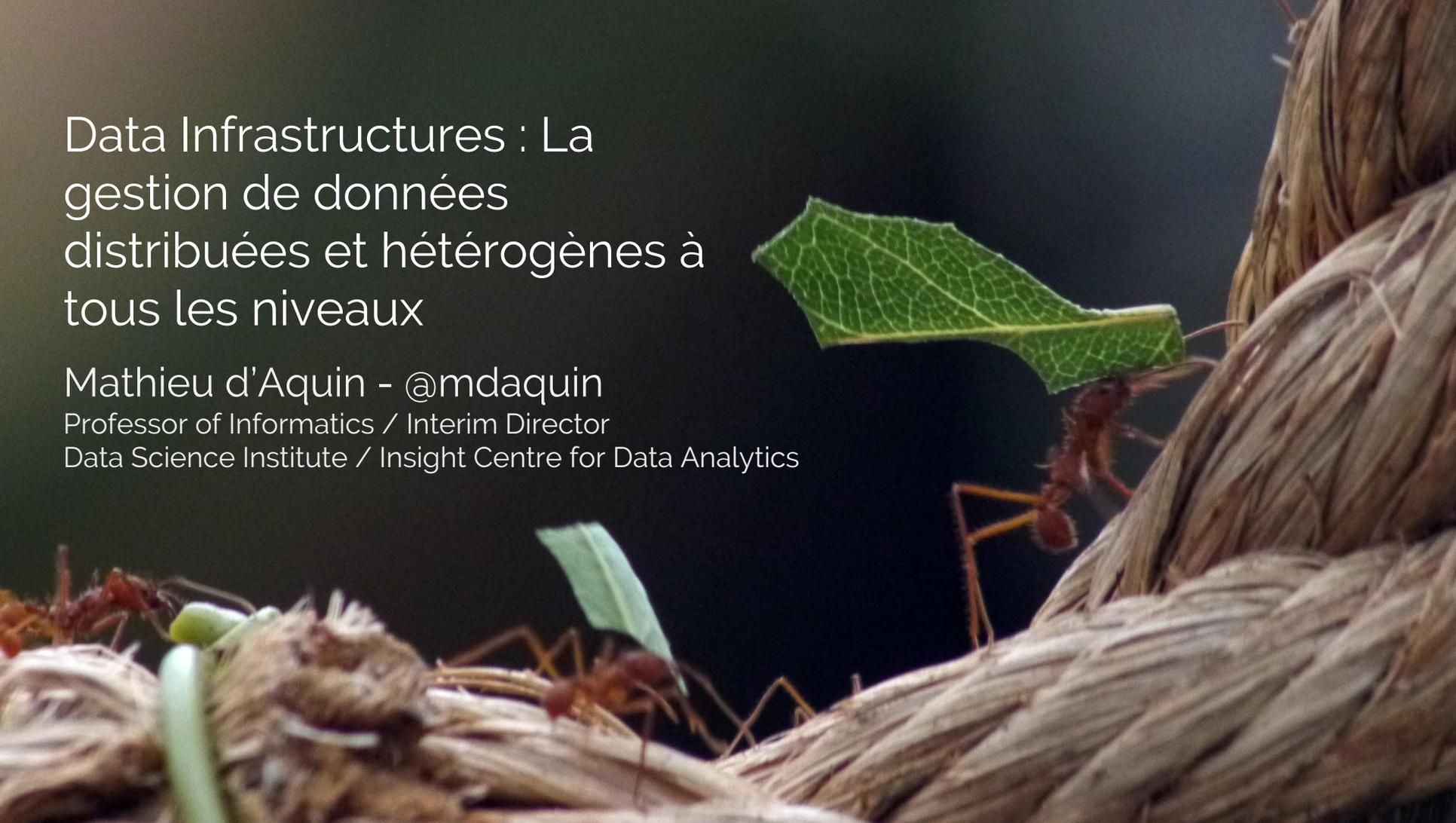
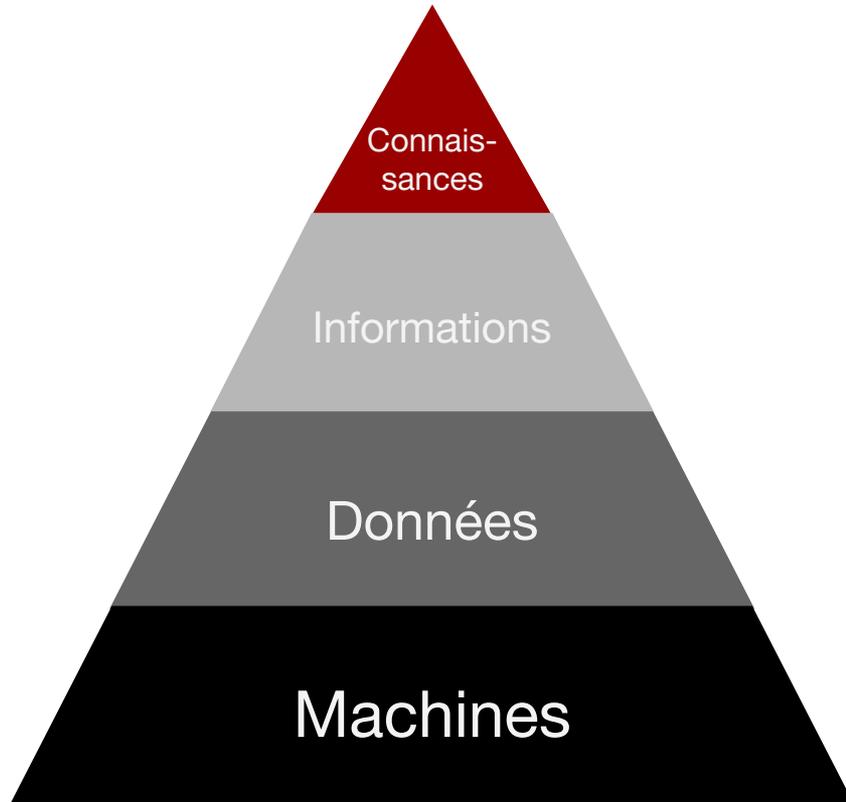


Data Infrastructures : La gestion de données distribuées et hétérogènes à tous les niveaux

Mathieu d'Aquin - @mdaquin
Professor of Informatics / Interim Director
Data Science Institute / Insight Centre for Data Analytics



Pyramide de l'information

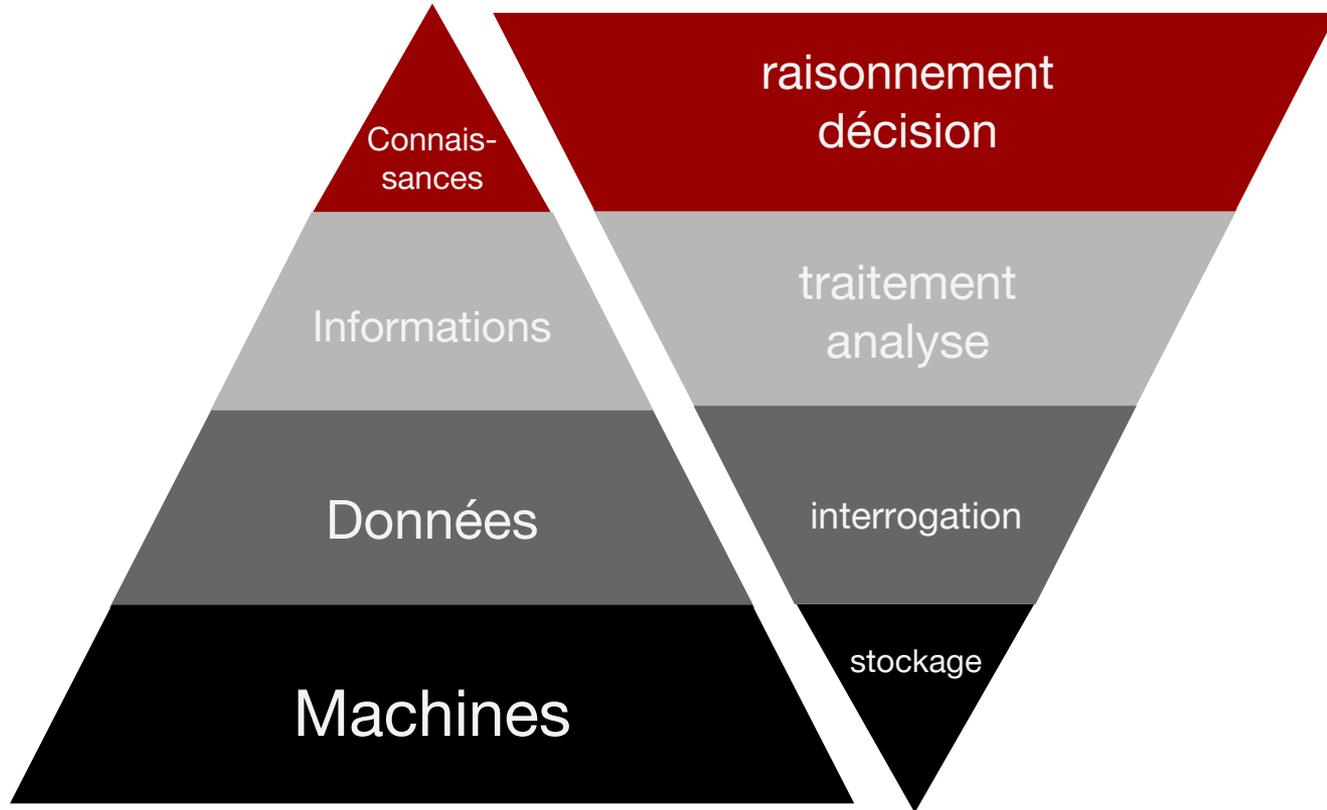


Information is not
knowledge. Knowledge
is not wisdom.
Wisdom is not truth.
Truth is not beauty.
Beauty is not love.
Love is not music.
Music is the best”

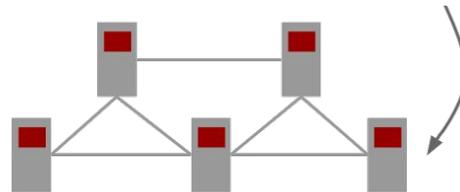
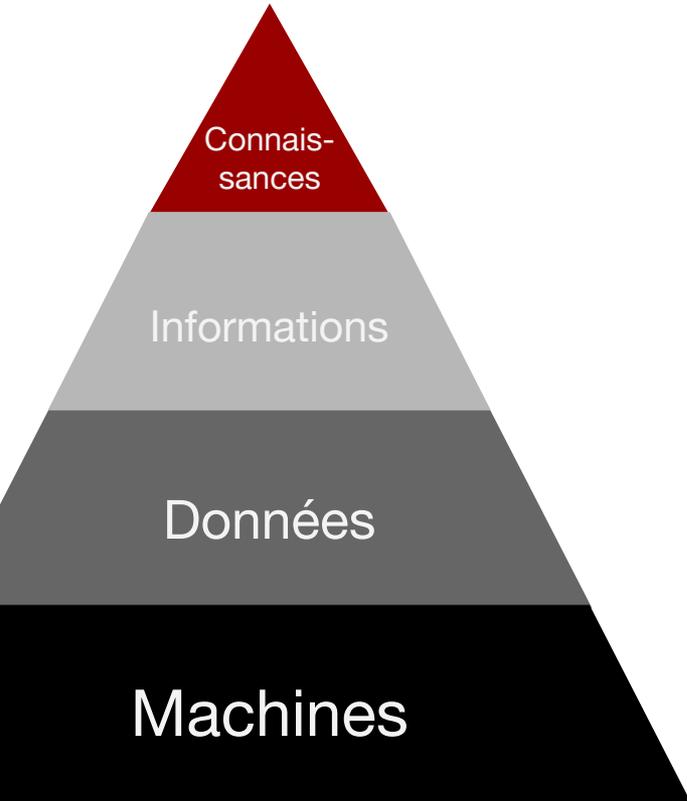
- Frank Zappa



Pyramide de l'information



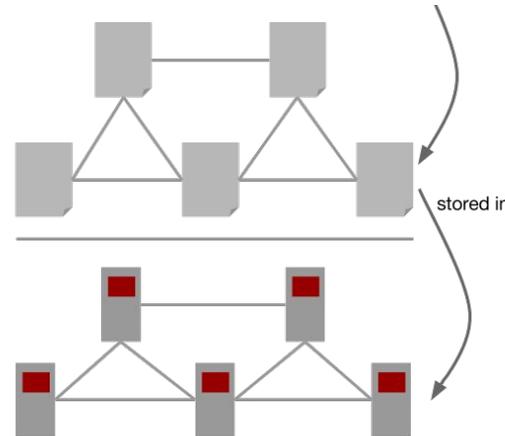
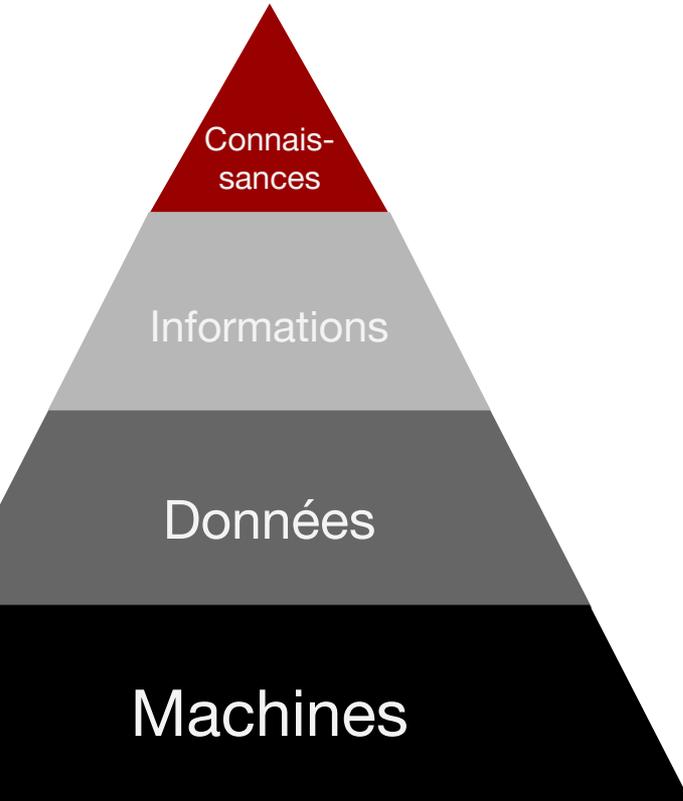
Pyramide de l'information et le Web (Sémantique)



Internet
Réseau de machines



Pyramide de l'information et le Web (Sémantique)

