

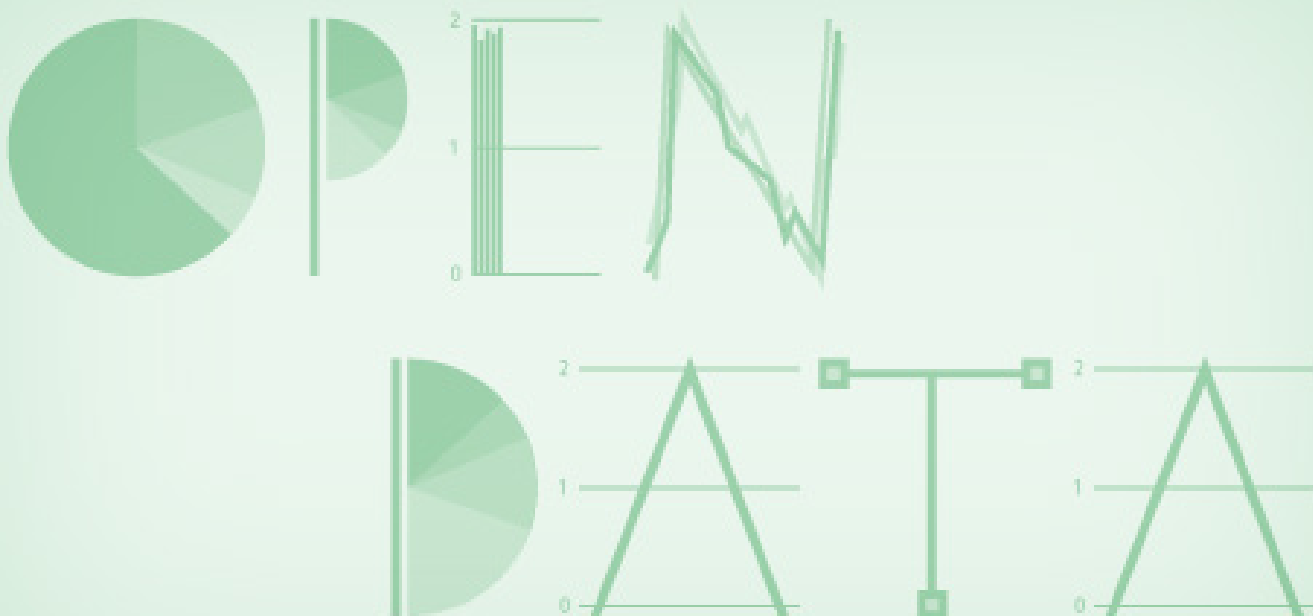
la collection numérique

de l'Agence de mutualisation des universités et établissements d'enseignement supérieur ou de recherche et de support à l'enseignement supérieur ou à la recherche



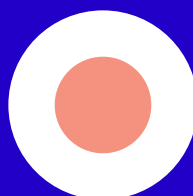
September 2019 (FR)
May 2020 (EN)

Open Data and french Higher Education : Opportunity to create new services



amue 

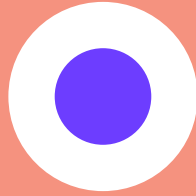
MUTUALISATION + SOLUTIONS



#05EN



Open Data and french Higher Education : Opportunity to create new services



Open Data and french Higher Education : Opportunity to create new services

Editorial

The Higher Education and Research community has a good grasp on the Open Science topic which is an important one for our research in terms of public policy. It may at times cover up the more global subject of Open Data which must be seen as a whole within all the contexts of our French higher education and research system.

Its openness and sharing nature results from an older and more crucial movement that echoes the doctrine of "open government" governance which embodies, among other things, the values of public life transparency, co-production of public services, transparency and promotion of public action.

Transformation vector

Open Data, like many other vectors of transformation, is based on a set of societal, legal, digital, cultural, organizational and citizen movements.

Led by the State, taken up at the level of ministries and many public operators, this movement would need to be amplified within our higher education institutions in order to respond to this public policy challenge.

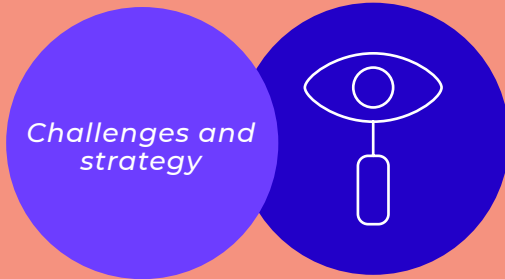
Implementation Strategy

This is not an additional burden but rather an opportunity to promote the action carried out by our institutions, to propose or enable new value-added services for public service users around our data. Let's ask ourselves about our information assets, let's enhance our action by opening up our data, documents and algorithms, let's imagine new services for our users, in a word let's build an Open Data strategy for each of our establishments. A strategy to be integrated into a broader strategy to enhance the value of the information held by our institutions.

Raise awareness

This digital collection aims to raise awareness about Open Data, help school governance to tackle this topic, discover new experiences and prepare oneself to contribute to this societal movement. A day organized by the « Amue » on November 14th will come to complete this approach that will enrich your reflection on Open Data, help you build your roadmap and present the work of the Amue in this specific field. I hope to see many of you there.

By Stéphane Athanase, CEO, Amue



1

By **Marin Dacos**, Open Science Advisor to the Director General for Research and Innovation, Ministry of Higher Education, Research and Innovation

The convergence of government open data and open science



The National plan for open science

The National plan for open science was made public by Frédérique Vidal Minister of Higher Education, Research and Innovation on 4th July 2018. It implements 9 measures to open scientific publications and research data. The first interest is a political one: science which is essential funded by the public sphere ought to be within the hands of the public, of everyone, and not confined to few campuses which have the means to afford very expensive subscriptions. Then, the interest is scientific. By overcoming the barriers, the universal access to publications improves the quality of science. Finally, it shows a society advantage: the more science is open, the more accessible it is to the society.

The National Plan has been initiating a convergence between the different opening movements which concern the public sector since the Budapest Declaration in 2002, which invented the concept of open access to scientific publications, and the Sebastopol meeting in 2007, which forges the principles of open public data. In 2016 and 2018, France took two initiatives that make these movements to converge.

Open as much as possible, close as much as needed.

The “Loi pour une République numérique”, i.e. the law on the new impacts of digital technology on society, grants a new right to researchers and teachers-researchers as authors. When the State funds at least 50% of the research, the author has the possibility to download the accepted author manuscript into an open archive for 6 (STM - science, technology, medicine) or 12 months (SHS, Human and Social Sciences) after publication. If the contract says otherwise, it is deemed unwritten. In the data field, the law introduces a new obligation, which has not been identified enough by the Higher Education institutions (higher education, research and innovation). The article 6 of the law sets the principle of opening by default all the administrative data, including research data. The public officials work is owned by default by the funder, and thus by the employer. However, the researchers and the teachers-researchers get exceptions regarding their published articles and books, which don't belong to the funder, but to the author. This exception is only valid for the publications and not for the genuine research material which constitutes the data. The latter belong to the funder, which provides institutions with a serious leverage to set up a real data policy. It should be noted that the article 6 was not essentially conceived for data research, but for public administrations in general. Nonetheless, it applies to the Higher Education institutions, both for research data and data available to institutions to ensure their



Of course, half a dozen exceptions are provided by the law, so as not to open in an abusive manner.

The National Plan doctrine for open science as well as the open science policy of the European Union goes as far as saying: “open as much as possible, close as much as needed”.

Convergence of movements

Besides, the Open Government Partnership (OGP) which is an international organization regrouping more than 70 countries, including France, publishes every other year national commitments as part of the transparency of public action. The OGP was born in 2011 from the open data movement. In its 2018 commitments, France decided to integrate to its “traditional” commitments of data openness, a chapter dedicated to open science. (“Commitment #18: Building an ecosystem of open science”).

This convergence of movements to open science and public data is a key to their success. Indeed, we must integrate this paradigm at the heart of our technical systems in order to make that opening a logical and integrated step in the data work processing, and not an additional “out-of-circuit” burden. To move in this direction, it is necessary to mobilize all the Higher Education players. The movement concerns the overall scientific process: the researchers, first, but also all the accompanying research occupations, at all the organizational levels of the research community. It seems that the collective determination exists, but the technical, legal and disciplinary complexity of the subject represents a significant challenge. This is why a national coordination, through the Committee for Open Science (Coso), has been set up. At its basis, more than 250 experts who have expressed their interest, and discuss digitally within the framework of the Open Science Forum. At the top, there are the presidents of the institutions who represent the entire higher education, research and innovation sectors. With this combination of field expertise, which we hope is representative of the diversity and complexity of situations and decision-making capacities at the highest level, we have a strong instrument of national coordination and international contribution.

Committee for open science

Should you wish to participate in the works of the Committee for open science, especially through its forum, please note that the call for participation is still open. We need every profession. Send your CV and a letter of motivation at coso@recherche.gouv.fr

You will find a library of indispensable resources and of news related to open science on <https://www.ouvrirlascience.fr>, including a legal guide of the research data openness.

CONTACTS

Mail:
marin.dacos@recherche.gouv.fr

On Twitter :
[@marindacos](https://twitter.com/marindacos)

EXCEPTIONS TO OPENING OBLIGATION

Data personal data without consent;
Data which shows risks for public security or the institution security;
National Defense Secret;
Professional secret;
Business secret;
Industrial and commercial secrets;
Data that may represent some risks for the protection of the nation scientific and technical potential





By
Romain Tales,
DINSIC

Etalab: A mission in the service of open public data

The Etalab mission, attached to the inter-Ministerial Department of Digital, State Information and Communication Systems (DINSIC and now DINUM), was created in February 2011 in order to promote France's open data policy, i.e., open access to public data. Less than a year later, on 5 December 2011, the national open data platform was launched: data.gouv.fr. This pioneering approach to open access was finally enacted into law in October 2016 for a Digital Republic. Henceforth, public data must be disseminated by administrations "as a matter of principle", and no longer only on request.

Open Data and french Higher Education : Opportunity to create new services



6

Principle of open data: a waste of time?

Administrations in the broadest sense of the term (ministries, local authorities, public institutions under supervision or public companies) often oppose the obligation to disseminate public data by arguing that they have insufficient time. Making data accessible would cause them to waste time better spent on their core services. On the contrary, making databases available as a matter of principle will save them valuable time: by proactively disseminating data on the national platform, the administration no longer needs to respond to individual requests, nor perform specific searches that would divert it from its primary activities.

