

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



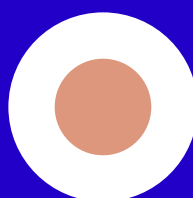
décembre
2021

L'ESR vu par saison 2 le prisme de la donnée universitaire

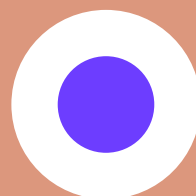


amue

MUTUALISATION + SOLUTIONS



#18



Directeur général de la
publication · Stéphane Athanase

Rédacteurs en chef ·
Bertrand Mocquet, David Rongeat

Secrétaire de rédaction · La com'

Graphisme & mise en page ·
@yay.graphisme

Photographie couverture ·
by Franki Chamaki on Unsplash

ISSN 2650-8494
La collection numérique
est sous Licence Creative
Commons CC BY-NC-SA 4.0

Ont collaboré comme auteur(e)
à ce numéro · Isabelle Blanc,
Valérie Le Strat, David Rongeat,
Bertrand Mocquet, Nathalie
Timores, Jean-Robert Petit,
Maud Lê Hung, Hervé Luga,
Yaël Salama, Fanny Lalleman,
Michel Cazaëntre, Sylvie
Haouy, Fériel Goulamhousen,
Antoine Blanchard, Romain
Pierronnet, Pauline Boyer,
Julien Gossa, Françoise Genova,
Bertrand Monthebert

Remerciements spéciaux
pour le réseautage ·
Nicolas-François Cazenave

Editeur · Amue · 103 boulevard
Saint-Michel · 75005 Paris

Fabriqué en France

Toutes les images et photos
de ce numéro sont © et libres
de droit, droits réservés
autorisation d'usage spécifique
à cette publication.

à télécharger!

tous les numéros de
la collection sont en
telechargement Amue.
la collection numérique, [ici](#) →

prochain numéro
de la collection numérique
(Février 2022):

Les Veilles. Vos propositions
de témoignage et retours
d'expériences dès maintenant
à numerique@amue.fr

Dès 2016, la France plaçait la donnée au cœur des enjeux stratégiques en promulguant la loi pour une République numérique, demandant à l'administration publique d'être plus engagée dans l'ouverture de ses données et codes sources.

En 2019, deux décrets relatifs aux échanges d'informations et de données entre administrations viennent finaliser et accélérer le principe « Dites-le-nous une fois ».

Au sein de l'ESR, chaque famille de données est multi-facettes. La donnée scientifique constitue à la fois la « matière première » indispensable aux travaux de recherche et le résultat des travaux de recherche. Pour la formation, elle représente une ressource aussi bien pour les étudiants qu'un bien pédagogique conçu par les enseignants et au service des enseignants. La donnée de gestion trop souvent enfermée dans les SI de chaque établissement devient un élément de simplification et d'efficacité des organisations quand elle circule davantage entre les SI d'un même établissement et plus largement avec d'autres. Qu'elle soit scientifique, pédagogique ou de gestion, la donnée n'a de valeur que par son usage, sans cela elle constitue un objet de valeur dormante, que seul son producteur connaît, utilise et stocke « à l'infini ».

Pour que simplification des processus, pilotage par la donnée, démocratisation de l'accès aux savoirs et aux données scientifiques se généralisent, les données doivent devenir un bien commun, en étant aussi ouvertes que possible, ou circuler dans une logique de partage maîtrisé au bénéfice de tous. Pour cela, le rapport de chacun à la donnée doit changer : chercher la donnée existante avant de la créer, décrire et mettre à disposition la donnée existante au lieu de la conserver jalousement. La circulation, l'usage et la réutilisation de la donnée constituent un cercle vertueux : sa qualité et sa description se perfectionnent, son accessibilité s'améliore, sa valeur d'usage augmente, et son empreinte environnementale diminue, par sa moindre réplique.

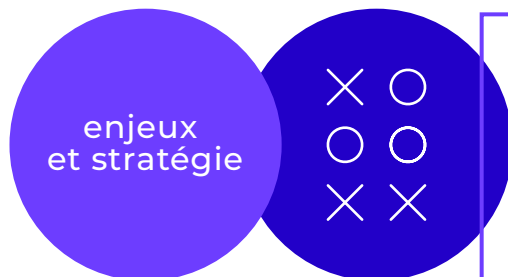
Isabelle Blanc

→ *Administratrice ministérielle des données, des algorithmes et des codes sources*

→ *Conseillère auprès de la directrice générale de la recherche et de l'innovation et de la directrice générale de l'enseignement supérieur et de l'insertion professionnelle*

→ *Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation*

@AMDAC_MESRI



auteure
Isabelle Blanc,
Administratrice ministérielle
des données, des
algorithmes et des codes
sources, Conseillère auprès
de la DGESIP, MESRI

@AMDAC_MESRI

Politique des données, des algorithmes & des codes sources

Quand le Ministère de l'Enseignement supérieur, de la recherche et de l'innovation fait de la donnée, un enjeu.