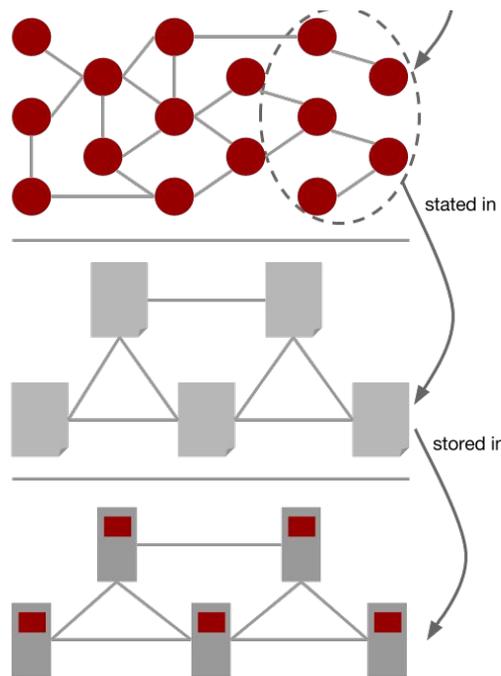
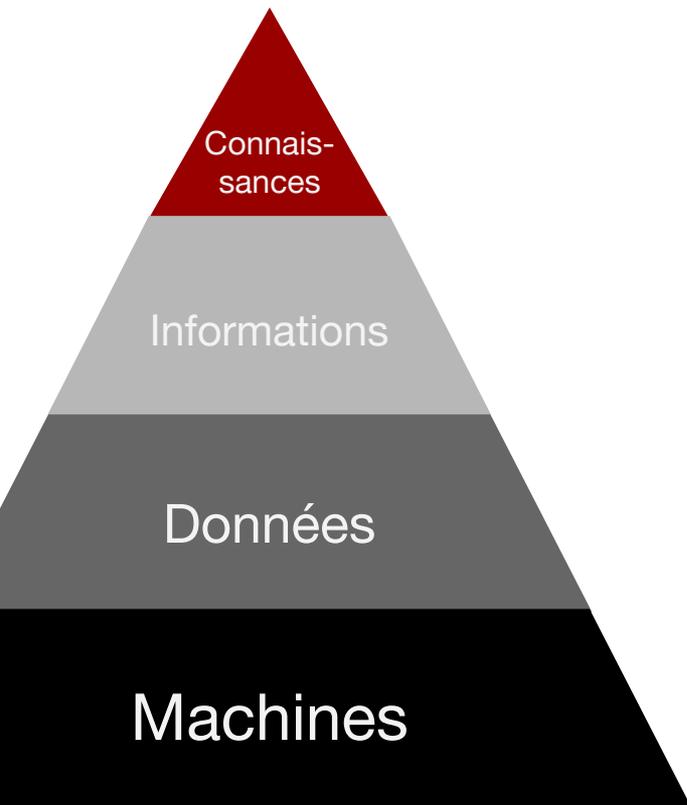


Le Web
Réseau de documents

Internet
Réseau de machines



Pyramide de l'information et le Web (Sémantique)



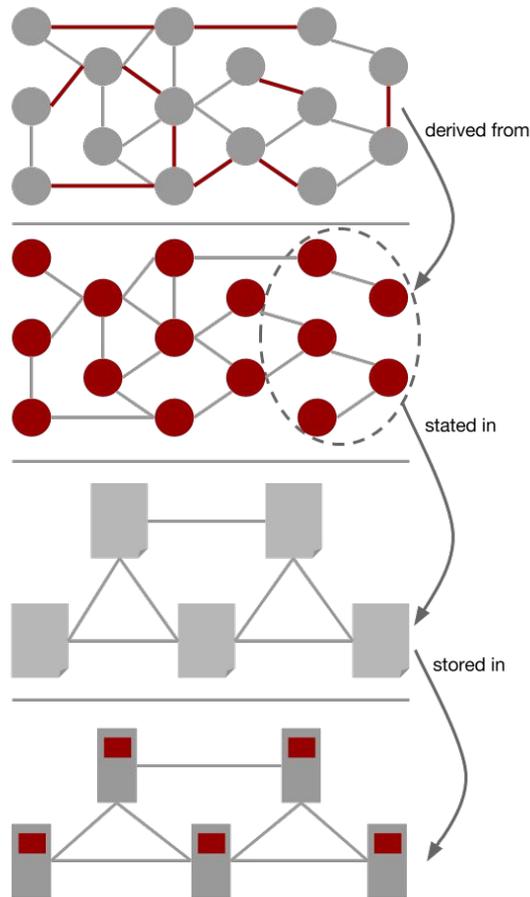
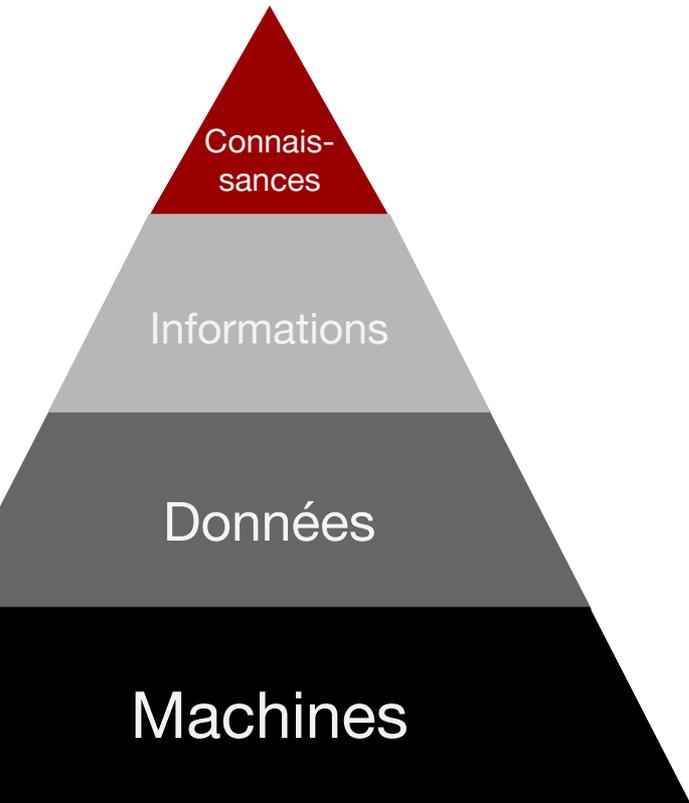
Les données liées
Réseau d'information

Le Web
Réseau de documents

Internet
Réseau de machines



Pyramide de l'information et le Web (Sémantique)



Le web sémantique
Réseau de connaissances

Les données liées
Réseau d'information

Le Web
Réseau de documents

Internet
Réseau de machines





MORGAN & CLAYPOOL PUBLISHERS

The Epistemology of Intelligent Semantic Web Systems

Mathieu d'Aquin
Enrico Motta

SYNTHESIS LECTURES ON

THE SEMANTIC WEB: THEORY AND TECHNOLOGY

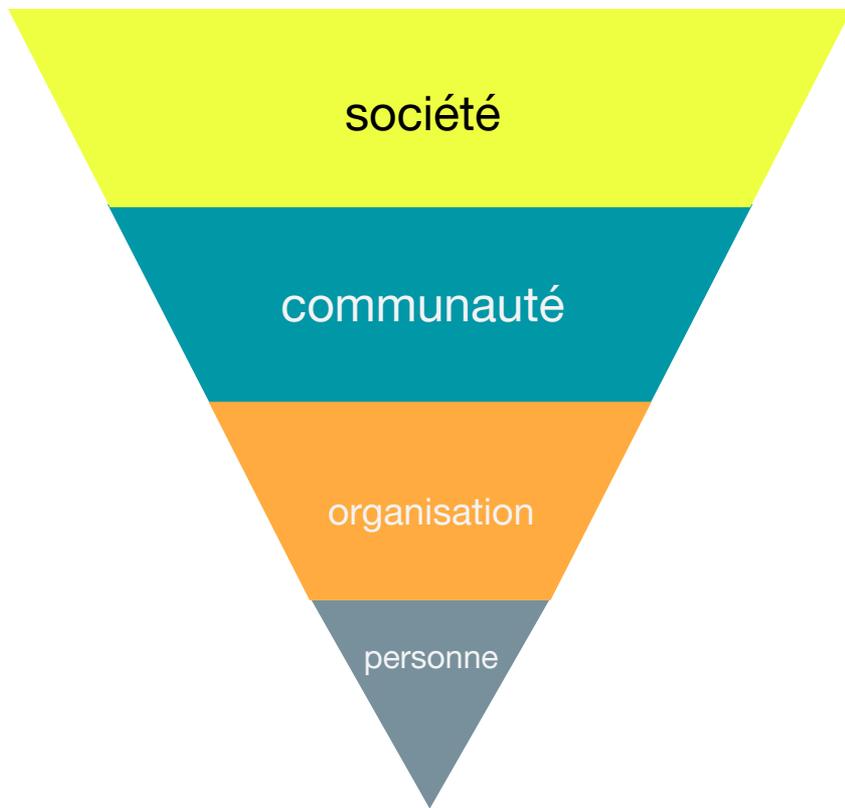
Ying Ding and Paul Groth, Series Editors

MORGAN & CLAYPOOL PUBLISHERS

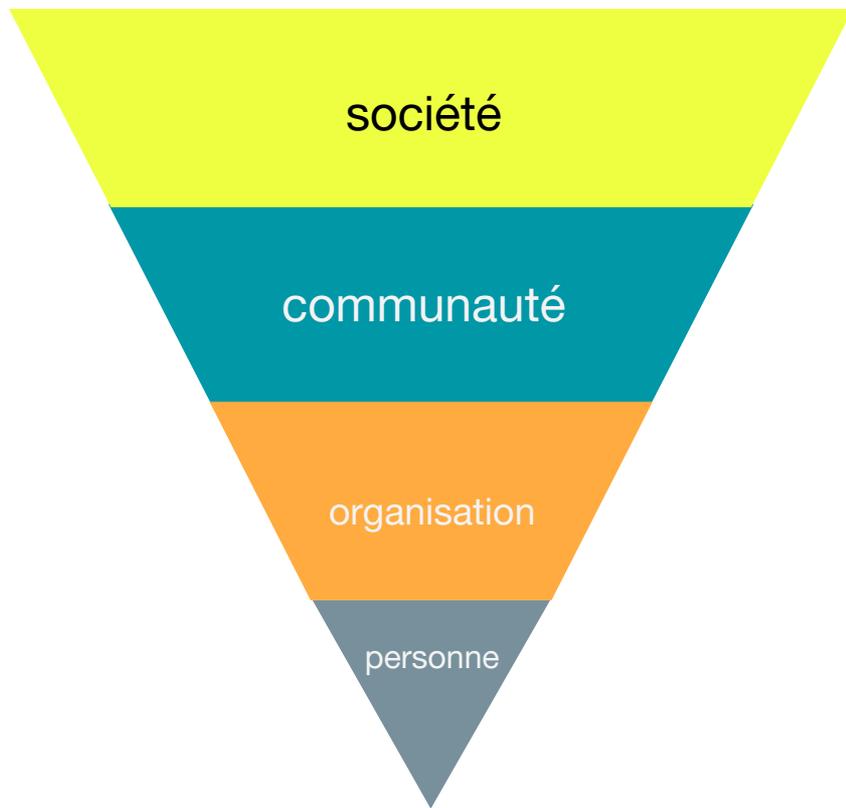
Epistemology of Intelligent Semantic Systems

AND TECHNOLOGY

Pyramide des implications du traitement de l'information



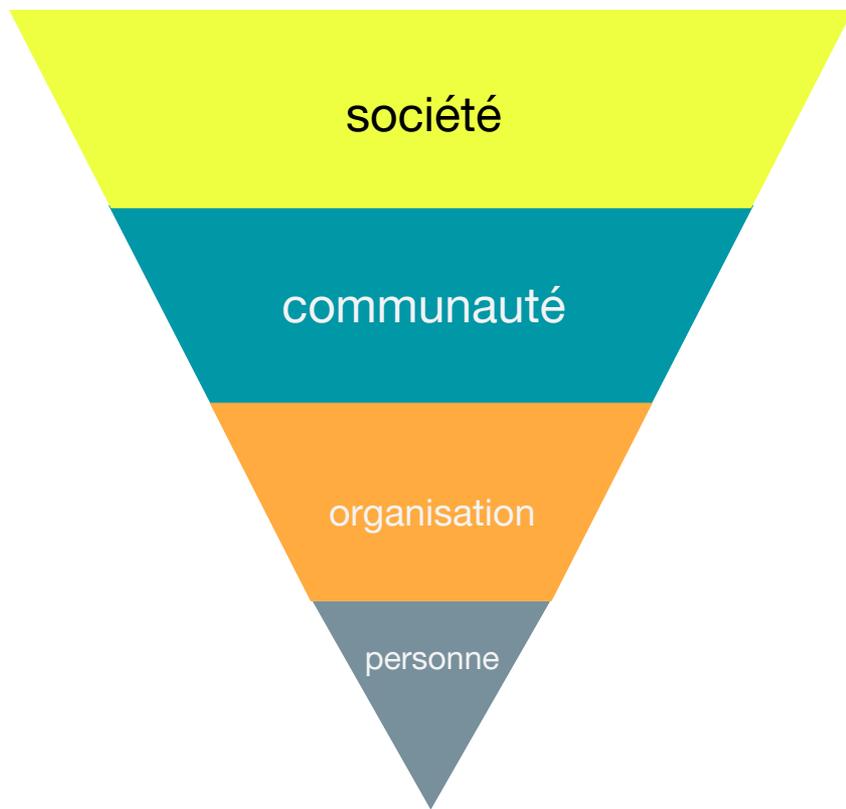
Pyramide des implications du traitement de l'information



Le traitement de l'information touche une personne, un individu.



Pyramide des implications du traitement de l'information

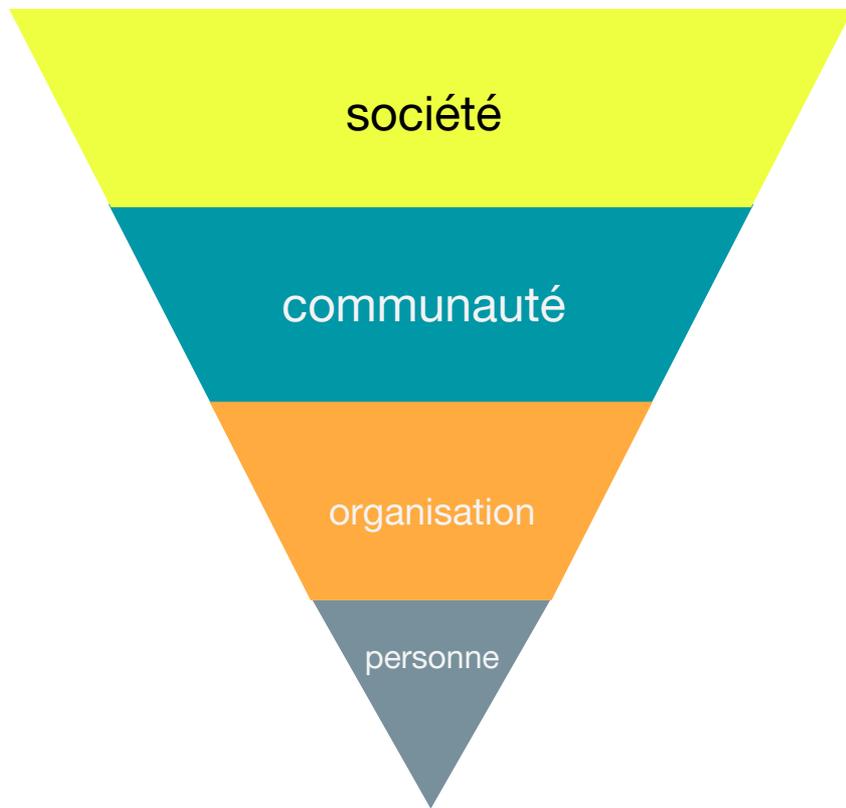


Le traitement de l'information touche une compagnie, une entreprise, une institution, etc.

Le traitement de l'information touche une personne, un individu.