A Case for Openness: Benefits for Administrations and Data Reusers

Open data first of all publicizes the administration's work and enhances its value. In addition to facilitating the daily life of administrations by saving them time in implementing their public service mission, it is also a means of increasing the effectiveness of public policy in general: the circulation of data creates synergies between offices, organizations, and other entities and improves the implementation of public policies through the harnessing of data. This dissemination is an open channel to re-users of public data, perhaps to supplement them, to detect anomalies, or even to correct them. In this way data, such as the database on traffic accidents resulting in injuries, have been improved in terms of quality, notably through feedback from re-users.

For these re-users, who can be both government and private, the analysis of data disseminated in open access creates innovations through cross-referencing, which can allow administrations to meet needs that they did not necessarily foresee when the data were initially produced.

Some examples of the benefits of open data among others...

The creation of a Public Data Service

Created by the law for a Digital Republic, the Public Data Service aims to make reference data available, with the intention of facilitating their re-use, i.e. those data with the greatest economic and social impact.

Reference data meet the three conditions defined by the law: they constitute a common database for naming or identifying products, services, territories or persons; they are re-used frequently by public or private persons other than the administration that holds them; and their re-use requires that they be made available at a high level of quality.

The list of reference data, established by decree of the Council of State, currently includes nine sets of data from the following domains:

- economic and associative: the Sirene database for the identification of establishments, national directory of associations, directory of jobs and professions,
 - geographical: computerized land registry map, national address database, official geographical code, ...
- administrative: reference system for the administrative organisation of the State.

The main objective of the Public Data Service is to facilitate the reuse of data with the greatest impact for the economy and society.

The dynamism of the Public Data Service is now reflected in the strong support of its users. The increase in the number of downloads illustrates this. According to Institut national de la statistique et des études économiques (INSEE), the french Institute responsible for the production, analysis and publication of official statistics in France, the number of users of the daily update of the Sirene database has multiplied by 20 since its integration into the Public Data Service.



The land register was previously sold (there was only one buyer in 2016) and distributed free of charge on request to local authorities and State offices. Its online publication, supported and equipped by Etalab, has not only enabled better distribution of data (7 million downloads of all types since the beginning of the year) but has also relieved the burden on the public financial services, which now send requests directly to cadastre.data.gouv.fr.

The reference platform: data.gouv.fr Source: data.gouv.fr

Professional data users have taken hold of public data services (TomTom or EDF, for example). Today, startups emerge and offer services that allow us to view and use reference data.

Local government collectives also benefit directly from using reference data (for town planning, housing, community life). In addition, most of them also disseminate certain data or information through their own portals, located within the municipality.

patible nationwide.

Restoring the information balance within the economy...

With the issuance of data on real estate transactions (Requests for Property Values) by the Direction générale des Finances publiques (DGFIP), Etalab has developed an application, app.dvf.etalab.gouv.fr, that allows data to be visualized in map form, to demonstrate the importance of exploiting open data. "The idea was to provide the general public, for whom manipulating gigabytes of data is not an option, with a simple, easy-to-use visual interface," says Marion Paclot, data scientist on the Eta-



2

Photo by
Jonathan Gray

Open Data and french Higher Education : Opportunity to create new services

Improving forecasting to prevent accidents at sea...

The Regional Operational Surveillance and Rescue Centers (CROSS) have published data on their offshore operations on data.gouv.fr. As part of Etalab's "Common Good Entrepreneurs" challenges, and in partnership with public administrations, data scientists have used those data in Predisauvetage. This interactive map of assistance and sea-rescue operations allows you to visualize the activities by applying several filters, including the date, the nature of the operation, the location, the severity, etc.

CROSS and many sea-rescue stakeholders generated data on their operations for external reporting but did not perform any internal analyses of that information to assess their own actions. Analyzing these data broadens our knowledge about accidents at sea and enhances our ability to prevent them. Rescuers also have greater knowledge of the risks in certain areas. The project has had an effect on unexpected sectors: for example, insurance companies offering personalized contracts for nautical activities. The Departmental Directorate of Maritime Territories has also improved navigational aids in some places, and exchanges information with university hospitals. The "Common Good Entrepreneurs", or EIG, are working to standardize data in order to make the Predisauvetage system com-

lab mission. The application, published in its initial version shortly after DVF data was released, has since undergone several upgrades, particularly in relation to user feedback. The site receives about 23,000 visits per day, and has been consulted 2.5 million times since its first release.

FOR MORE INFORMATION

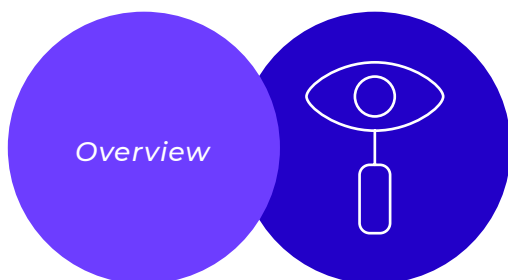
Etalab's website: <https://www.etalab.gouv.fr/>

The national platform for Open Data : <https://www.data.gouv.fr>

Data operations coordinated by the CROSS Secmar: <https://www.data.gouv.fr/fr/datasets/operations-coordonnees-par-les-cross/>

Predisauvetage sea-rescue site / SecMar map <https://carte.snosan.fr/>





By **David Rongeat**,
Digital
manager at the
Information
System
Strategy and
Programming
Department,
AMUE

Open Data: Overview



Let's start with a little bit of international history

Open Government: Open Data is a subject linked to "Open Government" which promotes the values of transparency in public life, citizen consultation, co-production of public service to improve the governance of States, bring out counter-powers and fight against corruption.

The main principles of open data were laid down / established in 2007 during the Sebastopol meeting in California. In 2013, the G8's open data charter creates the suitable conditions for the generalization of open data by default. Among the pioneers, the city of Chicago, which in 2005 disclosed information's about crime with the aim of objectifying it and allowing citizens to monitor its evolution.



Open Data: revolutionary?

Historically, the first signs of Open Data are on the French side with the 15th article of the Universal Declaration of Humans and Citizen's Rights (1789): "Civil society has the right to hold any public official accountable for his or her administration". / "Society has the right to call for an account of the administration of every public agent". Much later, in 1978 (via law 78-753 of the 17th of July 1978), the CADA (Commission of Access to Administrative Documents) was created. It is an independent administrative authority whose goal is to facilitate and control access by private individuals to administrative documents. For example, a citizen whose requested documents were not forwarded by an administration can approach the CADA.

During the Hourtin speech on the 25th of August 1997, Prime Minister Lionel Jospin declared: "Essential public data must henceforth be accessible to all free of charge on the Internet" and "the international dissemination of our public documents must be encouraged to this end". In 2011, the State created Etalab and opened the data.gouv.fr platform. Open data began its spread.

Legal framework

The 7th of October digital law drives / encourages the circulation of data and knowledge. Data of all public operators of more than 50 agents or municipalities of more than 3500 inhabitants (and 50 agents) must, since the application of this law, be open by default. Obviously, regulations regarding freedom and information technology, or the GDPR (General Data Protection Regulations) are in application in this context of transparency and participate in the protection of citizens. Reading this ticket allows the understanding of GDPR and Open Data's interactions.

Eventually, to conclude on this legal framework, the decree of the 12th of December clarifies which data and important documents containing personal data can be made public without anonymity. In our context of higher education and research, we can find for instance the results of the exams, or on the organization chart of public institutions.



FOR MORE INFORMATION

- *Loi pour une république numérique, Digital law of the 7th of October 2016 here*
- *Decroix, J. (2017). Private life and open data: Harvard proposes some advices at the website cnil.com*
- *Decree of the 12th of December 2018 regarding categories of administrative documents that can be made public without anonymity.*

Some key elements

Beyond this historical movement and legal framework, let's see what Open Data with some key notions is. The aim is to make data held by a public operator accessible to all (see article on data on page 8). It must be specified that behind this term "Data" we also find documents and software code, not only data from information systems.

Another important notion, it must be considered that a public institution has a double positioning with regard to Open Data: as a producer of open data but also as a consumer of data proposed by third parties.

Consumers of open data are citizens, NGO, companies that wish to create services around this data, the media and public operators. For Higher Education and Research, students and researchers are among the largest consumers of open data. By the way/Moreover, it is often described that Open Data has its roots in the academic world (see article on Open Science page 2).

However, it is also necessary to take into account some reticence related to Open Data. For example, access made possible to some of its data by a competitor; this applies less in the context of Higher Education and Research. Finally, implementing open data in an institution goes beyond the necessary respect/consideration with regulations. It consists in questioning its information assets, setting a strategy for establishing it and identifying the services that will can be created around this open data.

Strategy of implementation and benefits of Open Data

This implementation strategy can borrow into the many identified contributions of open data:

Enhance the work and production of the establishments and make it attractive,

- Serve general interests,
- improve the efficiency of some services,
- allow to question public policies,
- ease the mingling of data by relying on open frame of reference.
- For its Information System: know more about the data of our information system, improve data quality.

From the standpoint of the citizen and of research, Open Data gives transparency to public action and potentially user's implication in participative production's plans.

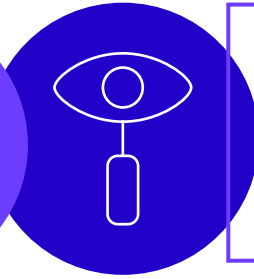
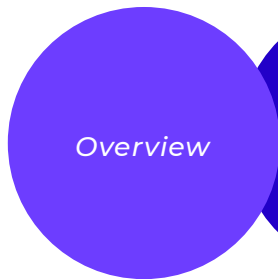
Opportunities to create services

Finally, and probably the main/fundamental point, opening data of one's institution makes it possible to create new services for users: some illustrations are available/ can be found in the article on uses page 11.

TO GO FURTHER

Open data Canevas is a wealth of open ressources
<https://opendatacanvas.org>





By

Emmanuel Weisenburger,
Department of decision-
making/support tools,
Department for Information
WSystems and Statistical
Studies (DGESIP), MESRI

Opening data to the Ministry of Higher Education, Research and Innovation

In September 2019, the Ministry of Higher Education, Research and Innovation (MESRI) proposed hundreds of open datasets of all types on its dedicated platform which is integrated into its institutional site and synchronized with the inter-ministerial / inter-departmental platform. Those datasets included student registrations, awarded diplomas, human resources, details of research infrastructures and major institutions or organizations, degree certifications, science award winners, results of calls for research projects, events, the agenda of the minister of Higher Education, Research and Innovation, ...