Depuis le début de la crise sanitaire, plus que jamais, la donnée s'est trouvée au cœur de la décision et du pilotage de nos actions et a montré l'importance, à l'échelle planétaire, de l'ouverture des connaissances scientifiques. La structuration, la circulation et l'ouverture des données, initiées depuis plusieurs années doit désormais se généraliser à tous les types de données, qu'elles soient d'enseignement, de recherche, de gestion ou de pilotage pour permettre un gain de valeur scientifique, économique et d'efficacité collective. Ce travail de fond mobilise tous les métiers de notre écosystème : administratifs, informaticiens, bibliothécaires et documentalistes, juristes, techniciens et ingénieurs en soutien à la recherche, enseignants-chercheurs et chercheurs... et également les étudiants que nous devons former à ces enjeux.

Innovation, confiance et simplification sont les 3 enjeux que sert la politique des données, des algorithmes et des codes sources du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation qui se déploie en 53 actions concrètes dans une feuille de route 2021-2024. Ces actions sont et seront rendues possibles grâce à l'ensemble des expertises et des initiatives déjà existantes au sein du ministère et chez ses opérateurs. Ainsi, la feuille de route constitue le cadre de travail commun à tout notre écosystème de l'enseignement supérieur et de la recherche.

Adaptative en fonction du contexte et des opportunités, cette feuille de route rassemble, capitalise, promeut, amplifie et mutualise les initiatives répondant aux attentes des usagers. En ce sens, elle soutient des initiatives de mutualisation et de transformation dans les domaines des ressources éducatives libres, de la production et de l'usage de logiciels libres, de l'ouverture ou du partage des données de recherche, de l'open-data en faveur de la transparence de nos politiques ou encore de la simplification administrative pour les étudiants, les enseignants-chercheurs ou les chercheurs.

Les actions visent à améliorer la gestion, amplifier l'ouverture par défaut (ou le partage en accès restreint si nécessaire) et la réutilisation des données de la recherche, de formation,

de gestion et de pilotage. Il s'agit de développer et mutualiser les services visant à garantir l'accès et l'exploitation de ces données par tous et à renforcer notre souveraineté. Les seules plateformes technologiques ne suffisent pas et ne suffiront pas à transformer les pratiques pour que la donnée devienne le bien de tous, produite par tous. Faire de la donnée un bien commun est un gage d'efficacité pour l'ensemble de nos organisations et de satisfaction pour tous les acteurs qui y contribuent (producteurs des données, intendant en charge de leur mise en qualité et utilisateurs des données en charge de leur analyse et leur valorisation). Mais cela ne sera possible qu'au travers de la confiance que chacun placera dans les données produites par d'autres, et nécessite plus de sensibilisations et de formations initiales et tout au long de la vie aux métiers de la préparation, description, mise en qualité, analyse des données ainsi qu'une meilleure reconnaissance de ces compétences.

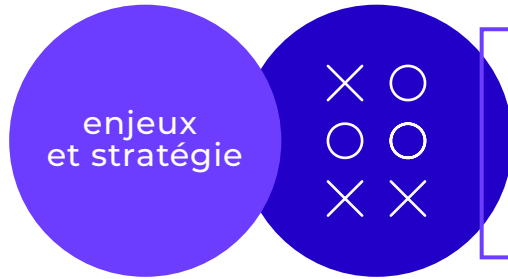
Les principes directeurs : mutualisation et souveraineté des outils, confiance dans les données, reconnaissance des producteurs de données, usages et réutilisation dans une approche écoresponsable des données comme des outils et simplification sont partagés par les initiatives comme Canal U, Callisto, Fun ressources, esup-portail, Recherche Data Gouv, appelsprojetsrecherche.fr, Software Heritage et le développement d'API.

Cette feuille de route est aussi un engagement pour plus de transparence sur les impacts de nos politiques d'enseignement et de recherche ou la mobilisation des ressources publiques. Cet engagement se concrétise par l'ouverture de jeux de données sur les moyens et les résultats de nos actions sur #dataESR et data.gouv, scanR et le baromètre de la science ouverte.



Le détail de ces actions
et de nouvelles à découvrir
sur : www.esr.gouv.fr/politique-donnee/

Pour aller plus loin
Pour comprendre la dynamique, sans précédent, autour des données, algorithmes et codes sources, nous vous invitons à lire cette page de synthèse. Vous y découvrirez les grands axes et actions majeures ainsi que les feuilles de routes de 15 ministères.