Pyramide des implications du traitement de l'information



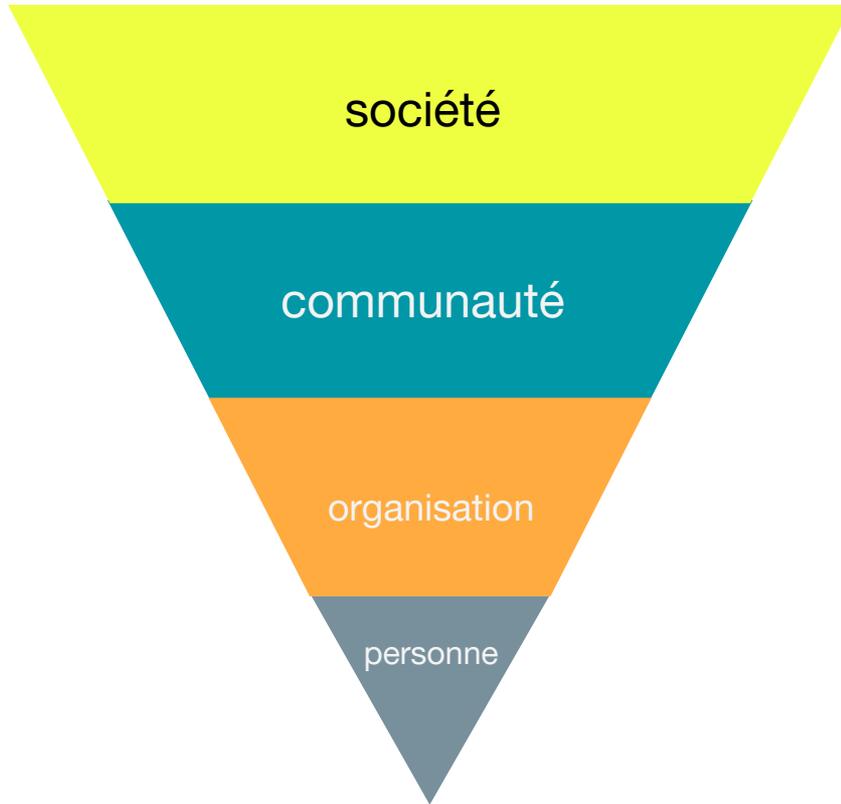
Le traitement de l'information touche des groupes sociaux, regroupements géographiques, etc.

Le traitement de l'information touche une compagnie, une entreprise, une institution, etc.

Le traitement de l'information touche une personne, un individu.



Pyramide des implications du traitement de l'information



Le traitement de l'information touche la démocratie, la liberté, l'égalité, la justice, l'éthique, etc.

Le traitement de l'information touche des groupes sociaux, regroupements géographiques, etc.

Le traitement de l'information touche une compagnie, une entreprise, une institution, etc.

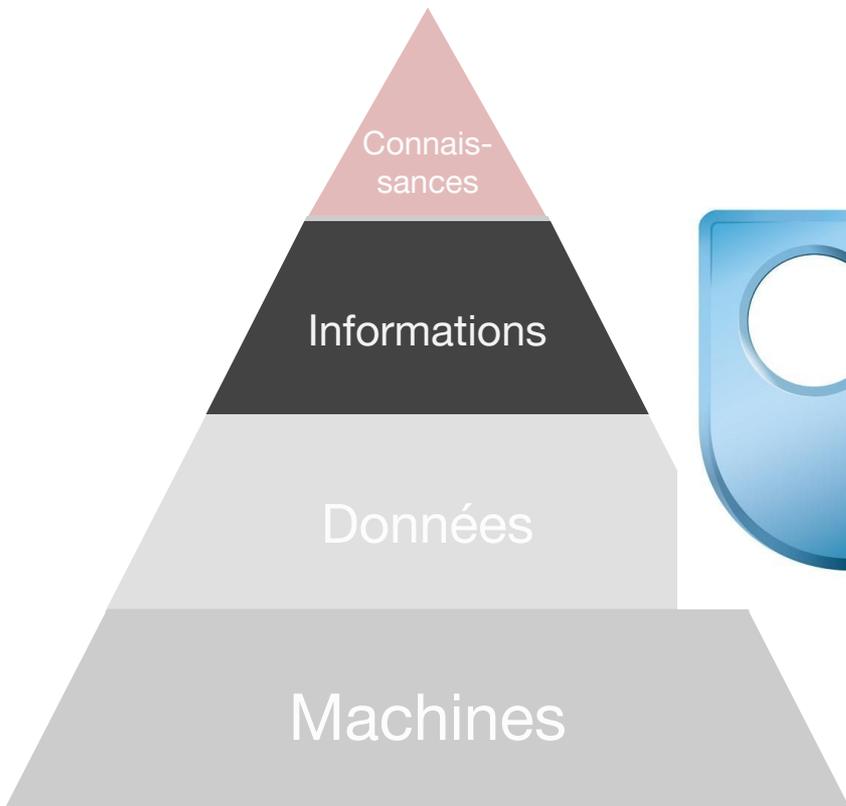
Le traitement de l'information touche une personne, un individu.



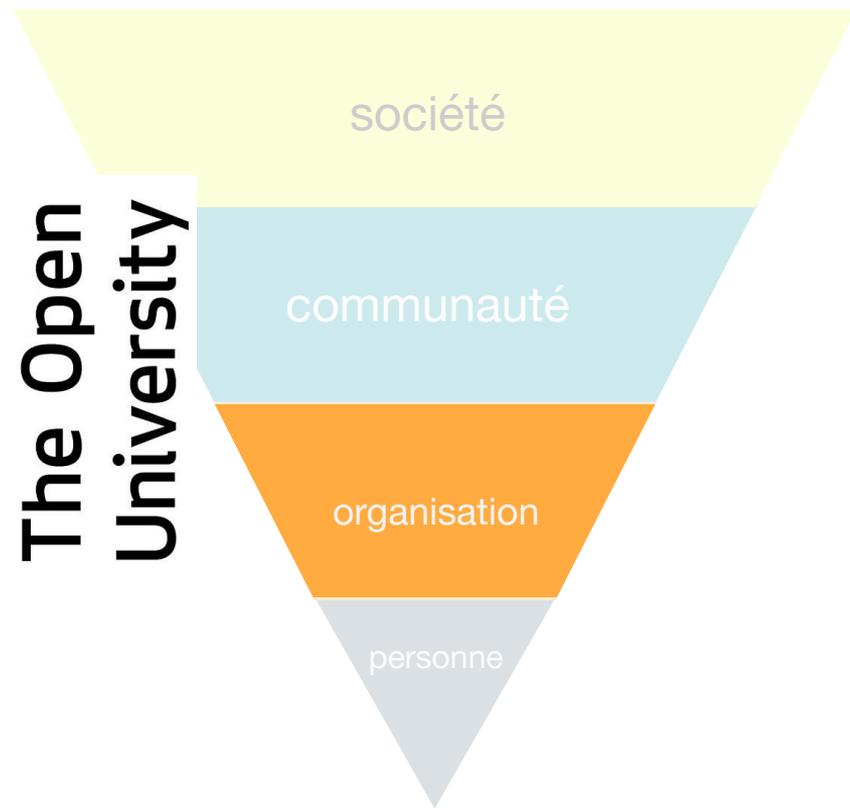
Des exemples d'infrastructures de données à différents niveaux



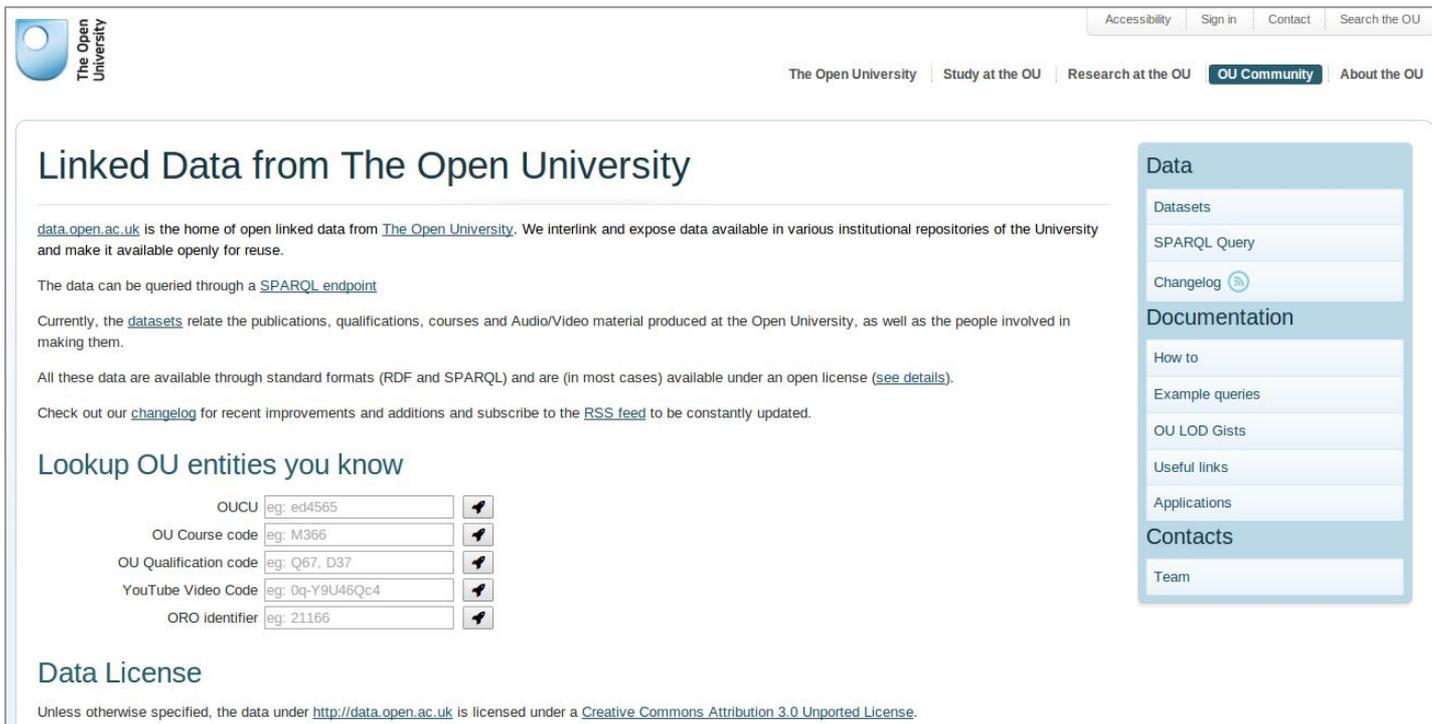
data.open.ac.uk



**The Open
University**



data.open.ac.uk



The screenshot shows the homepage of data.open.ac.uk. At the top left is the Open University logo. The top navigation bar includes links for Accessibility, Sign in, Contact, and Search the OU. Below this is a secondary navigation bar with links for The Open University, Study at the OU, Research at the OU, OU Community (highlighted), and About the OU. The main heading is "Linked Data from The Open University". The content area contains several paragraphs of text explaining the site's purpose, how to query data via SPARQL, and where to find documentation. A "Lookup OU entities you know" section features five input fields with example values and share icons. A "Data License" section at the bottom states that the data is licensed under a Creative Commons Attribution 3.0 Unported License. On the right side, there is a vertical menu with sections for Data (Datasets, SPARQL Query, Changelog), Documentation (How to, Example queries, OU LOD Gists, Useful links, Applications), and Contacts (Team).

 Accessibility | Sign in | Contact | Search the OU

The Open University | Study at the OU | Research at the OU | **OU Community** | About the OU

Linked Data from The Open University

data.open.ac.uk is the home of open linked data from [The Open University](http://www.open.ac.uk). We interlink and expose data available in various institutional repositories of the University and make it available openly for reuse.

The data can be queried through a [SPARQL endpoint](#)

Currently, the [datasets](#) relate the publications, qualifications, courses and Audio/Video material produced at the Open University, as well as the people involved in making them.

All these data are available through standard formats (RDF and SPARQL) and are (in most cases) available under an open license ([see details](#)).

Check out our [changelog](#) for recent improvements and additions and subscribe to the [RSS feed](#) to be constantly updated.

Lookup OU entities you know

| | | |
|-----------------------|--|---|
| OUCU | <input type="text" value="eg: ed4565"/> |  |
| OU Course code | <input type="text" value="eg: M366"/> |  |
| OU Qualification code | <input type="text" value="eg: Q67, D37"/> |  |
| YouTube Video Code | <input type="text" value="eg: 0q-Y9U46Qc4"/> |  |
| ORO identifier | <input type="text" value="eg: 21166"/> |  |

Data License

Unless otherwise specified, the data under <http://data.open.ac.uk> is licensed under a [Creative Commons Attribution 3.0 Unported License](#).

Data

- Datasets
- SPARQL Query
- Changelog 

Documentation

- How to
- Example queries
- OU LOD Gists
- Useful links
- Applications

Contacts

- Team

<http://data.open.ac.uk/course/m366>

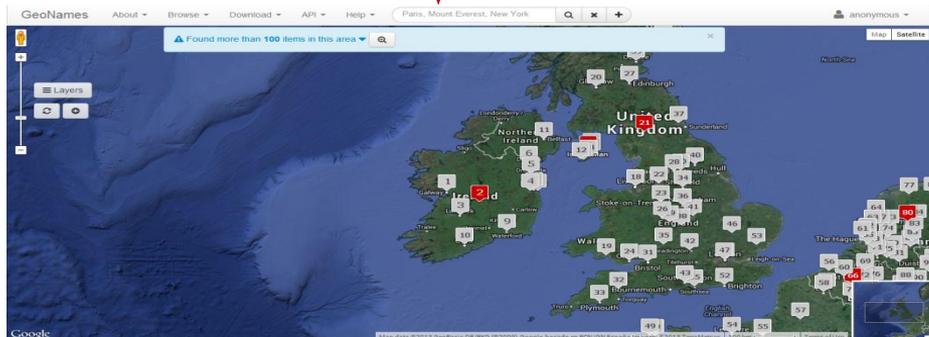
http://data.open.ac.uk/organization/the_open_university

Natural and artificial intelligence

<http://data.open.ac.uk/course/m366>

| Property | Object |
|-----------------------|---|
| Type | Course Module Learning opportunity specification Course |
| Label | Natural and artificial intelligence |
| Has assessment method | Computer-marked assignments (Online) Examination Tutor-marked assignments |
| Has courseware | M366 Course software Natural intelligence Intelligent machines Reflections Evolutionary computation Symbolic intelligence Neural networks Neural networks M366 Course software Natural intelligence Symbolic intelligence Evolutionary computation |
| Has title | Natural and artificial intelligence |
| Is taught present | true |

mlo:location



<http://sws.geonames.org/2963597/> (Ireland)

The Open University

http://data.open.ac.uk/organization/the_open_university

| Property | Object |
|----------|---|
| Type | http://purl.org/net/mlo/LearningOpportunityProvider http://xcon.org/profiles/catalog/1.2/provider |
| Offers | Reflecting on professional learning in education Working together for children Leadership and management in health and social care Diploma in Natural Sciences Continuing professional development in practice Postgraduate Diploma in Online and Distance Education Certificate of Higher Education in Social Care (Wales) BSc (Honours) Mathematics and Its Learning Structural integrity: designing against failure Managing in the workplace Postgraduate Certificate in Education: Geography Postgraduate Certificate in Education: Science (Physics) Introduction to financial services BA (Honours) Language Studies Chemical change and environmental applications Professional practice portfolio Professional Graduate Certificate in Education: MFL - French Certificate in Legal Studies Concept to clinic Understanding children's development and learning BSc (Honours) Psychology Diploma of Higher Education in Social Care (England) |

mlo:offers

owl:sameAs

HM Government

Linked Data API

The Open University

<http://education.data.gov.uk/id/school/133849>

| | |
|-------------------------|---|
| PFI | 133849 |
| unique reference number | 133849 |
| address | address 1 Walton Hall postcode MK7 6AA region Buckinghamshire town Milton Keynes |
| gender | Mixed |
| type of establishment | Higher Education Institutions |