A specific state of the art interface makes it easy to explore all datasets (search engines, filters), to export or integrate directly and dynamically its data into third party systems (using APIs).



Innovative Initiative

This offer is the result of an innovative approach initiated in 2013 by our department. This approach aims to align the Ministry's information dissemination principles with the emerging legal framework, which has gradually introduced the principle of opening up public data by default. Beyond that, by sharing information widely, the Ministry intends to stimulate public discussions / to encourage public debate about French higher education, research, innovation and public policies that support it. At the same time, the Ministry perceives/sees the development of open data as one of the possibilities of modernizing its operations by promoting a better internal flow of information.

By implementing this approach, the Ministry adopted 3 structuring principles: giving priority to quality over quantity of data sets, encouraging the reuse of information it offers by setting up support system, developing services and tools for datavisualization on key datasets.

Datasets are historically recorded and enriched by identifiers which facilitate the interoperability of the relevant information (introduction identifiers SIREN, UAI, Grid, Wikidata, Idref, RNSR,...). The meta-data shall specify the conditions for producing this information. Each set is designed to guide easily and to give enough information for its interpretation and for its proper utilization ; each variable is thus duplicated : a first one includes a code, a second one a standardized wording/label and comprehensible/ understandable for the public. For the most complex datasets, a document describes the structure of the dataset variable by variable and specifies conditions for a stringent/strict use of this information.

Viewing Data

For a large part of its Data, The Ministry of Higher Education, Research and Innovation offers a set of services and data visualization tools (e.g. scanR, Search Engine for French research and Innovation)

referenced within dataESR, portal listing all the data resources it offers. These tools provide re-users with framing elements for self-checking by comparing their results with reference aggregates. Their main purpose is to make open information accessible to all, which, if it were only accessible via often complex and massive raw data files, would be reserved to the only data specialists. Data visualization tools concern for example, indicators of success in bachelor's degree, PACES and DUT at university, or the provision of APB data, and soon Parcoursup, by higher education training. This certainly explains the success of making MESRI data available. The visit of the open data platform continues to grow. Since its opening, more than 650,000 data files have been downloaded, including 175,000 in the first 8 months of 2019. But the opening of data is a continuous process and MESRI continues its efforts to augment its offer, to always improve the quality of the proposed data and to automate the updates. It is also following with interest the works carried out in institution to adapt to the legal framework of the "La loi pour une République numérique" Digital Republic Law and to develop local data offers. In this context, it is all the more crucial that ministry and institutions organize together the conditions for a sustainable production of open data and to develop the interoperability of exposed data.

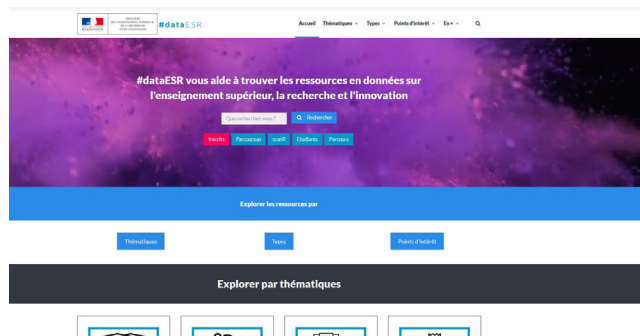
Open Data and french Higher Education : Opportunity to create new services



BUSEFUL LINKS

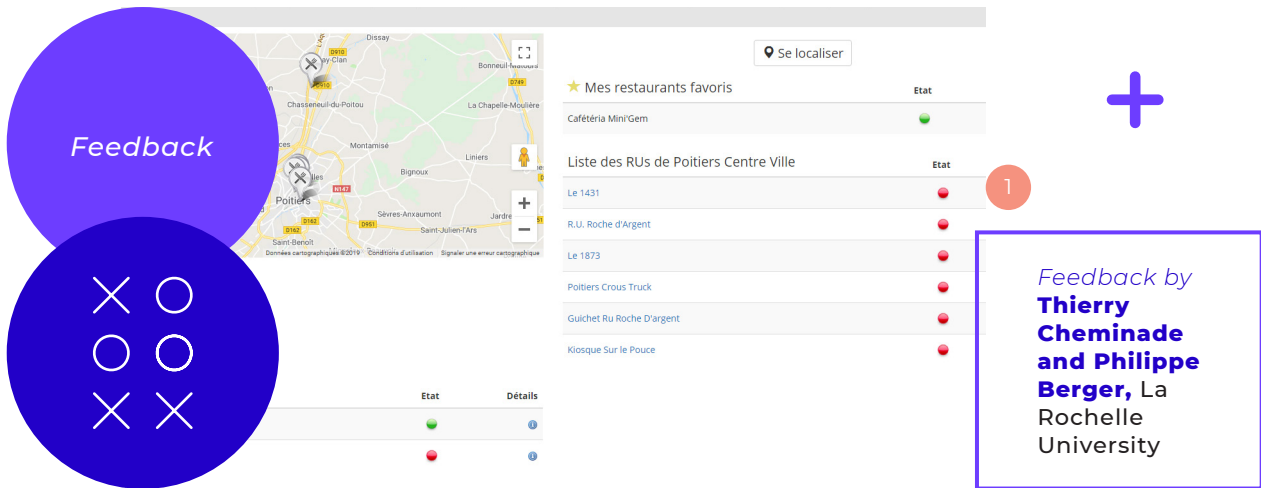
<https://data.enseignementsuprecherche.gouv.fr> : ministerial exhibition platform of open data

<https://data.esr.gouv.fr/> : portal federating all data resources on higher education, research and innovation (datasets, datavisualisation, dashboard, publications, API, ...)



<https://scanr.enseignementsuprecherche.gouv.fr/> : scanR, the engine of research and innovation which makes it possible to analyse the French research and innovation landscape

<https://data.esr.gouv.fr/data1ercycle/> : the career paths and success of holders in bachelor's degree, PACES and DUT



Use of open data from Cnous in the service of our students and staff

Historically: In 2014, the Centre national des œuvres universitaires et scolaires (CNOUS), the public establishment of a French administrative nature which manages university scholarships and housing, will make data available from different actors, mainly from its regional centers (the Centre régional des œuvres universitaires et scolaires (CROUS)) but also from its partners and individuals. This data, stored on a parking server, is accessible to all and available nationally. They concern jobs and student accommodation as well as descriptions of university restaurants and the menus on offer.

The CNOUS already offers access to this mass of information through the Jobaviz and Lokaviz centres, two portals that allow users to consult jobs and student accommodation. Unfortunately, these services suffer from a lack of visibility and are unknown to the student population.

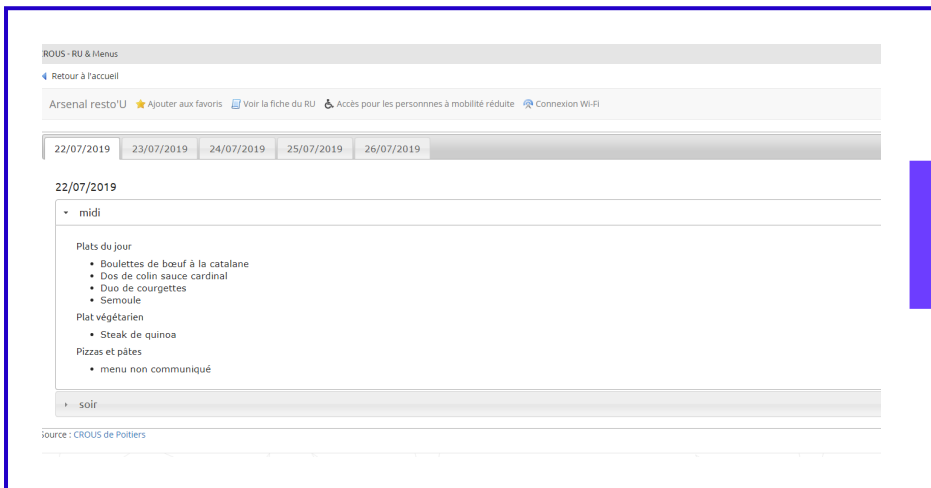


To solve this Poitiers's CROUS and La Rochelle University decided to make that information accessible as close as possible to the user by integrating it in the Environment Numérique de Travail (ENT) through new and specific services.

Beforehand, a common work consists of establishing and structuring the flow of data which has then to be used by those future services: at first in an HTML format, then in JSON to optimize the size of the files and thus allowing better performances.

Finally, La Rochelle University oversees the development of two web apps (of portlet type) based upon Twitter Bootstrap:

- **Esup-crousannonces: to display ads about housings and jobs provided by the data centers Lokaviz and Jobaviz**
- **Esup-dining: to display university restaurants and their menu**



Menu of a university restaurant

Focus on the "esup-dining" service

Few prerequisites are necessary for its installation: simply have an ENT ESupPortail / uPortal V4 or higher and a DBMS (PostgreSQL or MySQL).

Some settings are necessary before it is started: indicate the user role who will be able to access the service in administrator mode (back office), define the URL of the JSON stream distributed by the CROUS and corresponding to the geographical area of your establishment.

Once the portlet is deployed and configured, the esup-dining service can:

- Geolocate the nearest university restaurants (UR)

- List the UR available in the geographic area
- Display generic information for each restaurant: contact details, accessibility, opening hours
- Display the menus served

The user can also customize this service:

- By marking a restaurant as a favorite
- By indicating his nutritional preferences and / or the foods to which he is allergic (*)

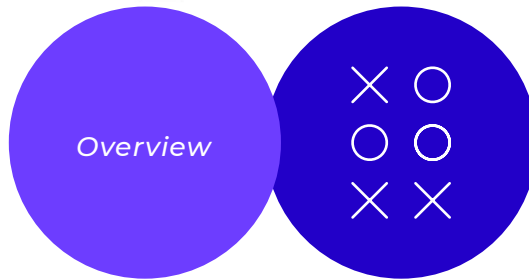
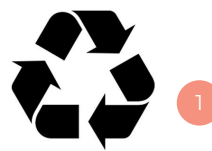
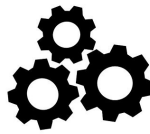
(*) Please note that this functionality is available in the portlet but the information may be missing from the feed: some authorities do not wish to take responsibility for classifying allergenic foods.

FOR MORE INFORMATION

Esup-crousannonces and esup-dining are branded "esup" because La Rochelle University put those services to the disposal of the consortium ESUP-Portail therefore they are proposed in their catalog.