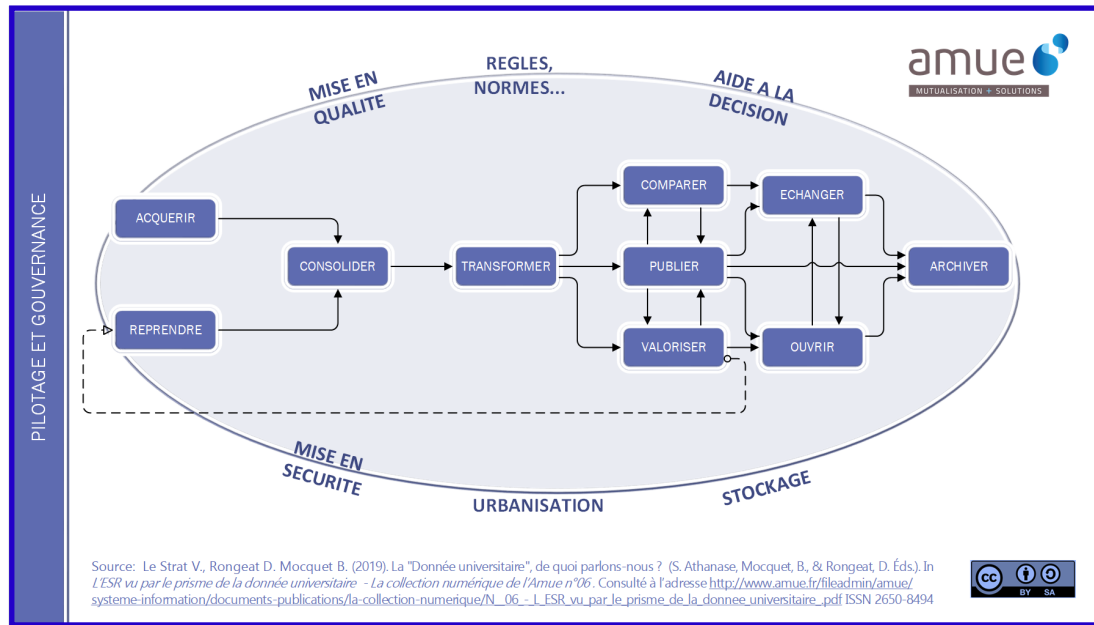


auteurs
Valérie Le Strat, urbaniste SI,
David Rongeat, responsable numérique et **Bertrand Mocquet**, expert numérique, Urbanisation et Numérique à l'Amue

La « Donnée universitaire », de quoi parlons-nous ?

PERCEPTION DU CYCLE DE VIE DE LA DONNÉE

Nous avons proposé de représenter le sujet de la donnée par le prisme de son cycle de vie lors de la collection numérique N°06 - L'ESR vu par le prisme de la donnée universitaire en novembre 2019. De son acquisition à son archivage définitif, le schéma, que nous avons proposé, positionnait différentes activités autour de la donnée (ex « transformer ») et différents sujets associés (ex « urbanisation »).



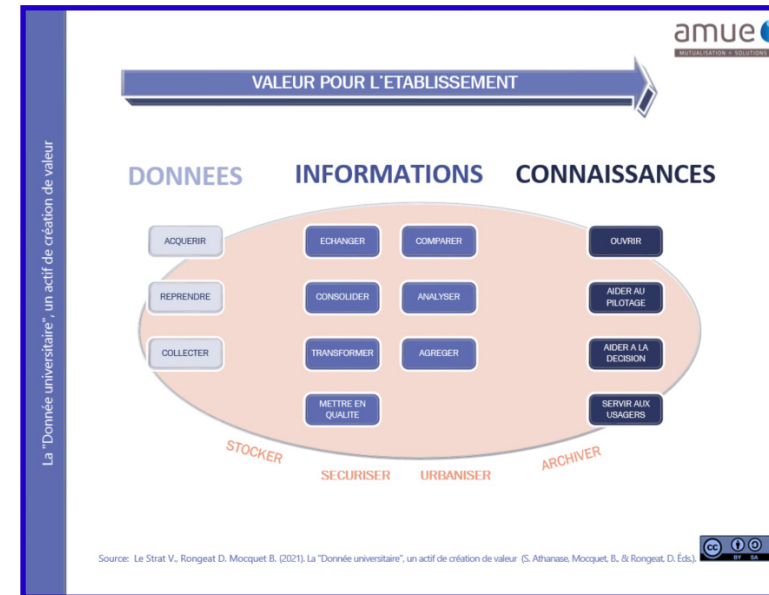
Dans ce cycle de vie, nous reconnaissons que notre intention était bien de permettre de disposer d'une représentation (vision) macroscopique de ce cycle, dans une optique de gestion de la donnée, d'aide au pilotage et de gouvernance de l'organisation qu'est l'établissement.