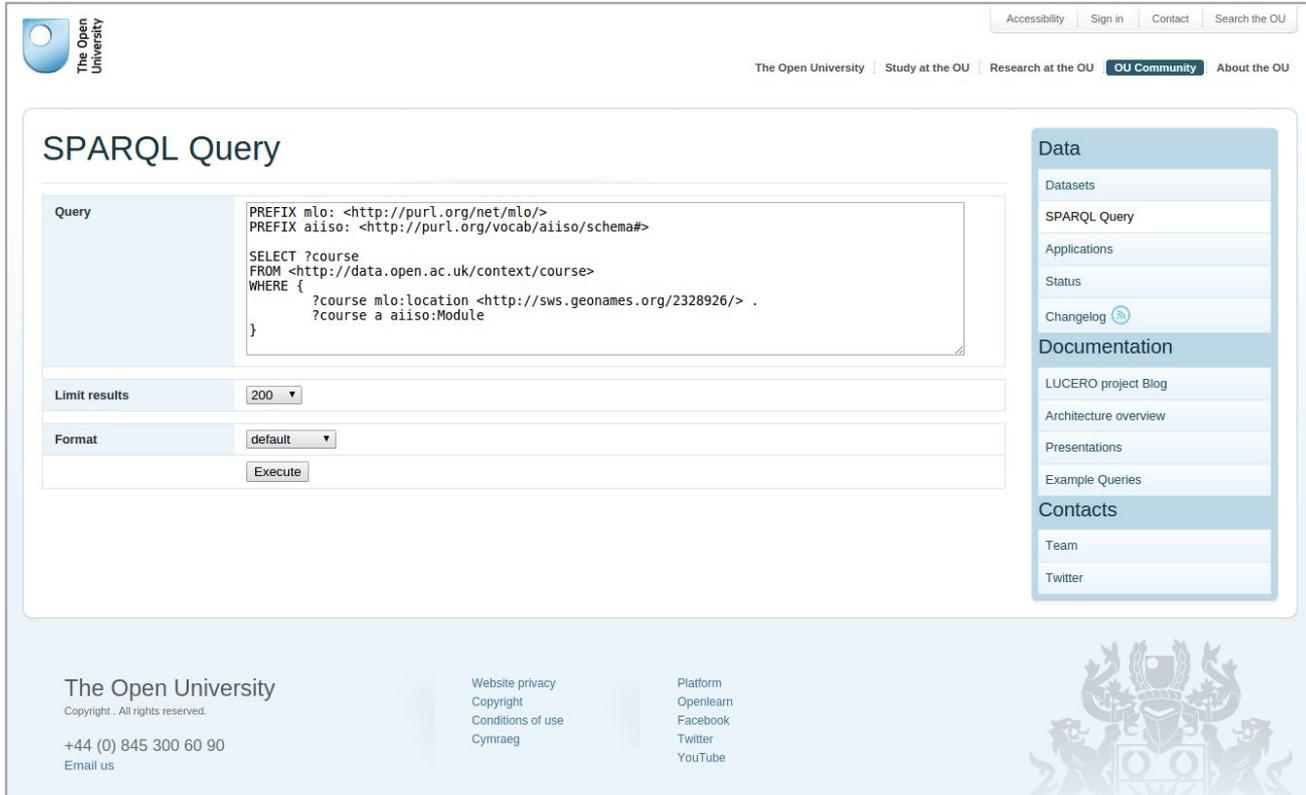
Browse: Census, School, Secondary schools, Primary schools, All schools

View: admin, location, performance, provision

PFI: unique reference number, address, postcode, address, region, address, town, gender, type of establishment

<http://education.data.gov.uk/id/school/133849>

Data -> Interrogation



The screenshot shows the Open University's SPARQL Query interface. At the top left is the Open University logo. The top right contains navigation links: Accessibility, Sign in, Contact, and Search the OU. Below this is a secondary navigation bar with links for The Open University, Study at the OU, Research at the OU, OU Community (highlighted), and About the OU.

SPARQL Query

Query

```
PREFIX mlo: <http://purl.org/net/mlo/>
PREFIX aiiso: <http://purl.org/vocab/aiiso/schema#>

SELECT ?course
FROM <http://data.open.ac.uk/context/course>
WHERE {
    ?course mlo:location <http://sws.geonames.org/2328926/> .
    ?course a aiiso:Module
}
```

Limit results 200

Format default

Execute

Data

- Datasets
- SPARQL Query
- Applications
- Status
- Changelog

Documentation

- LUCERO project Blog
- Architecture overview
- Presentations
- Example Queries

Contacts

- Team
- Twitter

The Open University
Copyright . All rights reserved.
+44 (0) 845 300 60 90
Email us

Website privacy
Copyright
Conditions of use
Cymraeg

Platform
Openlearn
Facebook
Twitter
YouTube



Applications

Mobile and Personal Semantics

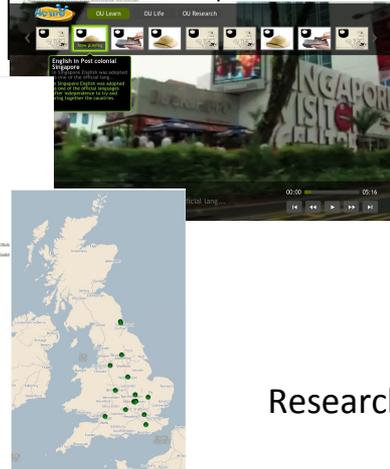


Social

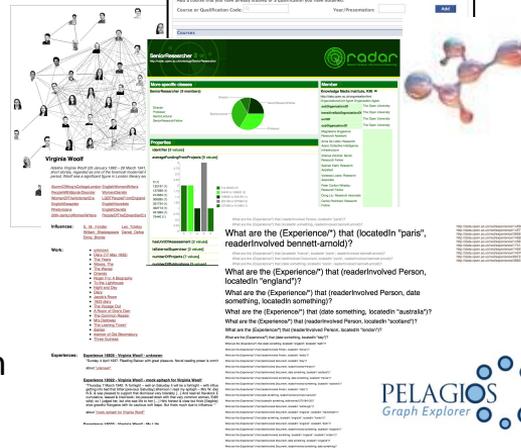
Resource Discovery



Exploration



Research



Au delà d'une organisation

Linked Universities

[Home](#) [About](#) [Members](#) [Vocabularies](#) [Tools](#) [Datasets and Endpoints](#) [Publications](#) [Search](#)



Welcome to LinkedUniversities.org

Linked Universities is an alliance of european universities engaged into exposing their public data as linked data. [Linked data](#) is a set of principles to put raw data on the Web, making them Web addressable and linkable, so that they can be easily accessed, discovered, connected and reused. The idea is that data from different institutions and organisations can contribute to a common data space on the Web: the [Web of Data](#).

There are only a few universities currently exposing their public data as linked data, using technologies such as RDF and SPARQL to give direct access to information such as their publications, courses, educational material, etc. These initiatives are currently often disconnected from each other. For every new site being developed, a lot of the efforts required to build a university linked data platform need to be re-deployed, and many of the lessons need to be re-learned. Also, we believe that the potential for linked data in education and research goes well beyond the individual benefit for each institution, as this potential can only be achieved through providing cross-university data that can be aggregated, integrated and compared. While linked data, much like the Web, relies on transparent distribution and a certain amount of self-organisation, we believe that sharing practices and collaborating in the development of university linked data platforms can significantly help towards this common goal, i.e., the creation of a Web of university data.

This portal should therefore essentially be seen as a collaborative space, where institutions and individuals involved in the exposure of university linked data can describe, share and reuse common vocabularies and practices. Our goals are therefore to:

- Identify, support and develop common linked data vocabularies, usable across universities for common concepts such as courses, qualifications, educational material, etc.
- Describe reusable recipes, and share reusable tools, for exposing linked data in universities
- Support, through experience sharing and reuse, initiatives towards exposing university data as linked data

If you want to obtain more information about *Linked Universities*, discuss your experience with us or become a member, please send an e-mail to info@linkeduniversities.org

Linked Universities

[Home](#) [About](#) [Members](#) [Vocabularies](#) [Tools](#) [Datasets and Endpoints](#) [Publications](#) [Search](#)



Vocabularies

This page list vocabularies commonly used to describe data related to universities, together with an explanation of the main concepts involved, and links to recipes showing how to express certain situations using these vocabularies.

Course and qualification description

MLO - Metadata for Learning Opportunities

Metadata for Learning Opportunities (MLO) is a European standardized model addressing metadata sufficient for advertising a learning opportunity. The MLO standard is also designed to facilitate semantic technologies and web architectures to support several mechanisms for exchange of the information and aggregation of information by third party service suppliers.

An RDF-S version of MLO has been [developed](#) by [Scott Wilson](#) and is currently available at <http://svn.ceis.ac.uk/xcr/trunk/bindings/mlo/> using the namespace <http://purl.org/net/mlo/>

See for example the use of MLO in course descriptions at the Open University: <http://data.open.ac.uk/course/204>

XCRI-CAP - XCRI Course Advertising Profile

XCRI is UK-oriented project to establish a specification to support the eXchange of Course-Related Information. A key activity for XCRI is the development of an XML specification, the XCRI Course Advertising Profile (or XCRI-CAP for short). Learning providers can publish their courses information in the standard XCRI-CAP format, so that it can be collected easily by organisations with course search services. Opening up the offerings of learning providers creates new possibilities for value-added services and information channels for universities, colleges, and training providers.

An RDF-S version of XCRI has been [developed](#) by [Scott Wilson](#) and is currently available at <http://svn.ceis.ac.uk/xcr/trunk/bindings/cap/> using the namespace http://xcri.org/profiles/catalog1_2/

See for example the use of XCRI in course descriptions at the Open University: <http://data.open.ac.uk/course/204>

TEACH - Teaching Core Vocabulary

TEACH, the Teaching Core Vocabulary, is a lightweight vocabulary providing terms to enable teachers to relate things in their courses together. The Teaching Core Vocabulary is based on practical requirements set by providing seminar and course descriptions as Linked Data. The vocabulary specification is available at <http://linkedscience.org/teach/> and it uses this same URI as the namespace. The suggested namespace prefix is "teach".

See an example of the use of TEACH in course descriptions at the University of Muenster: <http://data.uni-muenster.de/coretext/page/csa/course/171>

University as an organisation

AIISO - Academic Institution Internal Structure Ontology

The Academic Institution Internal Structure Ontology (AIISO) provides classes and properties to describe the internal organizational structure of an academic institution.

See for example the use of AIISO in course descriptions at the Open University: <http://data.open.ac.uk/course/204>

Bowlogna Ontology

The Bowlogna ontology originates from a lexicon defining terms related to the [Bologna Process](#) and aims at providing a standard schema for European universities involved in the Bologna Reform of higher-education studies.

Core organization ontology

This ontology is not specific to universities, but aim at supporting linked-data publishing of organizational information across a number of domains. It is designed to allow domain-specific extensions to add classification of organisations and roles, as well as extensions to support neighbouring information such as organisational activities.

Academic Publications and Communities

BIBO - The Bibliographic Ontology

This volume... in Education Seminar...
In the first part of the book, two chapters... the current use of linked and open data in education... technology and the topics that are being covered.
The second part of the book focuses on the specific, practical... that are being put in place to exploit open and linked data... today.

The goal of this book is to provide a snapshot of current... share and disseminate the growing collective experience... data in education. This volume brings together real-world... practical endeavors from initiatives spread across... the world. These initiatives are laying the foundation... data in the education movement and leading to... applications.

State-of-the-Art Survey

LNCS 9500

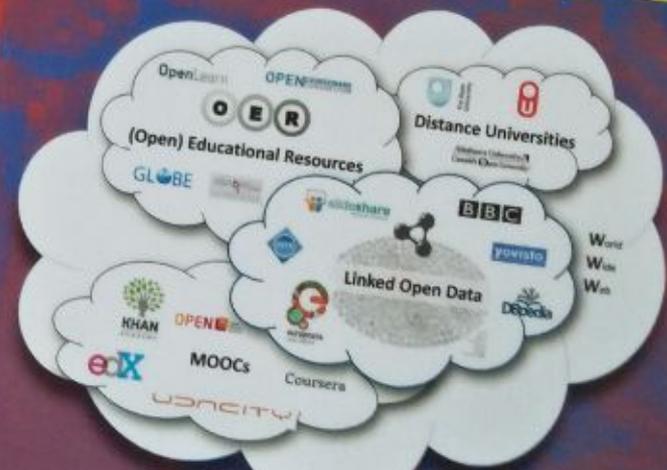
Open Data for Education

Linked, Shared, and Reusable Data for Teaching and Learning

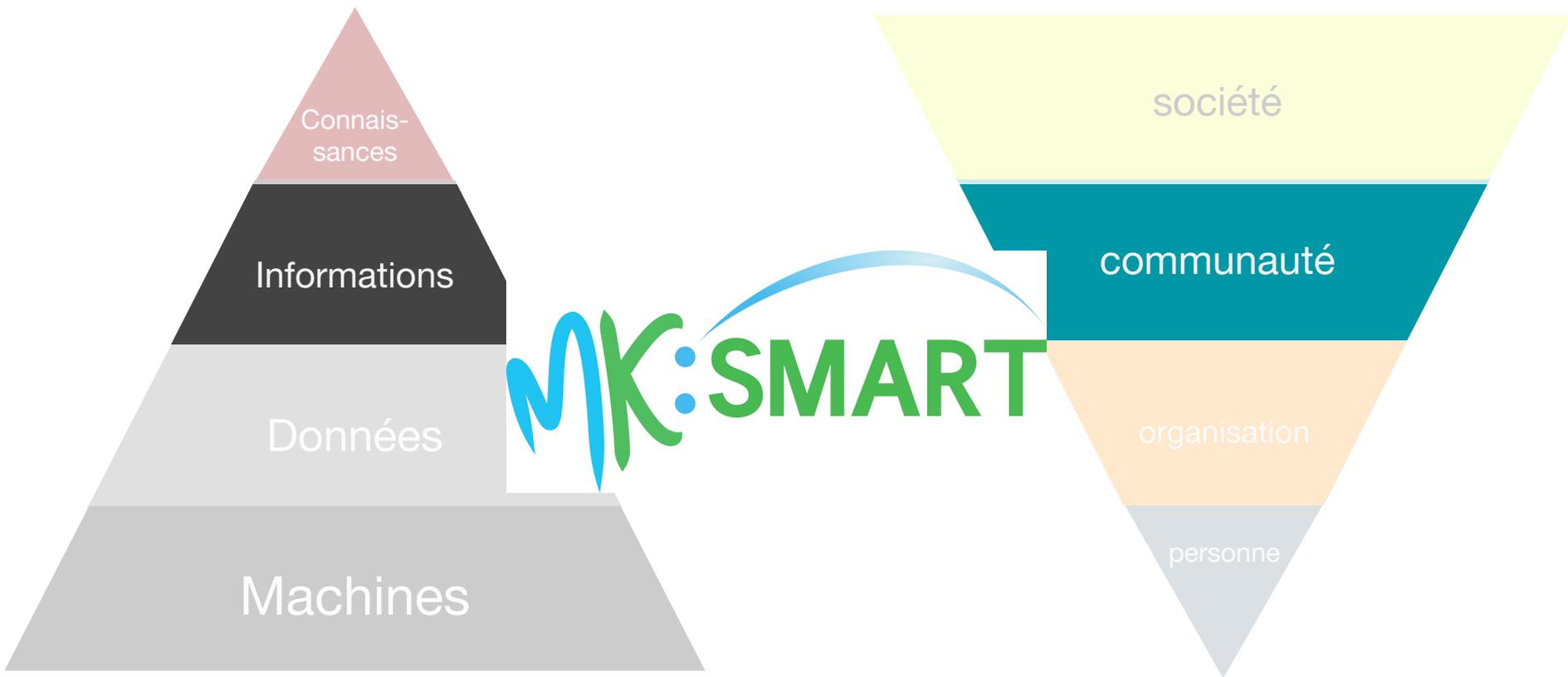
In parallel to the printed book... in LNCS Online.
Detailed information on LNCS... www.springer.com/lncs
Proposals for publication... LNCS Editorial, Tiergarten...
E-mail: lncs@springer.com