You can find the sources on ESUP-portail's github:
<https://github.com/ESUPPortail/esup-crousannonces>
<https://github.com/ESUPPortail/esup-dining>

F indable A ccessible I nteroperable R eusable



By
Amue-DSPSI

Data at the heart of Open Data

The acronym FAIR (Findable Accessible Interoperable Reusable) DATA alone summarizes some of the main characteristics that open data should have. We will detail them here.

Main characteristics

Open data must above all be primary data (in the sense that it has not undergone any transformation), not aggregated as long as they meet regulatory requirements (GDPR, anonymization for example). The timing is also an important point: open as early as possible (notion of freshness), it is important to respect the timing announced for its dataset to facilitate processing by consumers. These two notions join the theme of the data life cycle. Finally, the completeness and quality complement the characteristics of open data.



The eligibility of a dataset on opening is measured by one of these characteristics.

Open, so accessible

By using the data.gouv.fr platform (see article on page 8), the data producer is encouraged to document his dataset, thus making it much more comprehensible to a third party. This documentation is for the producer the opportunity to describe their dataset (diagram, publication frequency, ...) and to master the strategy of re-use of their data by one or more consumers.

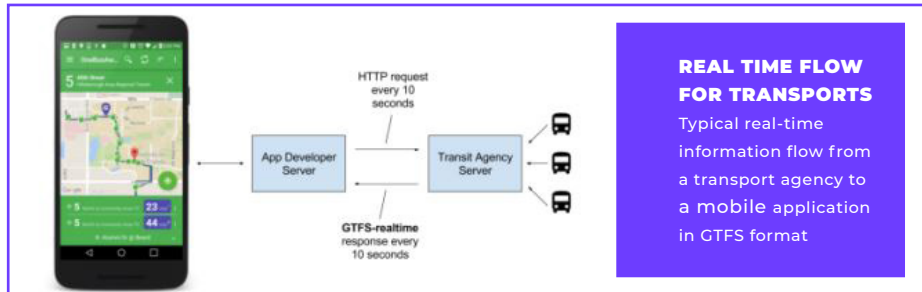


This platform also makes it possible to make the data sets accessible and usable by the greatest number of potential users and makes them more easily exploitable automatically by machines. Finally, datasets must not use proprietary file formats (such as xls for example) but open formats (CSV for example).



Standard and opening

Open data must be based on open licenses allowing its use by third parties. To choose the right license, follow the instructions at <https://www.data.gouv.fr/fr/licences>. For its open data, the Amue has chosen the Open State License 2.0 for its open data. Reuse will also be facilitated if the datasets are based on standards. Let's mention for example the GTFS standard for publishing transport datasets (see here) or for geographic data, please refer to the specific page of geo.data.gouv.fr.



Repositories

To simplify the use of datasets and, more importantly, the cross-referencing of datasets, it is important that the constituent data be based on repositories. This naturally makes the set of interoperable data. These repositories, themselves published in open data, are for example those published by INSEE, the national address database, the UAI (Unité Administrative Immatriculée) which replace the RNE code, etc.

In a decisional IS target

How to deal with aspects of standardization, repositories, data quality without mentioning the case of a business intelligence information system (BIIS) project in an Open Data context: it would be both a consumer of data opened by others (e.g. external repositories and indicators) but also a central point of opening data as long as they remain non-aggregated. These data will then benefit from the cross-functional consistency of a BI project (e.g. quality, source identification, data timestamping, standardisation). Depending on their strategy, institutions could even open their indicators to open data

TO GO FURTHER

*Tim Bunders Lee's 5-star deployment model > to be seen at : <https://5stardata.info/>
From Open Data to Big Data by Dataactivist Cooperative Society : <https://dataactivist.coop/dataliteracy>*



2

Photo by SangyalPundir

Open Data and french Higher Education : Opportunity to create new services





Openness to assert its sovereignty over public data

By **Jérémié Valentin**, Open Data Project Manager, Montpellier Méditerranée Métropolis

A first phase launched in 2011

Initiated in 2011-2012, the policy of opening up data in the Montpellier metropolitan area was above all focused on the potential creation of innovative services. In order to do that, the aim was to accumulate a large amount of data as quickly as possible, regardless of its quality and value. This first phase ignored the role of open data as a transparency factor in public action and as a tool for internal synergy that enhances the value of the work of the community's agents. Now, eight years later, the quantity and especially the quality of the data make it possible to affirm that an open data policy can, among other things, allow the creation of innovative services for the community and its territory.

A portal for the 31 municipalities

Nowadays, the Montpellier Metropolis has overcome, but not abandoned, this obsessive quest to create services thanks to the evolution of the regulatory framework in 2016 (Loi pour une République numérique) and the Metropolis is now in the process of proposing a basic public service, useful to businesses, researchers and citizens. This change of paradigm has led the Metropolis of Montpellier into a new phase of the project, by proposing the open data portal as a free service offer for the 31 municipalities of our territory. The idea is to regroup and homogenize the data of the 31 communes on a single portal. This work of acculturation to the data was accompanied by an awareness of the quality of the data and its freshness. This is why the Metropolis has multiplied initiatives to open up data with high added value as well as data that is updated in real time. These are the minimum conditions to see the creation of services with durable economic models.

Moving towards a smart city

Simultaneously, it was clear that the opening up of data alone did

not represent the stakes of data for a community, but that it was a brick in the intelligent city policy that the Montpellier metropolis had been experimenting with since 2010. Indeed, an open data platform is still a technical device and it is necessary to organize its administration in order to respond to the problems of the smart city, both internal and external.

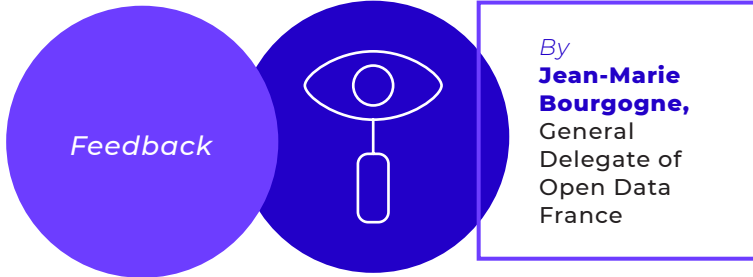
Today, in its Smart City roadmap, the Montpellier Metropolis is "turning towards urban data management, to ensure its quality, durability and dissemination". The metropolis is a trusted third party in terms of respect for privacy, independence from third-party actors and the dissemination of data. In order to apply this independence and its sovereignty over data, the Metropolis applies a non-privileged and non-discriminatory access to its data, via the unique open data diffusion channel.

L'OPEN DATA A MONTPELLIER MEDITERRANEE METROPOLE

The Open Data portal of Montpellier Méditerranée Métropole:

<http://data.montpellier3m.fr/>





The re-use of public data

Committed to open data for nearly 10 years, more than 450 local authorities have already opened their public data. Thousands of data sets are thus available to citizens, researchers or companies that operate them more easily, can create new uses or enrich existing services.



Open Data and french Higher Education : Opportunity to create new services

Communities are still struggling to communicate how these data are used in "open data". Some websites often present cases of exemplary uses but this is generally only a small visible part of the uses of open data. A more in-depth study reveals that many services regularly use multiple data sources including a large part available in open data. Moreover, re-users are not exclusively starts-ups but an infinity of public, associative or private actors including old companies or very powerful ones. As part of its territories open data observatory, OpenDataFrance published a summary of cases of reuse of public data through thematic (mobility, environment, employment, etc.) and interviews, thus showing the usefulness of the open data and the dynamism of the ecosystems related to their use (<https://frama.link/reutilisationsOD>).

What are the main lessons from this survey?

1. The underside of reuse

Open data uses look like an iceberg. On the one hand, there is a visible and obvious part, especially the reuse of the communities themselves. But on the other hand, especially when we overcome local platforms, we can find less known uses. Some of them (geographical data in particular) existed before open data: only the access mode has changed. Others are poorly identified due to the lack of downloads and Apis traceability. Also, if you look at the reported re-uses -as reported on community portals - they may seem weak. If you consider these "hidden" re-uses, they are much more significant.

2. Some sectors drive uses

Every day, millions of people ask for route calculation applications, like Google Maps or City Mapper and all those created by territories or start-ups. These applications are mainly based on open data. In the urban sector too, architects, planners and developers are big consumers of territorialized data that open data contributed to make more usable. The driving force of these areas is considerable and represent an oil task effect on related subjects. One will cite



parking, cycling or car-sharing for mobility. But this is also true for every data associated with the plot (regulation, addressing, network service...) or the events (public buildings, attendance, handicap...), for example.

3. A phenomenon of intermediation

Apidae, Navitia, Opendatasoft, Ozwillo, Openagenda... and all thematic data platforms are really appreciated by re-users to whom they simplify research and data exploitation. What is their strength? Gathering open data (and sometimes non-open), on a single platform and offering them as ready-to-use Apis or widgets. As a result, however, the communities are not familiar with the reuse of the data exploited via Apis, and in some cases, the re-users themselves do not know necessarily that the service (paying by some platforms) is produced from open data.

4. The "Mr Jourdain" of open-data

More and more services are created by programmers from « bricks », which are software or data. If these programmers know about the API, they are not necessarily sensitized to open data. The case of OpenStreetMap, boosted by the increase in GoogleMaps rates, is a perfect example: we went from a militant contribution to a massive contribution (including from institutions) without the contributors necessarily knowing that this is one of the largest open data bases in the world. To produce open data (and use it) without this knowledge is perhaps a sign of maturity coming from a sector which left activism behind.

5. Crowdsourcing and open data: A virtuous cycle

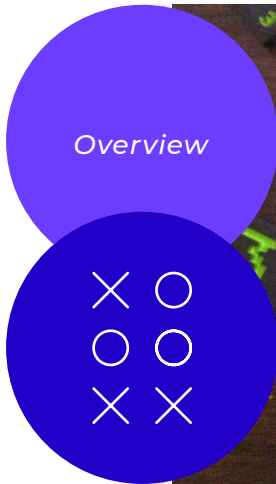
Beyond to OSM, more and more open data projects are associated to crowdsourcing, in others words to an engagement of residents or local actors to enrich and complete "common" data. This engagement which is facilitated and reinforced by the generalization of smartphones, is particularly strong in the areas of disability, mobility and tourism. The engagement of residents or local actors contributes to the creation of a virtuous cycle because the more a data is used, the more it is corrected and enriched, and the more it can, thus being more (re)used.

The association OpenDataFrance was created in October 2013. The aim is to bring together and support engaged communities actively involved in a process of opening up public data and to encourage all the procedures by this communities to promote open data.