ISSN 0302-9743

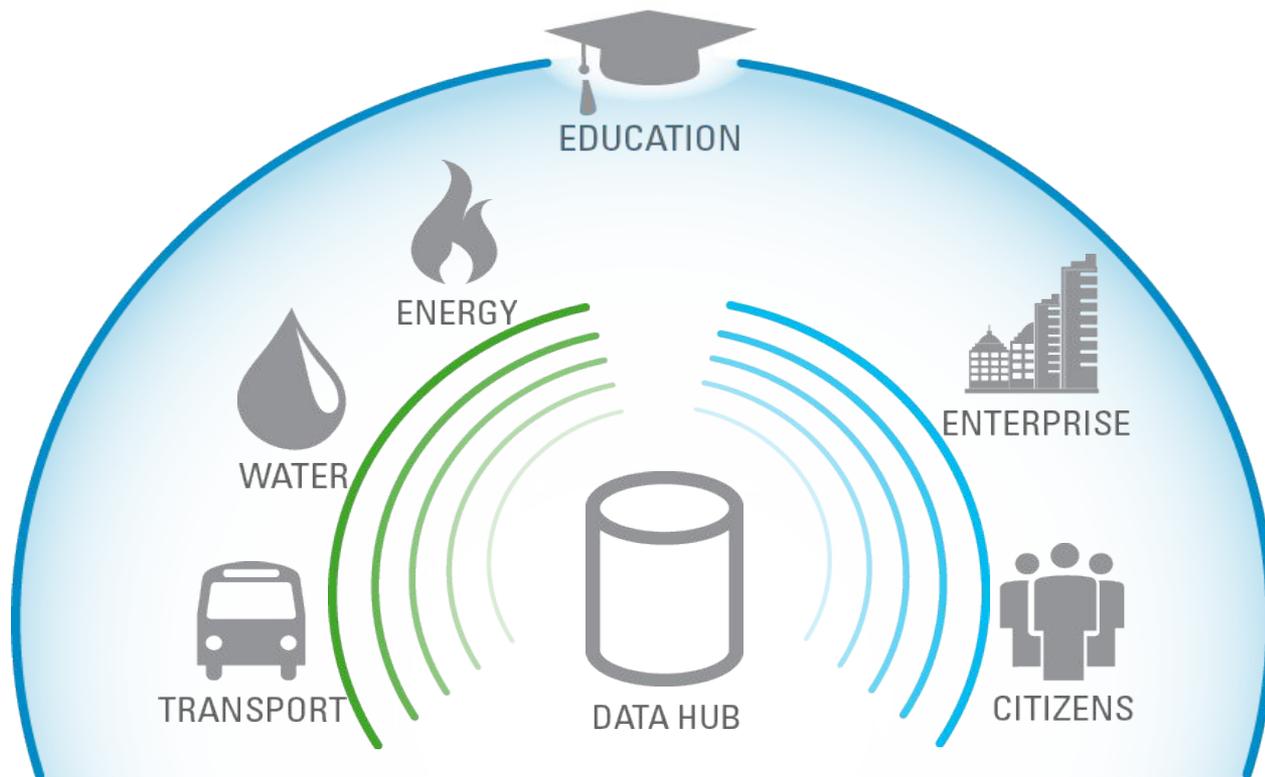
ISBN 978-3-319-11111-1



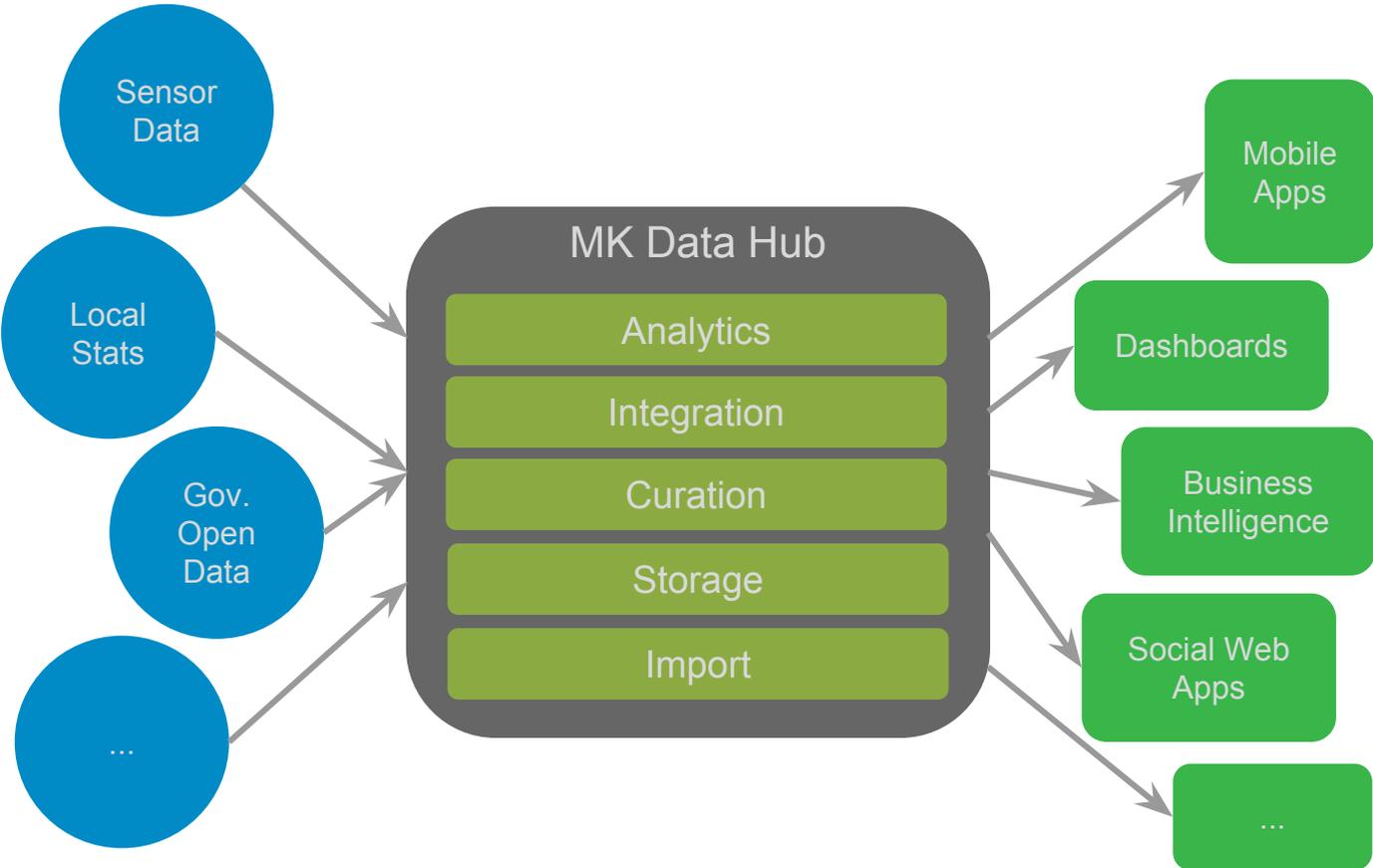
MK:Smart



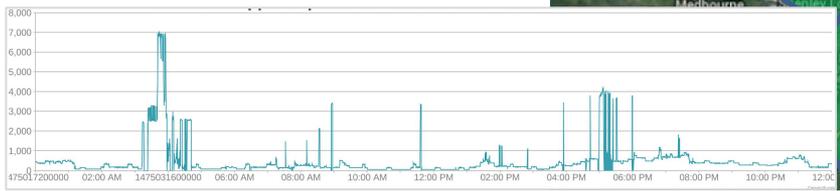
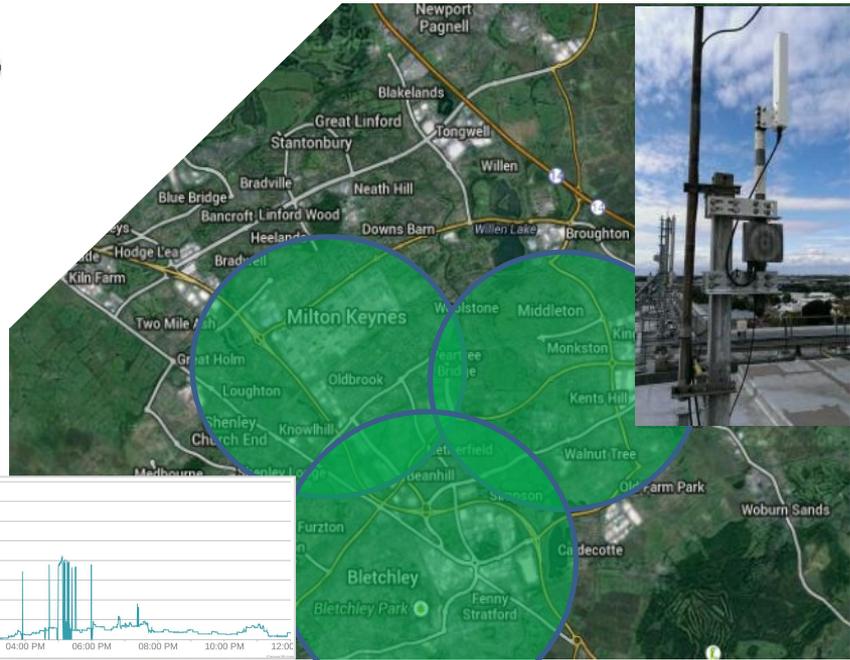
Les données au centre de la ville intelligente



MK Data Hub



Integre les données temps réel



Au travers d'API intelligentes

Entity lookup

With the **entity-centric API**, the MK Data Hub enables users and applications (e.g., mobile apps, etc.). Through this page, you can preview city data for your applications.

To obtain data about an entity, select its type from the drop-down box, enter the entity name, and click the **Lookup** button.

ward

Entity ID

https://data.beta.mksmart.org/entity/ward/walton_park

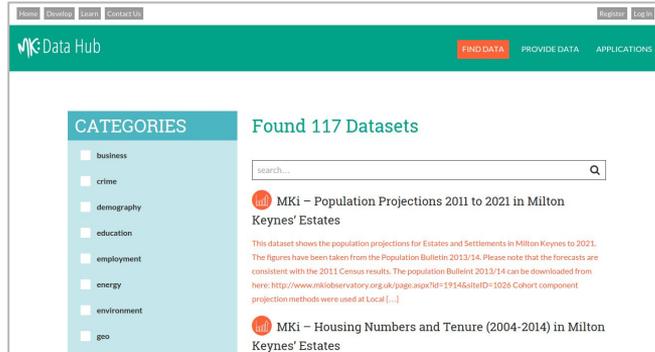
Data

```
{
  "global:religion": [
    {
      "global:sikh": [
        "76"
      ],
      "global:buddhist": [
        "91"
      ],
      "global:jewish": [
        "24"
      ],
      "global:no_religion": [
        "4902"
      ],
      "global:has_religion": [
        "8697"
      ],
      "global:religion_not_stated": [
        "960"
      ],
      "global:all_persons": [
```

Une API centrée entités, qui intègre les données de plusieurs sources autour d'un objet/d'une entité, sur le principe du Linked Data.

Compile les résultats “à la demande” (i.e. au moment de la requête) selon un schema virtuel (une ontologie) et évolutif.

Et d'un catalogue de données sémantique



Home | Overview | Learn | Contact Us | Register | Login

Data Hub FIND DATA PROVIDE DATA APPLICATIONS

CATEGORIES

- business
- crime
- demography
- education
- employment
- energy
- environment
- geo

Found 117 Datasets

search...

MKI - Population Projections 2011 to 2021 in Milton Keynes' Estates

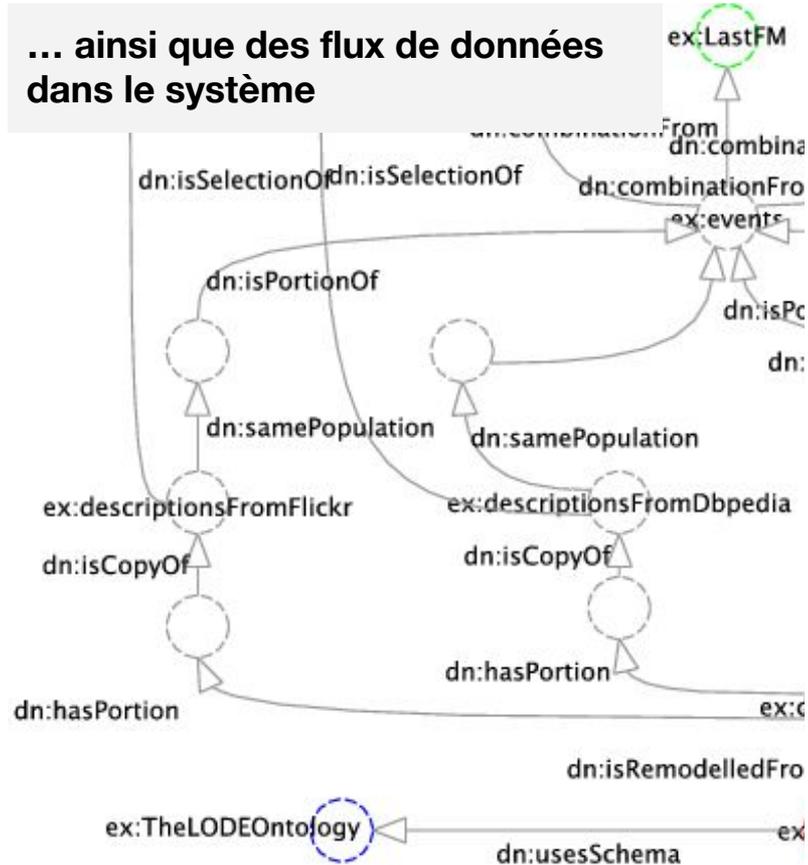
This dataset shows the population projections for Estates and Settlements in Milton Keynes to 2021. The figures have been taken from the Population Bulletin 2013/14. Please note that the forecasts are consistent with the 2011 Census results. The population Bulletin 2013/14 can be downloaded from here: <http://www.mkobservatory.org.uk/page.aspx?di=19146&siteID=1026> Cohort component projection methods were used at Local [...]