This association is both animator and federator, and acts on different projects:

- it provides its members with information, advice or other support necessary for the opening of public data ;
- it encourages common productions from working groups or workshops on themes defined in general consensus as providing answers to the questions that arise on open data and implementation. It has thus produced a pedagogic and frame of reference freely accessible online, set up an ambitious program of local support for local authorities (OpenData-Locale), and finally, it has implemented a set of tools aimed at simplifying production, improving quality and standardizing public data (OpendataFactory).
- it represents its members in all public and private authorities in order to ensure the harmonious development of open data.





By
Amue-DSPSI



1 Photo by justgrimes via visualhunt.com

In concrete terms, what would open data bring to institutions? A compilation of uses

It is impossible to be comprehensive in the field of Open Data, as the usages are so varied. But there is nothing like presenting actual cases, completed or in the planning stages, mainly in the field of Higher Education and Research, to help understand the subject of Open Data.

Harvesting open-access databases to improve the quality of their data

Harvesting data from open repositories, such as those released by INSEE (French national public company for statistics and economic studies) on company information, makes it possible to qualify, correct, and purge company-related data in its Information System (suppliers, CRM, repository for internships, etc.). In the same spirit of improving data, using the National Address Database provides quality control of its address data under a single 'address' for this purpose: <https://adresse.data.gouv.fr/>. Logic is to be taken into account as early as possible in IS projects to guarantee quality data from their initial input.

Also in terms of quality control, open-access data may encourage users to request corrections or additions to the published data. Provided that a process for taking this feedback into account is implemented, openness can become a means for correcting one's own data.

One example is the distribution of the geolocation of public transport stations for which the public could recommend updates, an interactive fix by users.

For better services to users

Notifying service operators of variations in population flows by campus (mass arrivals for an international confe-

Open Data and french Higher Education : Opportunity to create new services



rence, unusual exam times, temporary campus or site closures, seminars...) can help them adapt their services. Open Data is a possible means of transmitting this information. Public transport operators in medium-sized cities serving campuses can therefore optimize transit for a better quality of service to riders (students, staff and guests) by adapting timetables and transport frequency to meet the reality of the number of riders. This same standardized information can also help campus catering services by adapting their offerings as needed.

An advantage for the employability and attractiveness of a region

In order to attract companies that recruit executives, an institution could publish in advance the number and qualifications of its next cohort of graduates. Likewise, the next logical step would be to publish the regional data that showcase the laboratories, research subjects, number of researchers, major equipment, ... all to attract talent and companies to a geographical area.

An educational institution's publishing of its graduates' professional integration data is a way of giving value to its training courses, which can be then used to show the benefits of a diploma. Taking this logic further, recruiters could use the data (which jobs were filled over time by which graduates, anonymously) to determine which degrees were the most common in certain types of jobs and in this way target those degree holders to fill those jobs.

Information from the institution

It is easy to publish institutional records in Open Data (report of the Board of Directors for example), as well as social and HR data from the social audit report, allowing one, for example, to highlight an optimum parity. In the logic of transparency and with a singular strategic willingness, an institution could publish the subsidies it grants to different associations.

Information on events organised by the institution or by university associations can be useful for regional operators (transportation, civic improvement, cultural services, etc.) or for applications intended for citizens, etc. The systems are already in place (an advantage of open data) that could be used to disseminate the information through the OpenEventDataBase project.

The University of Paris Nanterre publishes its training courses on the site data.gouv.fr. Although this information is generally accessible on institutional sites, open data enables the data to be used for analysis, data reconciliation, etc.

As for registration and graduate data, they are automatically published in open data by The Higher Education Ministry via its #DATAESR site (see article on page 6).

More possible uses

Taken to its extreme and generalized to the whole of the ESR, open-access data could become a means of responding more easily to surveys, and would allow, without too much administrative effort, self-assessment in relation to similar institutions, from an organizational and mission point of view.

Some uses outside of Higher Education and Research

Mobile applications are large consumers of open data and today optimize journeys according to the parking spaces actually available

(open data in real time), mark locations of self-service bicycles and their availability, and update themselves according to real-time traffic information, alerting you to the imminent arrival

TO GO FURTHER

Data opened by Pôle Emploi
<http://www.pole-emploi.org/home/news/news/infographics/open-data-des-data-for-elaj.html?type=article>



2

val of a tramway. We refer here to existing uses in cities. The logic of "Smart cities" that exploit digital technology to improve the quality of services is an inspiration for future intelligent campuses, which have yet to be invented.

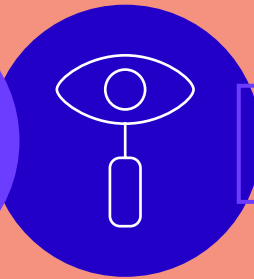
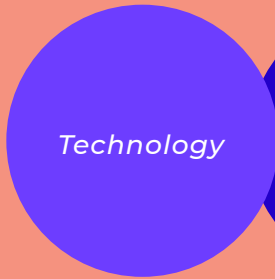
Recently at the top of the rankings of most widely used data sets, the DVF "Land Value Query" allows the values of real estate transactions by address to be known. Heavily used by sellers, prospective buyers, and professionals to know the real prices of the real estate market neighborhood by neighborhood, this open source has been a minor revolution in the world of real estate sales. Finally, major projects such as OpenStreetMap (an online and global mapping project, an open competitor to Google Maps) or openfoodfacts (a directory to food products from around the world) combine Open Data and collaborative citizen contribution (productive participation or "crowdsourcing").

In conclusion

These examples can be supplemented by your contributions and good ideas to be sent to numerique@amue.fr. All these cases and feedback could lead to the publication of "the beautiful stories of open data ESR", highlighting the benefits for users for example.

TTO IMAGINE YOUR SERVICES

You can go to data.gouv.fr and look at some of the 38,000 data sets.



A recommendation
by **Amue-DSPSI**

Hosting, technology, format: the dark side of open data



Locally...

In building its open data strategy, each institution could look for its own data distribution platform. A market solution specializing in data dissemination, a domain on the institution's site, a dedicated server open to the world, a partnership with the city, ...

These are all valid questions for the long, even medium term, but in the first phases of implementing open data, it seems more important to deal with other aspects than the creation of a dedicated and individualized technical environment. Amue is considering partnering with an institution in order to introduce a distribution framework that can be configured by each member according to its own unique digital strategy.

... or nationally on the site data.gouv.fr

The simplest way to enhance and share open data is to use the platform supported by Etalab (see article on page 3): data.gouv.fr. In France, this is the reference platform, where each operator can deposit their data and where users come first to gather it. It is easy to use, the site is documented, and the reporting mechanism for a dataset encourages good structuring and documentation, thus facilitating its use and referencing.

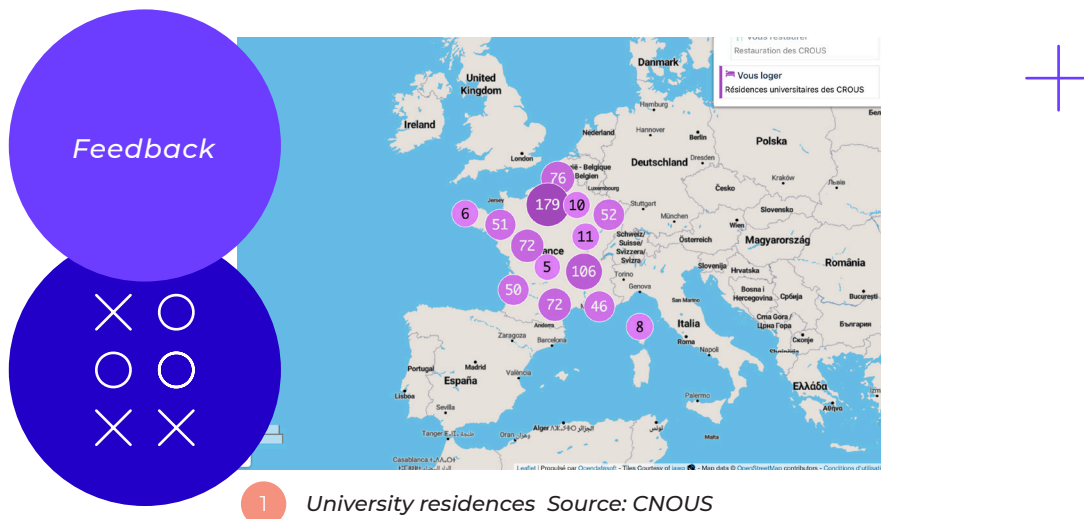
Publishing, but in what format?

This State platform also includes all the documentation needed to choose the technical format best suited to its context and to each of its datasets. Formats / standards such as CSV, JSON, XML are defined but also formats specific to the nature of datasets (transport data, geographic data, ...).

Some technologies

In addition to the manual filing and updating of data sets, the data.gouv.fr site offers the use of harvesting or API technologies for example. This makes it possible to automate the interactions between its Information System and the Open Data reference point that is data.gouv.fr.

Conversely, if you wish to display one or more datasets on a web page, this is documented here. You will even find a JavaScript library, MetaClick, allowing you to present your Open Data catalogue from data.gouv.fr.



The open data catalogue of the CROUS (Centre régional des œuvres universitaires et scolaires) network

By **Stéphane Papon**,
OpenData Referent at the CNOUS

The CROUS network has had several web services and open data streams available to the higher education and research community for several years.

The beginning of our reflection on the availability of our data came about following a complaint from a student who had seen information on a university site about accommodation that were erroneous. A conscientious staff had manually re-entered the information from the local CROUS website with precision, but over time the page of the site that was left blank, without updating, became a source of obsolete information.

It was therefore necessary to allow our partners to have our captured data as close as possible to the information.

This should be done easily, freely and with a real-time updating process to ensure that the data is always reliable.

The first "restaurant" feeds were generated by the CROUS pilot site in Poitiers for the purposes of working together with the University of La Rochelle as part of the development of



its Digital Workspace (ENT: Espace Numérique de Travail) (see article Page 7). The data flows thus make it possible to display in the university ENT information of the restaurants (schedules, geolocations, information & menus). In the extension of this project, we have declined the catering flow for the other services of the CROUS on the Poitou-Charentes region, thus were generated among others the flows housing, news, job announcements, housing announcements...

Our broadcasting system has always been designed and thought out for a national and not just local reach. So we were able to duplicate these feeds very quickly for all the Crous.