MKI - Housing Numbers and Tenure (2004-2014) in Milton Keynes' Estates

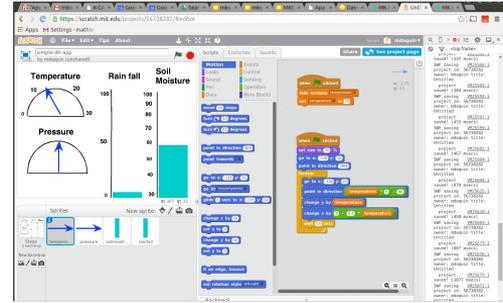
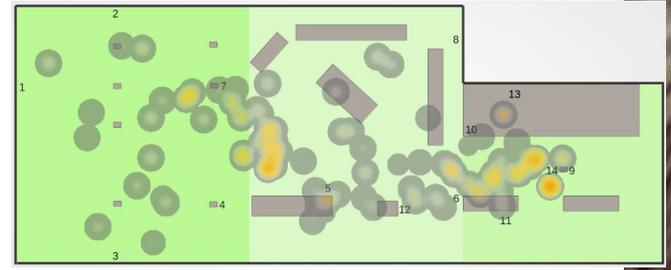
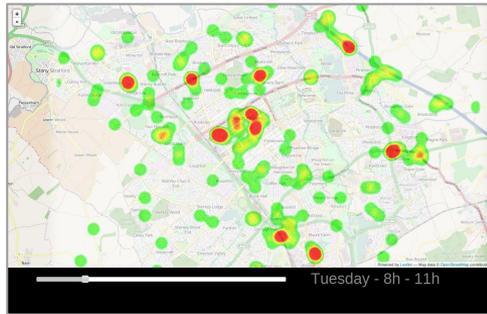
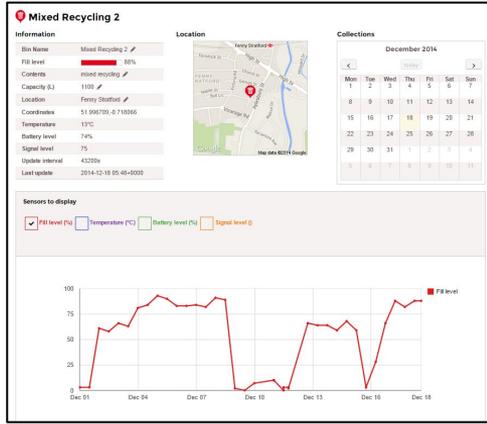
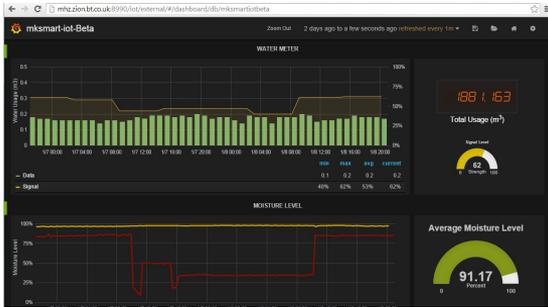
Représentation explicite et raisonnement sur les licence de données...

```
lic:cc-by-nc4.0 a odrl:Policy ;
rdfs:label "CC-BY-NC" ;
odrl:permission [
  odrl:action cc:Distribution,
  ldr:extraction , ldr:reutilization ,
  cc:DerivativeWorks , cc:Reproduction ;
  odrl:duty [ odrl:action
  cc:Attribution , cc:Notice ] ] ;
odrl:prohibition [ odrl:action
cc:CommercialUse ]
```

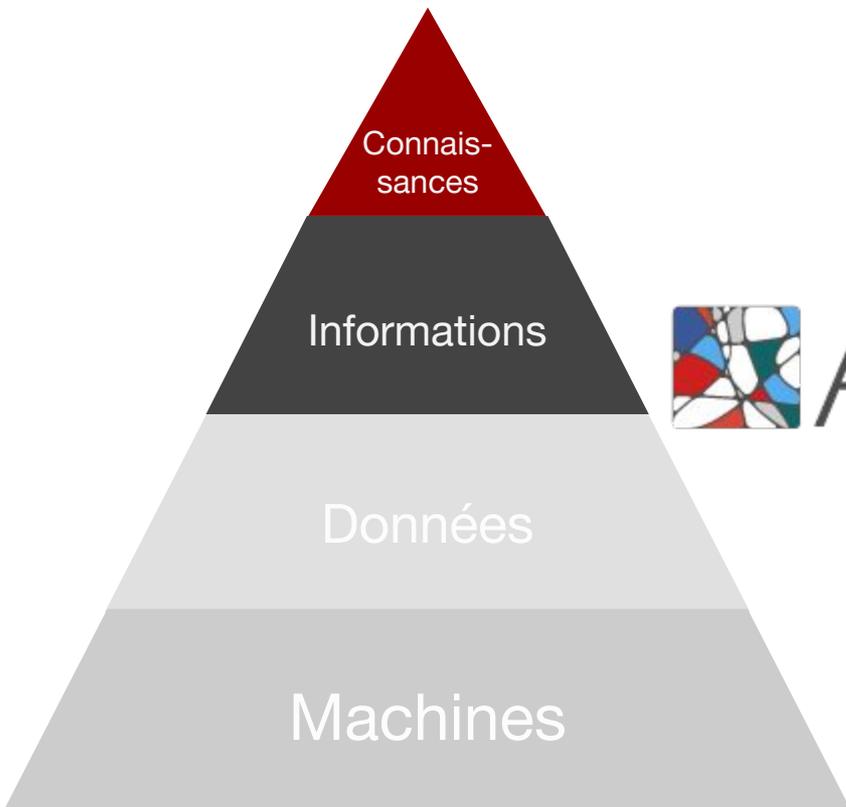
... ainsi que des flux de données dans le système



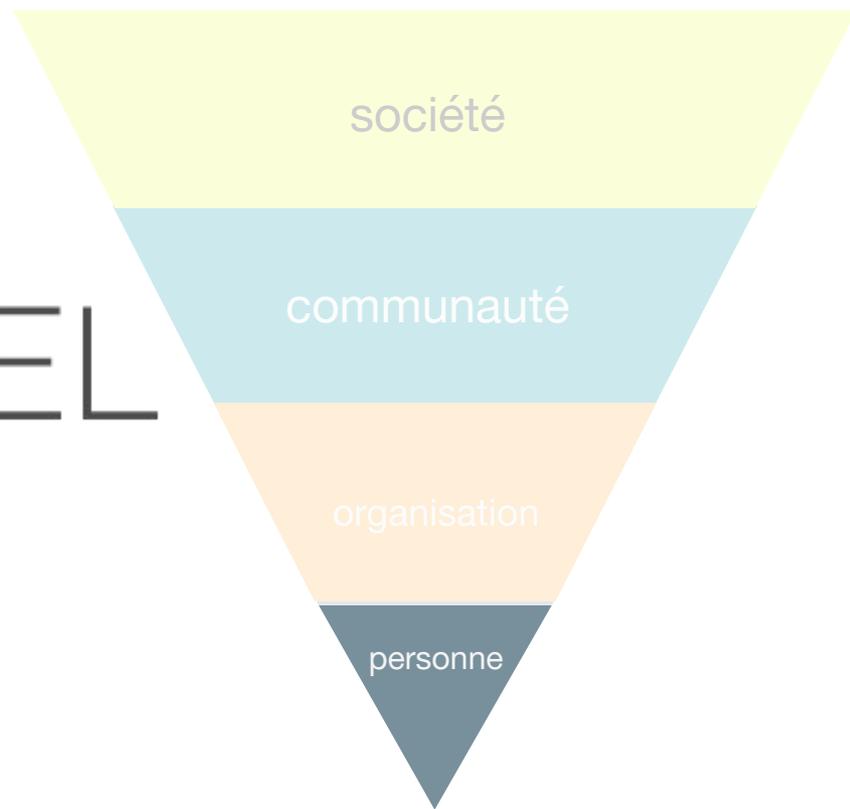
Applications



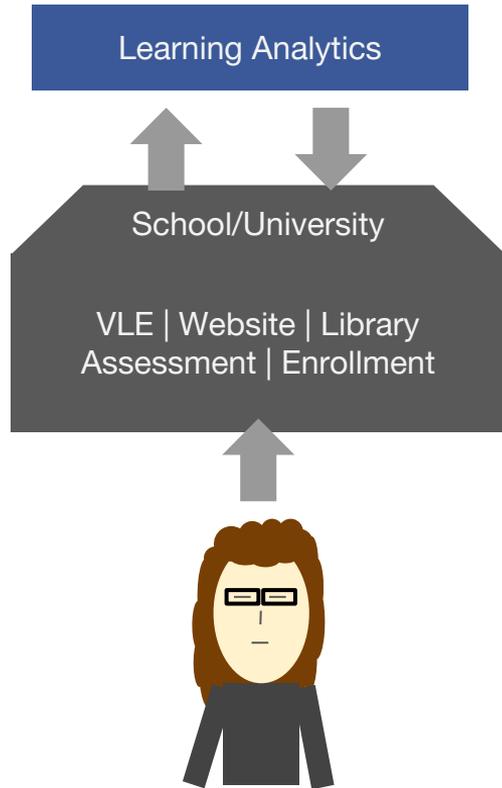
AFEL



AFEL



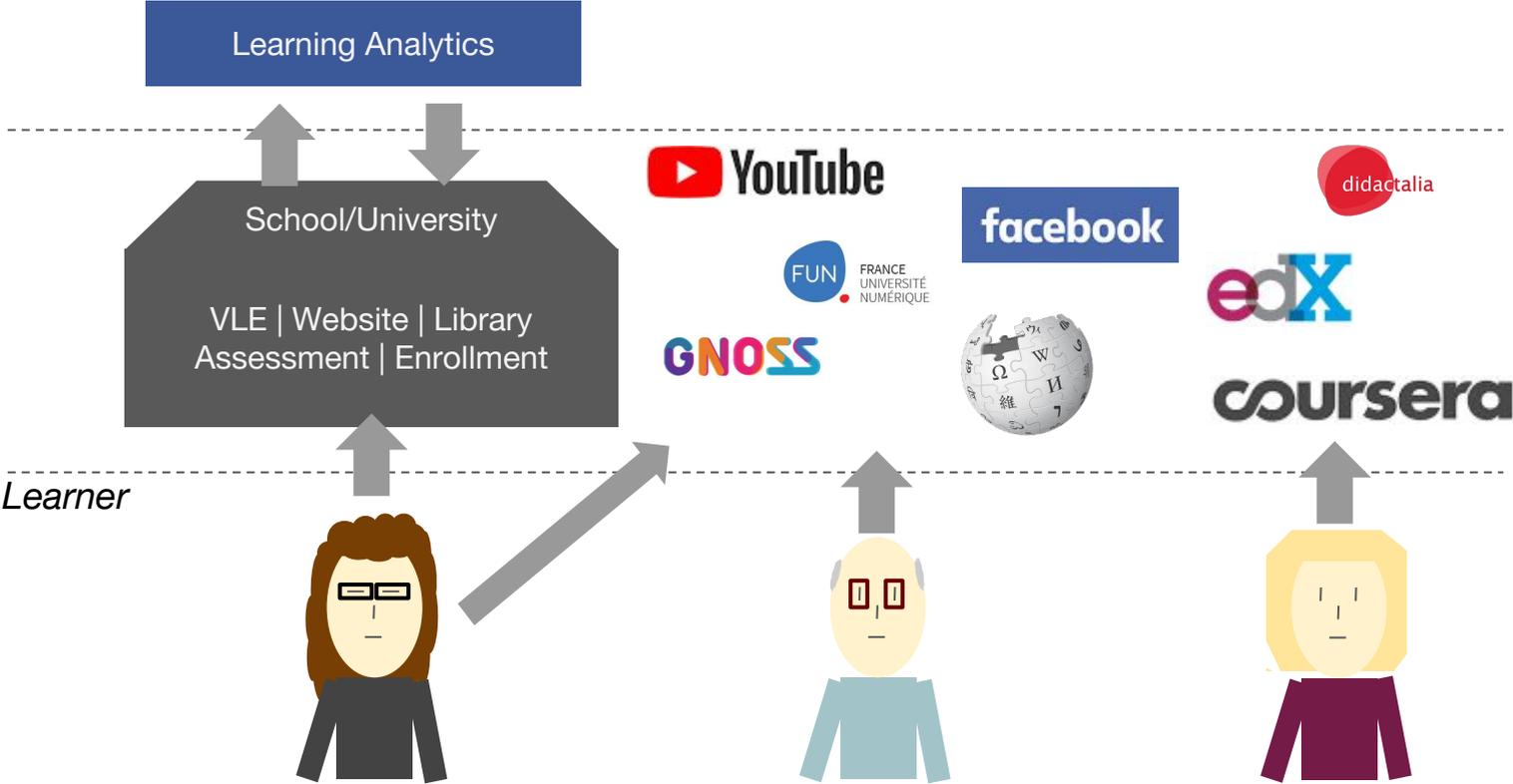
Le learning analytics (L'analyse de données d'apprentissage)



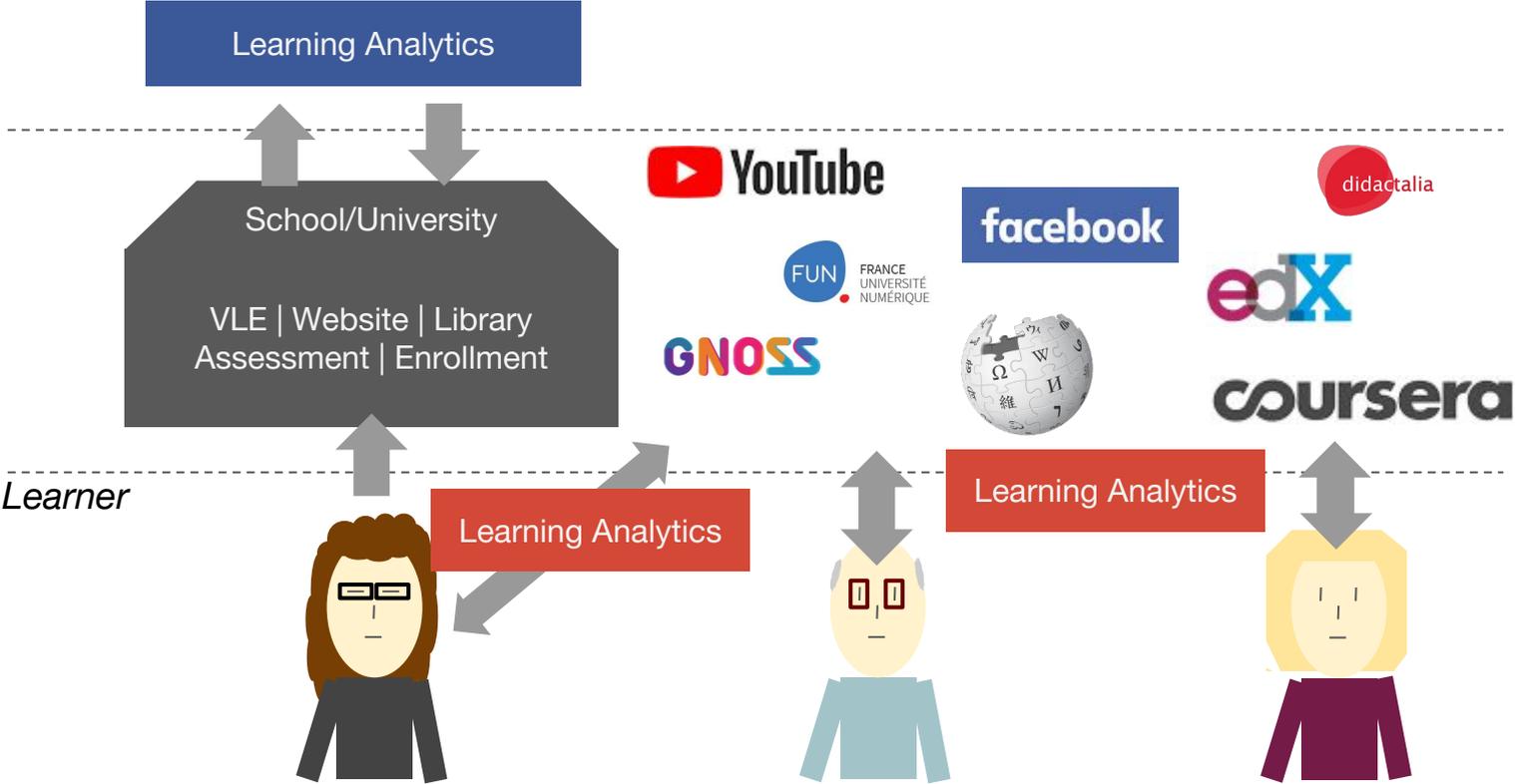
Typiquement, une université utilise les données sur les étudiants et sur leurs activités au travers de leur système informatique dans le but de prédire la réussite des étudiants afin de pouvoir l'améliorer...

Et si on a de la chance, aussi d'améliorer les enseignements...