Today, this open data is mainly consumed by our information system (CROUS Mobile, digital signage, website ...) but also by different partners: universities for their ENT or mobile applications, municipalities or communities of agglomeration for their websites, digital signage or mobile applications. Our data is now generated by specific applications closest to the field and information, and stored on our own servers. We have a replication on data.gouv.fr of certain feeds (do not forget to register as a subscriber when you consume our data, this allows us to keep you informed of developments).

An example for restoration/catering

The restaurant secretary's office has an application developed by the CROUS network that allows it to enter all the information necessary for the operation of the restaurant (schedules, menus, opening/closing, events, etc.).

This application, among other features, to generate its menus in PDF for sending emails to partner sites, but above all it generates the feeds that are sent to the open data server of the Crous and consumed by the website, mobile applications, digital signage...

The efficiency of this system makes it possible, for example, when a supply shortage is noted in the kitchen - effectively impacting the menu of the day - to modify the display in real time on all our channels with a single input modification by the restaurant's secretariat.

Available resources nowadays

- The available data catalogue of September 2019 is composed of flow and web services: catering/restaurant flow (XML/JSON), geolocation, schedule, information... , program flow (XML/JSON), housing flow (XML), geolocation, schedule, information..., news flow (XML) regional news from the CROUS (images and text), scholarship flow (XML) (information), event flow (XML being finalized).

- web service restaurant: Geolocations, menus, schedules, information...

- web service to announce housing: display of the X latest housing announcements, possibility to choose the area (France/region/department/university cities)

- Job advertisement web service: display of the X latest student job advertisement, possibility to choose the area (France/region/department)

Available resources of tomorrow

Several development and evolutions are to come upon the last trimester 2019 and the first semester of 2020:

- the migration of every XML flow into JSON upon the same principle as restaurant flow,

- the implementation of web services on all flows,

- the migration of the server with the implementation of an open data server, in tune with the latest developments and the current needs,

Our flows are lively so don't hesitate to inform us if you exploit them or wish to do so; it will allow us to keep you in touch with the latest developments (without our service being interrupted).

We are at your service if you need more content, or if you wish to implement new CROUS data. We are favorable and open to exchanges that might allow our flows to perform better technically and qualitatively speaking.

We can make specific flows on demand if you have specific needs, if those flows can be reused for the whole community (for instance, flows combining restaurants and housing for the whole country and that serve as support for the mapping of the etudiant.gouv.fr website.)

**To contact the
CNOUS on this
matter:**

opendata@cnous.fr
stephane.papon@crous-poitiers.fr





1

*Etablissement étrangers en relation avec Paris 13
Données: Wikidata
Carte: contributeur OpenStreetMap*

Collaborative frame of reference in open data

In our information system, we have references to data our establishments don't have control over. For instance, high school references. If those high school attribute (name, address, ...) change, we are not made aware of it. That's why we would rather adopt a national frame of reference, rather than to accept the charge to keep those data up to date. Internally, it is very uncomfortable for a service to be responsible of data maintenance without being involved in the process of modifying information. These frictions can be handled by improving the process, opening the silos, interconnect applications...

*By
Pierre Boudes,
VP SI université
Paris 13*



Some frame of reference without a central authority?

But what if data are not produced under the control of a central authority? This is the question raised with foreign higher education establishments, which we have to reference, for instance for the composition of examination boards of thesis. A good solution is to switch to an open and collaborative international frame of reference.

The essential infrastructure of knowledge

Wikipedia, created in 2001, turns out to be particularly efficient to collectively produce information that are reliable and qualitative. Some of this information are semi-structured and BDPedia benefits from it to make it available as linked open data (see text in the box), thus constituting since 2007 a central brick in the silent revolution of semantic web.

Wikidata completes the process by allowing contributors to directly produce and improve the quality of structured data, according to the logic of the frame of reference. The assumptions about an element are often supported by a link to the source of information, and to contribute is especially easy.

At the university of Paris 13, we had to revisit our data on foreign establishment, as part of the deployment of Sinaps (an Amue solution for the management of data reference). We chose to use Wikidata as an external frame of reference. As such, we do not endorse the responsibility to track changes concerning establishments (name, address, merging...) and a simple Qwant research with the key word &wd enables to find an establishment quickly. Before/Formerly, the entire/complete data had to be produced from documents in foreign languages, which led to mistakes; for instance, a British university with a Columbian code (will you manage to find the issue here?). When European universities are built, Wikidata provides a common frame of reference, both multilingual and trustworthy.



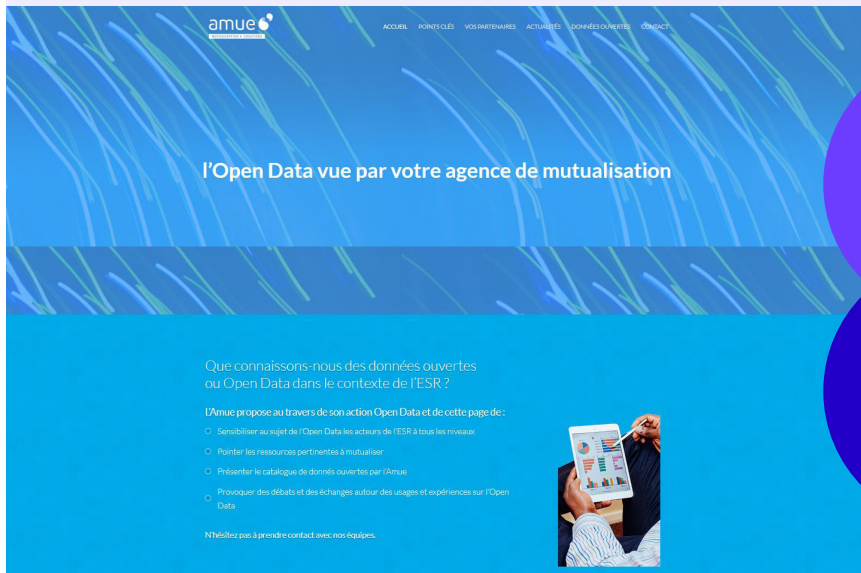
2

Wikidata

WIKIDATA

Every data is a triplet subject-property-object which, in a table, would be equivalent to the cell at the 'subject' line, containing the 'object' value; but without a table! It makes it easy to organize heterogenous data and to combine multiple values to the subject for a same property. Those triplets are communicated in an open and interoperable manner, making arbitrarily complex and multi-based requests possible. The users ID (of the URI) are requestable in HTTP and designate subjects themselves and not their codifications in a base. A lot of data are exploitable directly with the tools produced by the community around Wikidata (60 millions subjects), which you can find at <http://tools.wmflabs.org/>.





1 The Amue's OpenData page

Amue and Open + Data now in unison

The Amue has positioned itself on Open Data on the basis of two types of actions: awareness and anticipation.

By
Amue-DSPSI,

Awareness

The first line of action consists in supporting institutions through awareness rising and experience sharing operations, of which this issue of the digital collection is the first achievement. An "Open Data" day is scheduled (see insert) on 14 November 2019 at the Maison des Universités. The morning will be dedicated to an overview of Open Data in relation to the Higher Education and Research context. Workshops will be organized in the afternoon. They aim to collectively produce a first list of data sets to be opened by analyzing added value, ease of implementation, sensitivity, ...

Anticipation through the pooling of tools and ...

The second axis aims at anticipating the requests for opening data contained in the Amue software solutions. Indeed, part of the information legacy of the institutions is managed through these tools. By pooling the extraction and format, it will be an efficiency advantage for the institutions that implement their Open Data policy. These works will allow existing solutions (Apogée, Sifac, Siham, Sinaps, ...) to propose data publications under the control of field experts. It will also allow solutions in construction (Pc-Scol/Pégase, Caplab) to be anticipated in their roadmap.

**TO REGISTER ...
... to the Open Data day organized by the Amue: November 14, 2019 >register**



...data patterns

Open data are characterized by documented data patterns (what are the different fields, how are the data represented, what are the possible values, etc.). These will obviously be pooled regarding the work related to the software solutions of Amue or co-constructed with partners (Cocktail and CNRS). It also seems interesting that institutions mutualize and share data patterns that would not be directly in the management domains covered by the Amue solutions. Finally, the community decisional IS project will also be involved so as to share its expertise in terms of data quality, data pattern pooling and especially in the transversal coherence of the data.

The more the patterns are common and shared throughout the territory, the better the consumers can offer value-added services through the consumption of these data. This approach to sharing pattern is in line with Etalab's one through its website <https://schema.data.gouv.fr/>, which allows data patterns to be referenced (see insert).

This logic of working together by and for the institutions would make it possible for institutions to create a common foundation to facilitate their Open Data approach. Working groups could co-construct this base.

A new website: data.amue.fr

As any public institution of more than 50 agents, the Amue is concerned by Open Data for its own data. So, in a progressive logic, 3 first data sets are open. These are the list of our members, the digital collection and finally a dataset proposing the action plan documents and activity reports. These datasets are published on the State website data.gouv.fr. One page on the open data topic has been put online: data.amue.fr: it briefly presents the topic and references the datasets published by Amue.

FOR MORE INFORMATION

site for referencing public data patterns for France supported by Etalab <https://schema.data.gouv.fr/>



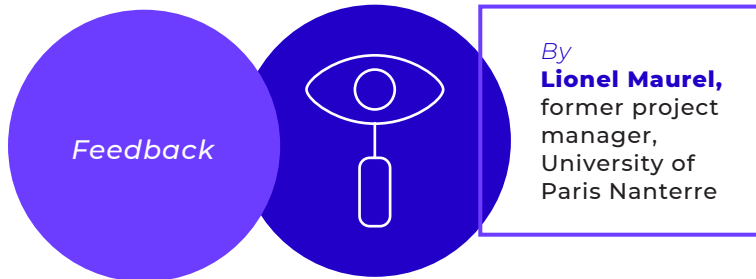
```

  0:
    name: "id_lieu"
    description: "Identifiant du lieu de covoiturage, délivré par le point d'accès national"
    example: "35238-C-001"
    type: "string"
    constraints:
      required: true
      pattern: "^[013-9]\\d|2[AB1-9])\\d{3}-C-\\d{3}$"
  1: {}
  2: {}
  3:
    name: "com_lieu"
    description: "La commune / le lieu-dit du covoiturage"
    example: "Rouen"
    type: "string"
    constraints:
      required: false
  4:
    name: "insee"
    description: "Le code INSEE de la commune d'implantation"
    example: "76540"
    type: "string"
    constraints:
      required: true
      pattern: "^[013-9]\\d|2[AB1-9])\\d{3}$"

```

2

Example of a data schema



Feedback on an Open Data project in an institution



Open Data: let's not forget administrative data!