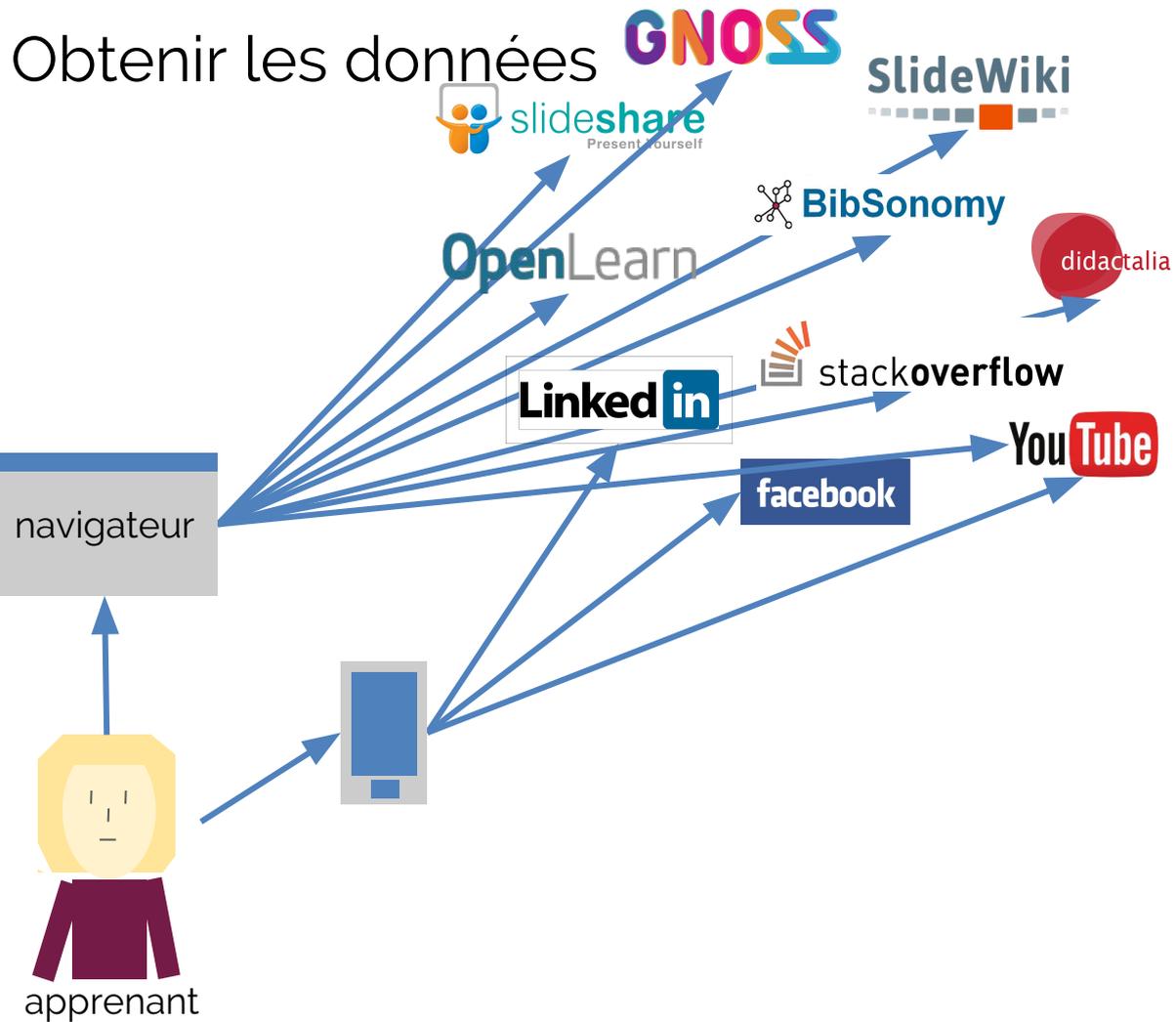
Mais : Everyday learning



Mais : Everyday learning

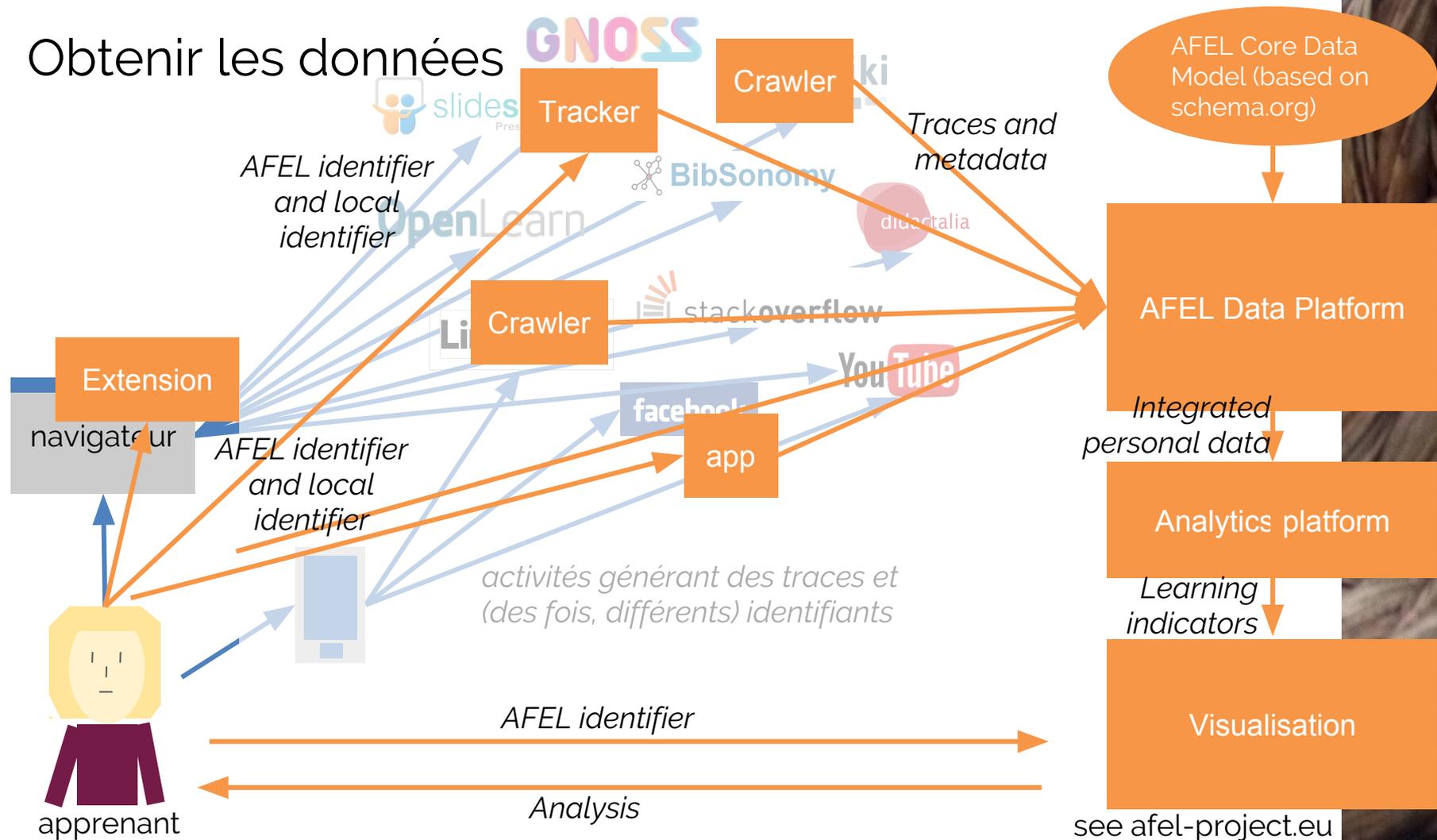


Learner

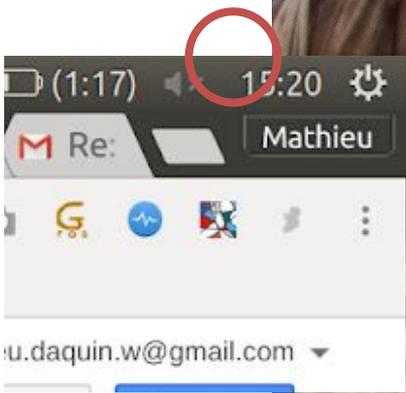
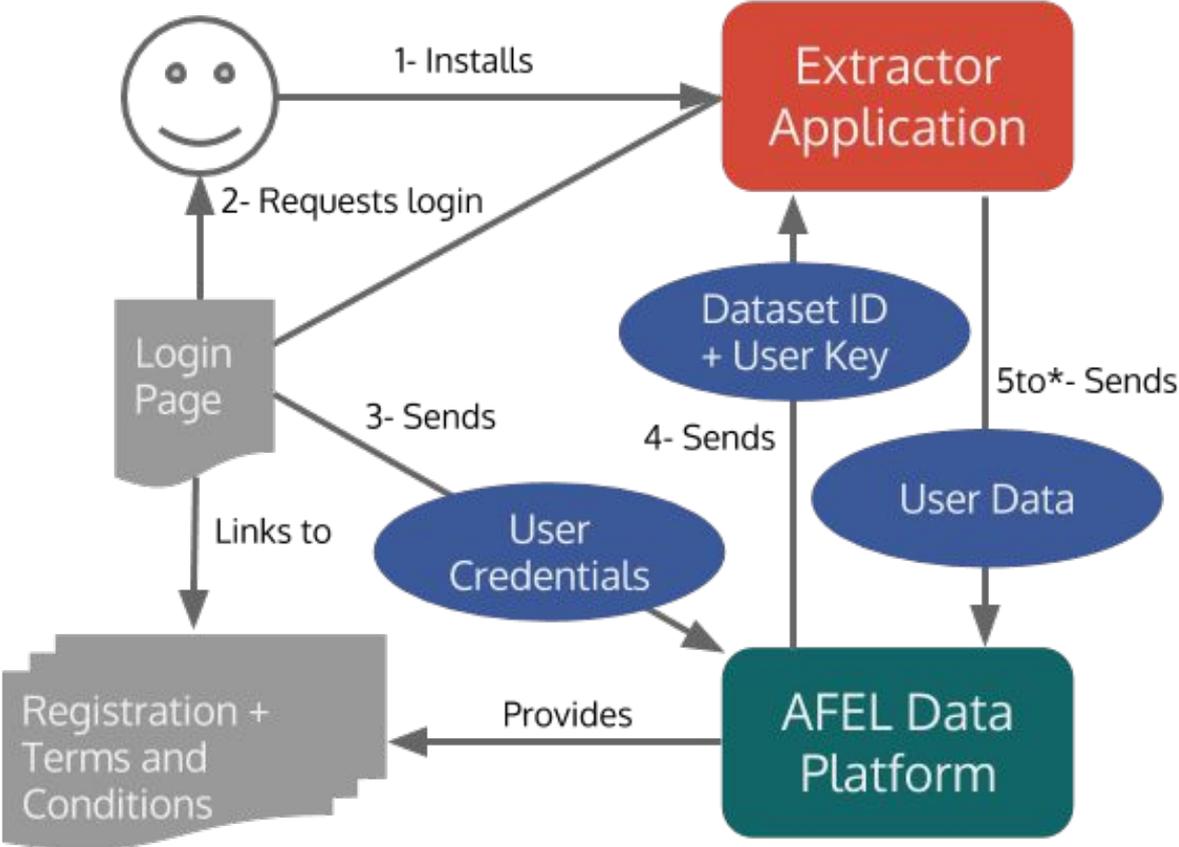


Obtenir les données

GNOSS

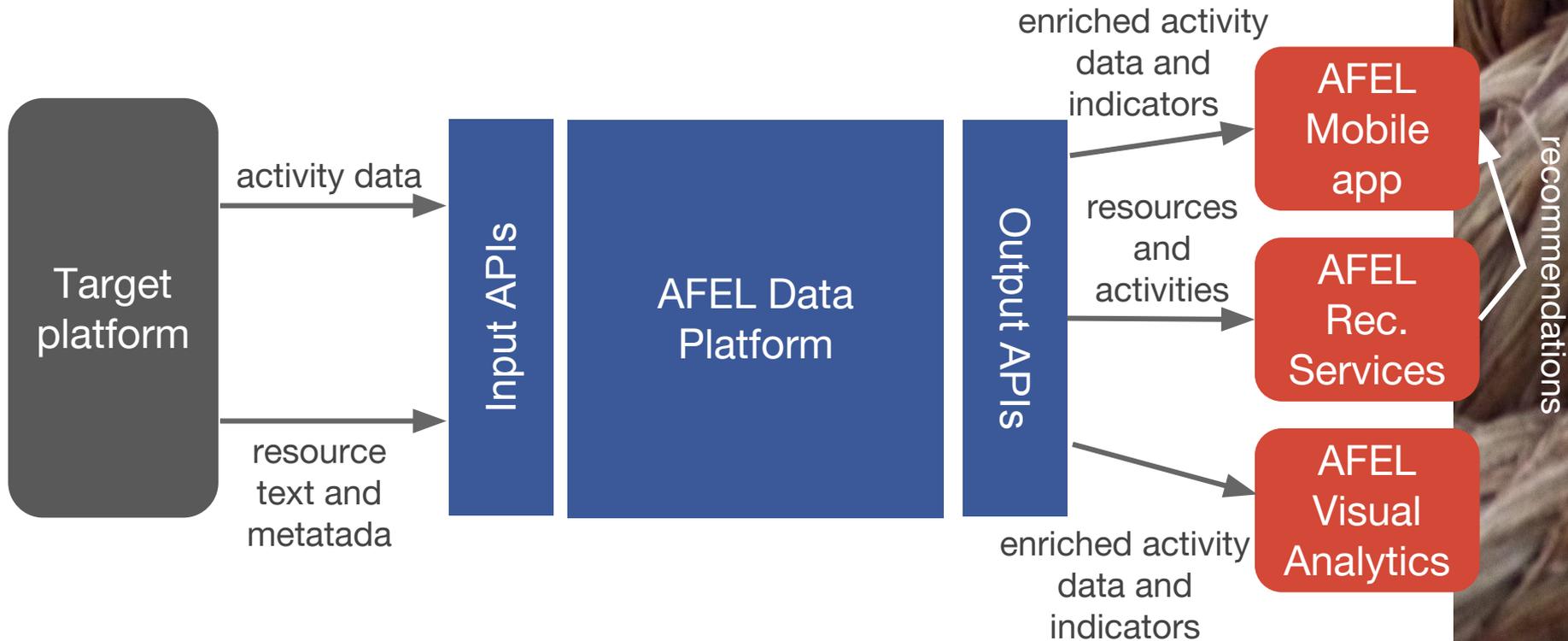


Exemple: Plugin pour navigateur



see afel-project.eu

Architecture pour l'analyse d'activités personnelles



Application

/D/

mixed intensity coverage
diversity complexity

Various

probability statistics

Topics of learning resources by mix of indicators

Topics of learning resources by mix of indicators

time spent number of tries
score

European Union's capital cities

European union countries

Games played by number of times playing

/D/

Scope: probability

diversity intensity
complexity

Top 10 activities...

- Calculating Probabilities of Combined Event? At grade
- Combination? Basic
- Recognizing Permutation? At grade
- Probability and Permutation? Basic
- Probability and Combination? Basic

/D/

Goals for scope probability

close

Increase coverage monthly (remove)

Increase diversity weekly (remove)

Increase work daily (remove)

Increase

/D/

New Goal in learning scope probability

work more increase coverage
increase complexity
increase diversity

The Didactalia AFEL app monitors your activity on Didactalia. It can remind you to focus on the topic you have chosen and to simply read more, watch more and do more on that topic.