When the issue of Open Data is raised in the field of higher education and research institutions, it is usually about research data. The National Plan for Open Science, published by the Ministry of Higher Education, Research and Innovation in July 2018, thus contains an axis entitled "Structuring and opening research data", while a growing number of institutions are developing sharing policies and institutional repositories to store and disseminate them.

This movement in favor of openness is commendable, but it is unfolding while leaving aside what constitutes the basis of the Open Data policy, as we have seen it develop in France for several years now. Indeed, it is primarily the administrative data of higher education and research institutions that should be disseminated online and made freely reusable. Since 2016 and the adoption of the Digital Republic Law, this has even become a legal obligation, since this law compiles public administrations to a "default principle of openness" of the information they produce within the framework of their activities.

By focusing attention on them, research data play somewhat the role of the tree that hides the forest of academic and university data. By "administrative data", we mean the current operating data of higher education and research institutions: budget and public procurement data; statistics on student populations; lists of training courses offered; data from the social audit relating to administrative staff; information relating to the location of buildings and the geolocation of services; documentary data relating to the activities of libraries, documentation centers and university presses; information on partnerships with companies and international actions, etc.

This list could be even longer, as all the departments of a university - right down to the university canteen! - actually produce data. Surprisingly enough, however, Open Data projects involving administrative data in universities are few and far between, and the institutions that are embarking on them are the exception.

A pioneering project at the University of Paris Nanterre

Even before the adoption of the Digital Republic law, the University of Paris Nanterre initiated an administrative Open Data project, linked to the particular local context to which it belongs. The university is located in the Hauts-de-Seine department, which is one of the local authorities that developed a proactive policy of opening up public data very early on. Moreover, the two entities were linked by a partnership agreement including a digital section in which a point relating to Open Data was included as early as 2015. The department proposed to make its Open Data portal available, which would exempt the university from having to set up a dedicated website. In return, the university undertook to make a certain number of data sets available each year.



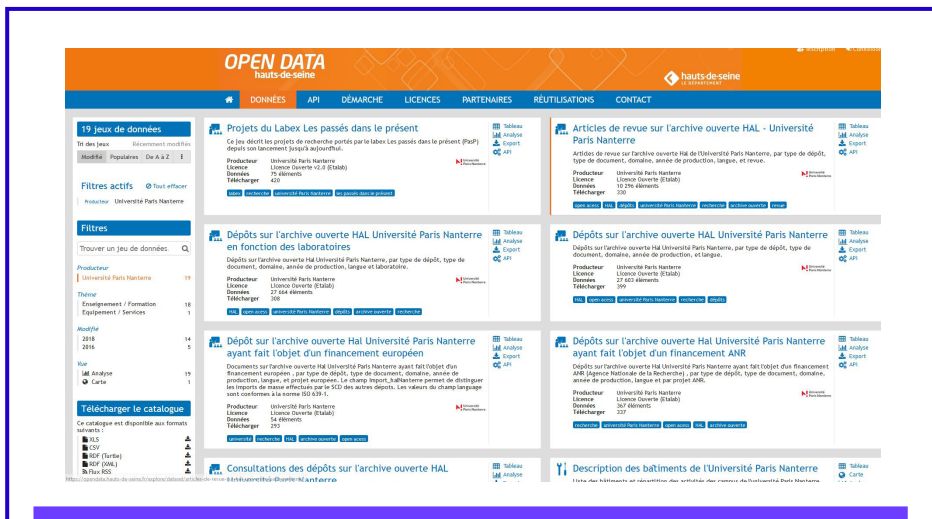
To be carried out successfully, this approach was conducted in project mode, bringing together a team combining various skills - administrative staff, researchers, librarians - and relying on the participation of students through internships. The task of this group was to make an inventory of datasets that could be opened and to collaborate with the producing services so that they could be reworked and documented for dissemination. An important aspect was also to obtain validation from the political bodies of the university before the data were published.

The Hauts-de-Seine Open Data portal currently distributes 19 university datasets, placed under the Open License allowing their free reuse, each having been downloaded several hundred times. In selecting the data, the emphasis has been placed on information relating to the training and diplomas awarded at the university, the activity of the doctoral schools, the research structures linked to the establishment, as well as statistics provided by the Common Documentation Service.

Such a project has also revealed the difficulties and obstacles that an Open Data approach can face within an academic context. More than a technical issue, openness is above all a governance issue as it requires the involvement of many departments related to all parts of the university that are not necessarily used to working together. The cost in human labour to be provided is also important, not only for the collection but also for the annual data updating. Certain imperatives must not be neglected, such as the protection of personal data and privacy, which requires anonymization of the games, which is sometimes complex to obtain.

These considerations would argue for universities to work together to jointly develop Open Data policies. Following the example of what has been achieved for local authorities within the framework of the Open Data France association, one could imagine, for example, that a "common university data base" could be defined, which could be the first step to be taken by institutions wishing to launch open data projects.

FOCUS ON DATA.HAUTS-DE-SEINE.FR



The data are available on the pendata.hauts-de-seine.fr website.

Open Data and french Higher Education : Opportunity to create new services





1 The developer view: image by Julien Delmas

By
Julien Delmas,
teacher

Annuaire-education.fr, an experience of open data of the world of education

Teacher for about fifteen years, web developer in my spare time, I recently discovered the large quantity of open data available on the platform of the Ministry of National Education data.education.gouv.fr.

Aggregating open data

I noticed that many sites were using this data, but in a very fragmented way, such as the school yearbook, the performance indicators of high schools or the organization of school time carried out by the Ministry of National Education. The Caisse des Dépôts et Consignations, a public group at the service of the general interest and economic development, proposes eCarto "the observatory of territories and educational digital" and many private companies are interested in the results of exams, to better classify high schools.

I wanted to create a directory of all French educational institutions: schools, colleges, high





I wanted to create a directory of all French educational institutions: schools, colleges, high schools, local public educational institutions, that they are public, private, under and outside contract. annuaire-education.fr aggregates all existing information of more than 65,000 educational institutions: from contact information to results examinations, and even school curricula, maps of the same city institutions, male-female teachers' proportion or computer equipment.

The site was launched on August 26, 2019, and a few weeks after its opening to the public, many teachers or managers of National Education understood its interest. Its purpose is not to propose a comparison or a classification, but only raw information that everyone will interpret as he wishes to.

Of the 10,000 visitors in the first week, I had a lot of feedbacks from National Education staff which has indicated to me the importance of all this information when choosing an institution in which to apply for a transfer. They were unaware of its existence until this day.

Development perspective

From a more technical point of view, this directory works with the French content management system, SPIP. It queries no less than fifteen duplicate databases from the Open Data site of the Ministry of National Education. This represents nearly two million lines and 400 MB of data.

I was confronted with several difficulties in the development of this tool. Some databases were not really updated, despite daily maintenance operations. After I reported the problem, it was solved. The second problem, and not the least, is the poor structuring quality of the databases with duplicate fields and particularly obscure headings.

Promote open data

This directory is certainly the most complete site on French schools and all its information comes solely from the Ministry's open data. In view of the reactions received, it is to be hoped that this modest personal initiative will make the existence of this public data known and that it will encourage schools to update their information to the Ministry and to respond more regularly and more accurately to the various surveys, in order to make this open data even more reliable.

FOCUS ON ANALYSIS

These Open Data allow to put forward certain information and are very significant of the dynamics of a school establishment. For example, when we observe that some schools receive 30% of Latin students while others, much more favored, only convince 5% of students to choose Latin, we can wonder about the reasons for these differences. This is also the case with the percentage of pupils in the third year [in France] who are behind in school, which is an objective indicator of the real difficulties of pupils in a school.



Open Data and french Higher Education : Opportunity to create new services



L'annuaire de l'éducation

Autour de moi Pourquoi cet annuaire ? Contact @annuaire_educ

Entrer un nom d'établissement, ville, UAI, code postal, téléphone...

Collège Jacques Daguerre de Corneilles-en-Paris (0951099D)

Les informations affichées ci-dessous concernent l'établissement Collège Jacques Daguerre de Corneilles-en-Paris et elles proviennent directement et sans modification du site officiel Open Data du Ministère de l'Éducation Nationale qui les rend publiques et accessibles à tous gratuitement.

La dernière actualisation de cette fiche date du 30/08/2019. Pour demander sa mise à jour, cliquez ici pour compléter le formulaire sur le site du ministère.

COORDONNÉES

Collège Jacques Daguerre
Rue Emy les Prés
95240 CORMEILLES EN PARISIS

Code UAI (ex-RNE) : 0951099D
SIREN/SIRET : 19951099100011

Téléphone : 01 39 78 75 01
Fax : 01 34 50 64 54
Adresse de courriel : 0951099D@ac-versailles.fr
Site web : www.clg-daguerre-cormeilles.ac-versailles.fr
Fiche ONISEP : Fiche établissement

INTERACTIONS

Partager WhatsApp Twitter Facebook QR code

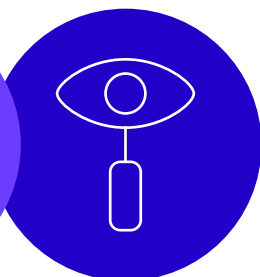
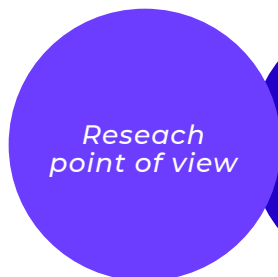
Naviguer Maps Google Maps Bing Maps

TOUS LES ÉTABLISSEMENTS DE CORMEILLES-EN-PARISIS (17)

Carte des établissements de la commune

2

One page in the education directory



By
**Bertrand
Mocquet,**
digital expert
at Amue

How the release of data could contribute to digital governance of the university or the establishment?

+

+

It is impossible to be comprehensive in the field of Open Data, as the usages are so varied. But there is nothing like presenting actual cases, completed or in the planning stages, mainly in the field of Higher Education and Research, to help understand the subject of Open Data.

Free flow of information

Open data is inspired by this goal of information circulation. When opened, data become exploitable, exploded, and there is a value creation around services based on these data. For instance, the “Find your campus” application was created in 2014 by StrasWeb, an association composed of students and young graduate and, until 2017, it allowed students from Strasbourg University to visualize which means of transportation were available to them to go to their campus. This is a good example of “putting data into circulation [which] renews modes to (...) while relying on third parties who generate services from these data” (Courmont, 2016), being in this case, a student association.