Timing:

monthly weekly daily

go

Recommendations

- Descartes: [Philosophica: Enciclopedia filosófica on line]
- Paises de Europa@es||Countries of Europe@en||Europako herrialdeak@eu||Länder in Europa@de||Paesi dell' Europa@it||Pays d' Europe@fr||Paises de Europa@gl||Paisos d'Europa@ca||Paises da Europa@pt||
- Velázquez, La Venus del Espejo (Artehistoria)
- Experimento: La jaula electromagnética de Faraday (madrimsd.org)
- Ríos de Europa@es||Os rios da Europa@pt||Els rius d'Europa@ca||Europako ibaia@eu||Os rios de Europa@gl||Fleuves d' Europe@fr||Die europäische Flüsse@de||Fiumi d'Europa@it||Rivers of Europe@en||

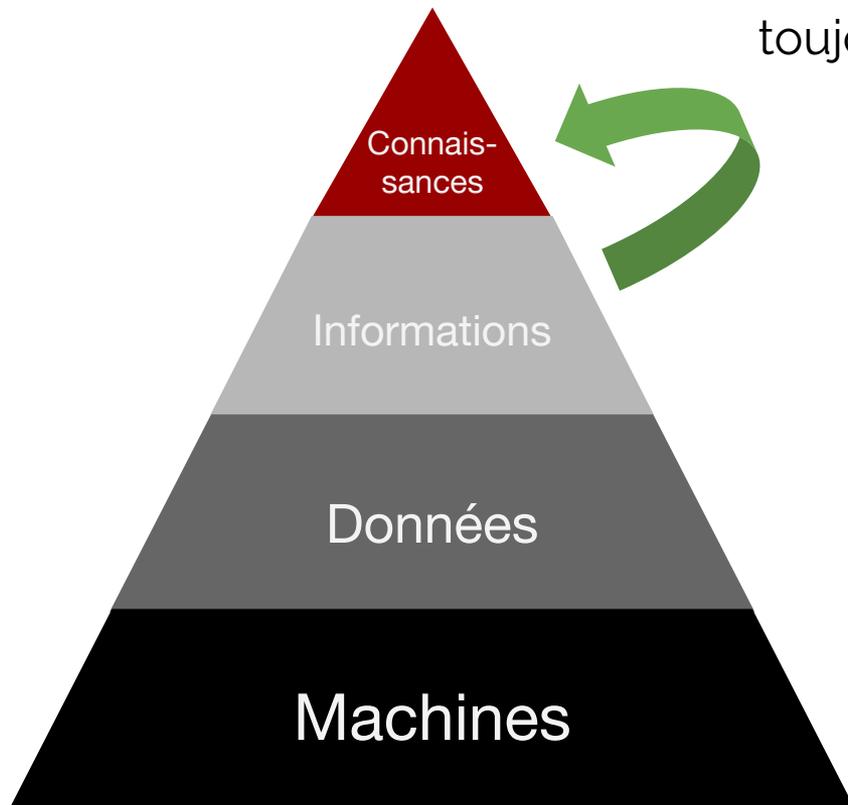
Qu'a t'on appris ?

La taille des données et l'hétérogénéité sont importantes, mais ça on le savait déjà.

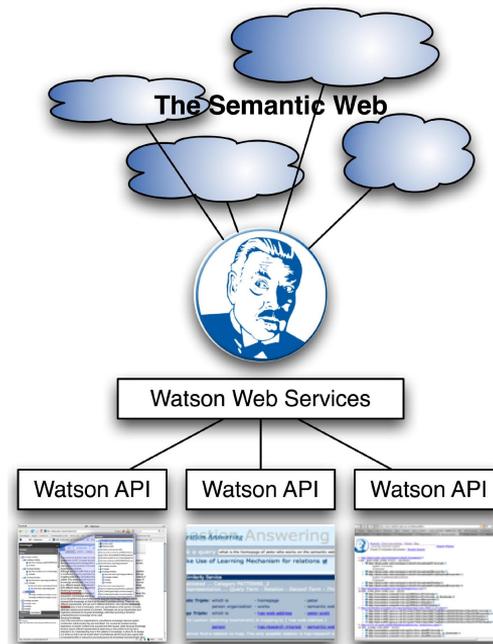
La diversité des données et de leurs usages est toujours un problème.



Qu'a t'on appris ?

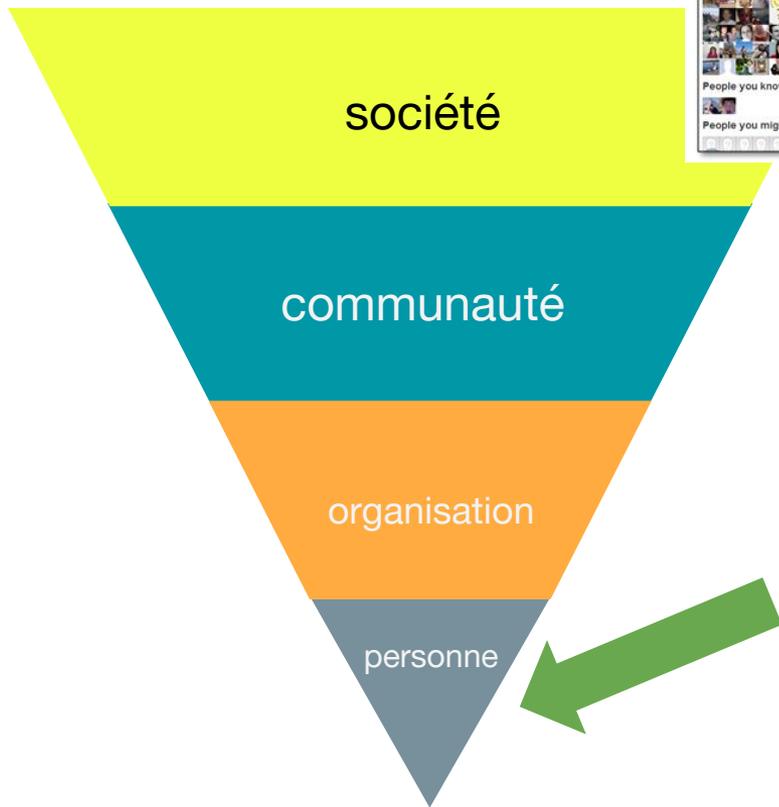


Passer de la à la est toujours difficile...



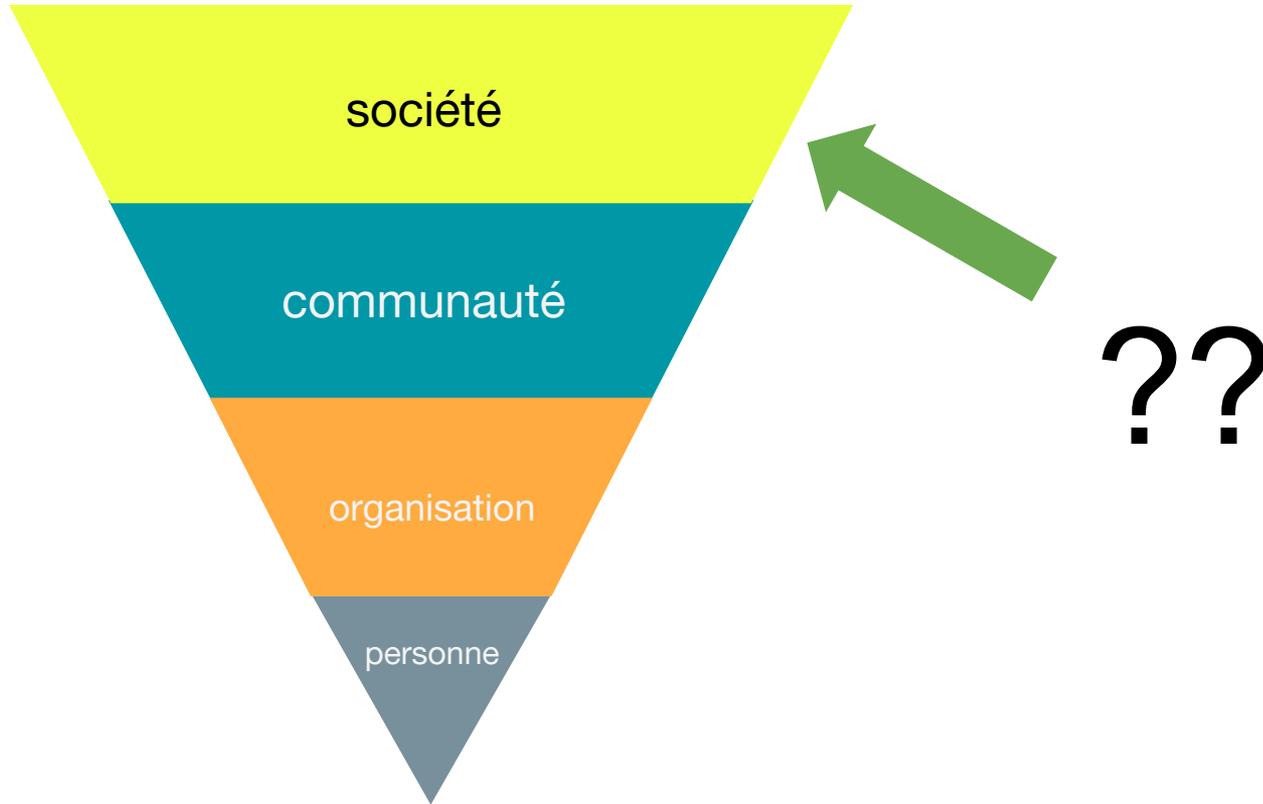
see papers from 2006 to 2009

Qu'a t'on appris ?

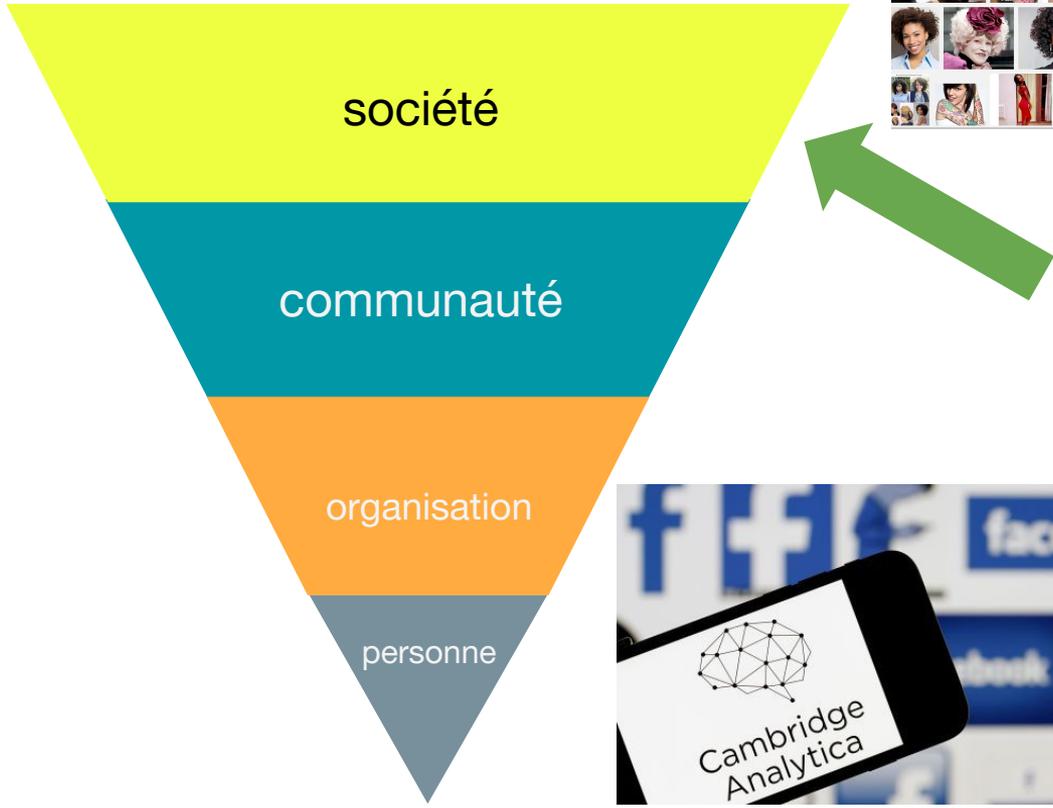


Ce niveau la est plus que jamais un problème !

Qu'a t'on appris ?



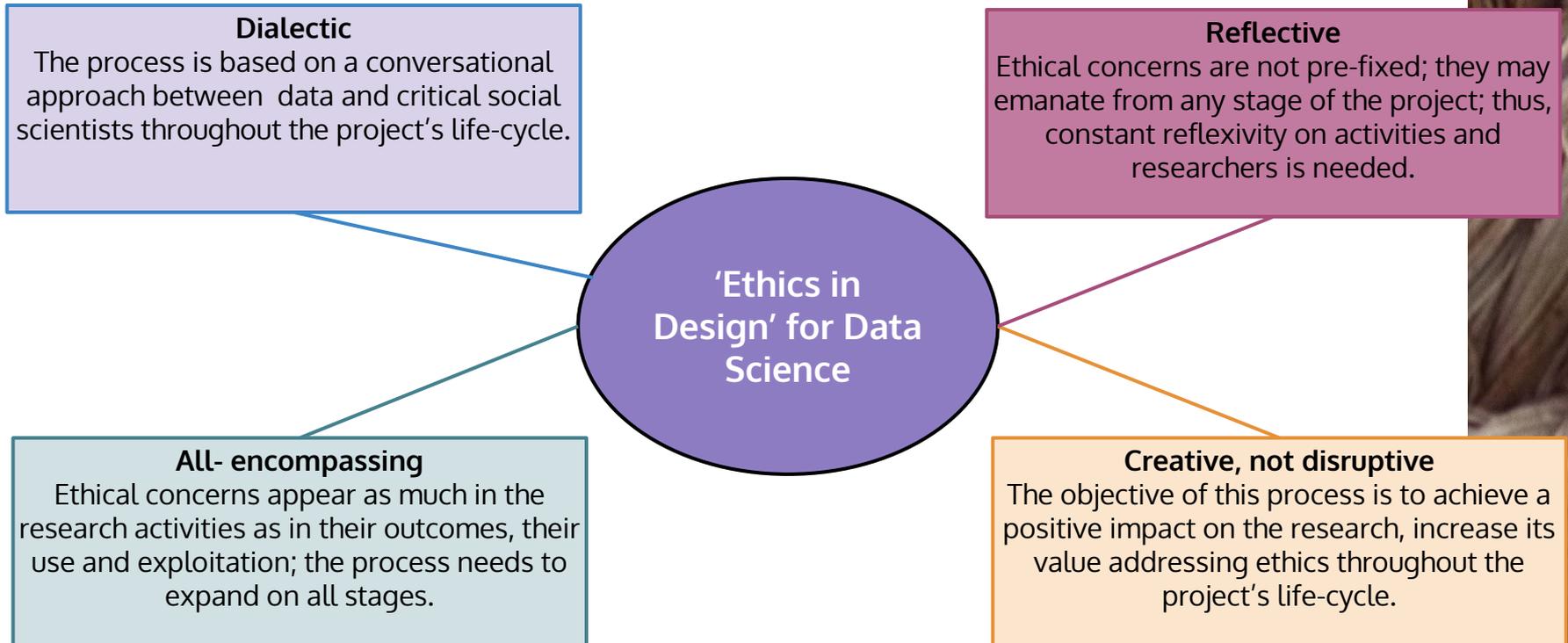
Qu'a t'on appris ?



??



Data Ethics - Comprendre les implications du traitement des données.



see paper at AIES 2018

Data Ethics - Comprendre les implications du traitement des données.

Design Fiction - utilisation de méthodes de narration pour extraire and comprendre les implications possibles des technologies, et pouvoir intégrer des méthodes préventives dans leur conception. Voir :



'Ethics in Design' for Data Science

Workshop "Re-coding Black Mirror" à The Web Conference (WWW 2018).

Reflective

Ethical concerns are not pre-fixed; they may emanate from any stage of the project; thus, constant reflexivity on activities and researchers is needed.

Creative, not disruptive

The objective of this process is to achieve a positive impact on the research, increase its value addressing ethics throughout the project's life-cycle.

see paper at AIES 2018

Conclusion

On est toujours de plus en plus à la recherche des éléments d'infrastructure qui permettent de distribuer et de rendre exploitables les données.

Dans les années 1900, Nicolas Tesla investissait dans la recherche dédiée à la transmission d'électricité sans fils. Une des raisons est qu'il ne semblait pas faisable de construire l'infrastructure globale nécessaire pour amener l'électricité dans les foyers par des fils.

Concernant les infrastructures de données, peut être somme nous, de façon similaire, à la fois trop et trop peu ambitieux...



Merci

mathieu.daquin@insight-centre.org

mdaquin.net

[@mdaquin](https://twitter.com/mdaquin)