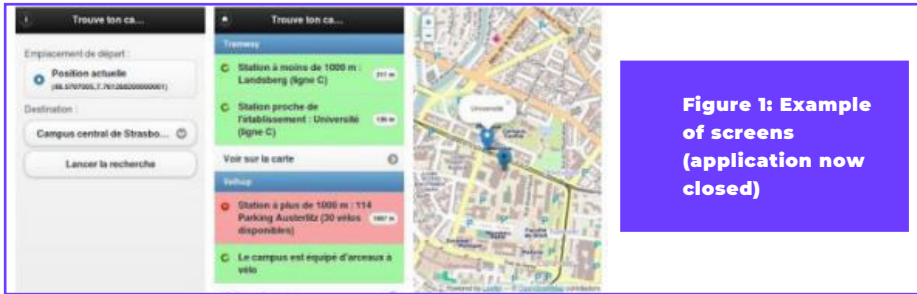


Figure 1: Example of screens (application now closed)

Three effects for this strategic topic

For our governance, in the sense of “governance or a set of arrangements implemented by the company to achieve effective coordination” (Como, 2013), when will digitization become a strategic challenge for the university or establishment? For us, it is a matter of acknowledging that “open data became a government issue the moment they were put into circulation”. (Courmont, 2016). This is the consequence of the opening of

data that we advise to capture from a university or institution digital governance point of view. Let’s look at some of the effects.

One of the first effects is to focus strategic attention on the quality of the organization a,d to perceive their wealth within its own organization: “attached to producers, preserved in “business” silos, data are instruments at the service of a public policy sector”. (Courmont, 2016). Opening them up, first internally, helps to de-compartmentalize its internal activity and to think about it in a more systemic and global way: this is how we easily become the consumer of one’s own open data.

One of the second effects for an organization is the "ability to control the use made of its data". (Courmont, 2016), and, as a consequence, the implementation of a “strategy” for regulating data flows. By choosing the modalities for putting data into circulation (licenses, formats, etc.), it encourages the attachment of its data to certain users or, on the contrary, to others” (Courmont, 2016).

Finally, one of the third effects would be to empower citizens, due to the fact that more and more people hope that “more and more actions and behaviors will be regulated by information flows” convinced that “if the public actor does not control these flows, he loses control of its public policies”. (Courmont, 2016)

These three effects, but the list is not exhaustive, of the opening of data on the digital governance, shows us that today might be the opportune moment to understand the contours of open data, and to look at it from a strategic point of view rather than from a legislative point of view.

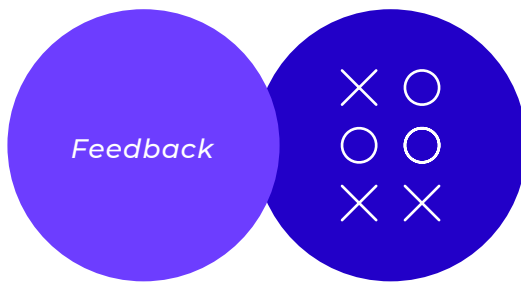
FOR MORE INFORMATION

Courmont, A. (2016). Politiques des données urbaines. Ce que l’open data fait au gouvernement urbain. Sciences Po, Paris.

Côme, T. (2013). La gouvernance des universités. Gestion et management public, Volume 2/n°1(3), 15.

Loveluck, B. (2015). Réseaux, libertés et contrôle: Une généalogie politique d’internet. Armand Colin.





By
Pierre Boudes,
Chairman
of the UNIF
Strategic
Committee

Experimenting the opening up of data on study courses



It is impossible to be comprehensive in the field of Open Data, as the usages are so varied. But there is nothing like presenting actual cases, completed or in the planning stages, mainly in the field of Higher Education and Research, to help understand the subject of Open Data.

A great absentee: the follow-up of actual study courses

However, in all the freely available data, there is a lack in terms of the following-up of individual study paths which indicates the sequence of steps actually completed, from year of study to year of study, diploma after diploma, and which takes into account inter-institutional mobility, or even interruptions and resumption of studies.



This tracking is simple to build using the National Student Identification Number (NSI). It provides with important information when it is set up within the institution and it would clearly highlight inter-institutional relations if it were available at this level. For example, one can discover internally that over the last ten years, 376 newcomers (those who arrived in the year of their baccalaureate) validated such and such a bachelor's degree in three years and how many of them chosen to continue their studies at the same university. By cross-referencing our data, we will also be able to find out how many went on to study at another institution.

Cohort follow-ups, generally carried out by student life observatories, provide indicators for piloting. Systematizing and automating these follow-ups, without statistical filters, will provide with new information on the functioning of the training (courses) and their relations. It is not a substitute for the expert analysis work to be carried out afterwards, but to provide reproducible and verifiable data, resulting directly from internal processes.

For guidance, an access to such data via suitable applications would enable the topic to be clarified in preparation for an individual interview and would provide a reference framework on which to base advice, which must be personalized in order to be relevant.



Let's do it!

The strategic committee of the UNIF proposes to every interested university (from the Île-de-France or not) to cross-reference their data on registrations and on students' success.

The goal is to quickly and experimentally make a first publication under an open license in view of perpetuating this process in the form of a permanent participation to the essential infrastructure of data that modern digital services need.

We are particularly thinking about third-party departments using diverse sources of free-of-charge data. For example, for a project concerning a guidance GPS offering a first-person view on continuity studies, real prospects and connection to people on the job, the UNIF considers giving necessary data to a partner in a nonexclusive way by making it free to use, in open data.

It is also important to know about the various mobilities between institutions to produce appropriate digital services, especially when data portability is concerned or if you make the more radical choice of self data, where every user controls his data.

The data we are considering for publication come from schooling Information Systems like Apogée. We can retrace the academic career of every individual (identified by the cryptographic mark of his national identifier code (INE in French) thanks to his yearly registrations. By grouping identical studies, we can count them and forget about individuals. The specificity of his baccalaureate or other data can be part of his academic career. We can also decide to disregard the year of registration and take into account the chronological succession only. We thus obtain marks that can overlap several institutions.

To avoid publishing re-identifying information, when there are so few identical academic careers, we can omit the exact cardinality (anonymity by generalization). For instance, when fewer than ten people have done the same studies, we will be in a position to publish the mark without stating the exact number (or by giving a mark of 1 person as in the marks published by the university of Paris 13). More accurate private data, including exact cardinality, will also be communicated to institutions that take part in the experiment.

To participate, you can contact us at the email address: <mailto:opendata@unif.com>

FOCUS ON UNIF

The Île-de-France digital university (originally called Digital University Paris Île-de-France) is an actor of mutualization on the digital topic that has taken the form of an inter-university service, nowadays supported by Paris 13 University. Its members are the higher education public institutions of the Île-de-France region, among which there are about fifteen universities and 4 Comues.



UNIVERSITÉ NUMÉRIQUE ÎLE-DE-FRANCE



Are you interested in the Open Data HE, would you like to contribute?

In a logic of pooling, experience sharing, and co-construction of solutions, the contribution of the of institutions is necessary for the Amue. You can contact the Amue digital team by email at numerique@amue.fr or participate in the Open Data Day on November 14, 2019.

September 2019

Managing Director of the publication: Stéphane Athanase

Editors in chief: Bertrand Mocquet and David Rongeat. Editorial secretary: Mire. ISSN 2650-8494

Couverture and model: @yaygraphisme

Editing by Lucas Champion

The digital collection is licensed under a Creative Commons CC BY-NC-SA 4.0 License.

Contributors to this issue: (a great thank you)

Stéphane Athanase, Marin Dacos, Romain Tales, Emmanuel Weisenburger, Thierry Cheminade, Philippe Berger, Jérémie Valentin, Jean-Marie Bourgogne, Stéphane Papon, Pierre Boudes, Lionel Maurel and Julien Delmas.

Acknowledgements to the Dataactivist cooperative company

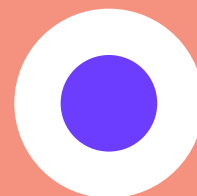
Thanks to the University Bordeaux Montaigne, and more particularly Soufiane Rouissi, in charge of the Master's Degree in Digital Educational Devices (DNE) at ISIC, and to the students of this Master's Degree for the translation into English of number 5 of this digital collection : Alizé Sibella, Apolline Gondeau, Elina Alves, Jennifer Gil, Jérémy Bessonies, Julie Métraud, Léa Bourasseau, Maï Ly Lifante, Maria Genetay, Rebecca Meddahi, Sophie Margnes, Thomas Bilancini; proofreading by Aadel Benyoussef, pro-fessional speaker (VP Europe Exceletrate Systems).

Publisher: Amue - 103 boulevard Saint-Michel - 75005 Paris

Made in France

All images and photos in this issue are <Copyright> and free of rights, rights reserved permission for use specific to this publication.internal processes.

For guidance, an access to such data via suitable applications would enable the topic to be clarified in preparation for an individual interview and would provide a reference



NEXT ISSUE OF THE DIGITAL COLLECTION

NEXT ISSUE OF THE DIGITAL COLLECTION (NOVEMBER 2019)

What about looking at the HE through the prism of data: your proposals for testimony and feedback at numerique@amue.fr

Source: data.amue.fr

Open data and ESR: Opportunity to create new services

Contact us

[Amue-DSPSI](mailto:Amue-DSPSI@numerique.fr)

numerique@amue.fr



ALL ISSUES OF THE COLLECTION ARE AVAILABLE FOR DOWNLOADING.

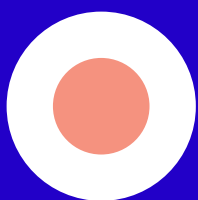
Amue. (2019). The digital collection, at <http://www.amue.fr/systeme-dinformation/metier/la-collection-numerique/r>

September 2019 (FR)

May 2020 (EN)

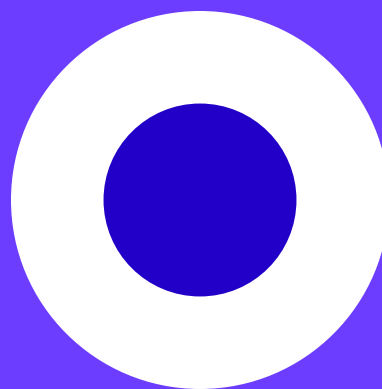


+



amue.fr

+



+

103 bd Saint-Michel + 75005 Paris
Nos réseaux sociaux : @Amue_com