QUELLE PLUS-VALUE CONCERNANT LA DONNÉE POUR L'ÉTABLISSEMENT ?

Dorénavant principalement numérique, la Donnée s'appuie sur des données de l'établissement qui existent sous plusieurs formes (dactylographiées, manuscrites, non écrites, ...). Elles ne sont pas exclusivement des données issues des Systèmes d'Information (SI) mais sont également des documents, des codes logiciels, des informations non structurées, ...

La Donnée peut être présentée comme un actif de création de valeur pour l'établissement, première étape de construction d'une information puis d'une connaissance. On comprend bien que ce dernier point résonne particulièrement dans le contexte de l'Enseignement Supérieur et de la Recherche. S'en saisir d'un point de vue stratégique c'est accepter ou faire en sorte de se donner les moyens pour accepter précisément la capacité d'une donnée à devenir une information (lorsqu'elle est structurée), information qui produira de la connaissance lorsqu'elle sera analysée et visualisable. Une donnée brute produit moins de valeur, contrairement à une connaissance.

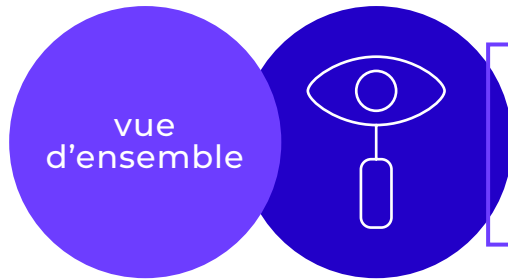


Chaque verbe du schéma précédent est associé à son étape correspondante (donnée, information, connaissance) et permet de matérialiser comment on passe de l'une à l'autre. La première étape de ce cycle de valeur réside dans la collecte de la donnée par acquisition ou une reprise. La donnée est ensuite consolidée et subit une succession d'étapes qui en feront une information. Il s'agit ici de proposer concomitamment des comparaisons, des publications de la valorisation dans une optique de diffusion vers l'échange ou l'ouverture.

Cette maîtrise des activités alimente des sujets de l'organisation publique du Supérieur, telle que le stockage, la mise en qualité, la conformité à des règles ou normes existantes, la mise en sécurité d'un point de vue de la sécurité du système d'information. Et c'est aussi des sujets plus complexes comme l'urbanisation et l'aide à la décision, deux sujets pour lesquels une mauvaise interprétation d'un jeu de données pourrait apporter une mauvaise décision et un risque probablement augmenté pour l'établissement.

Quel que soit le stade de maturité de la donnée, il faudra la stocker, la sécuriser, l'urbaniser mais également l'archiver.





auteure
Nathalie Timores, chargée de domaine patrimoine immobilier et développement durable, Amue

Les enjeux de la DATA immobilière en ESR

Avec les schémas pluriannuels de stratégie immobilière et la dévolution du patrimoine, le pilotage de l'immobilier par la donnée devient un enjeu majeur de la stratégie globale des établissements. On en découvre tous les aspects.

ENJEUX DE LA DATA IMMOBILIÈRE EN ESR

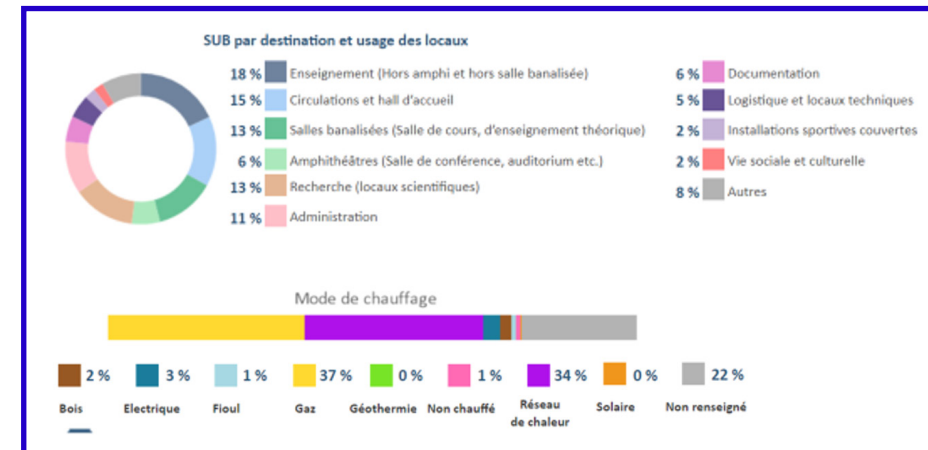
Le 2 décembre dernier, l'Amue a présenté son nouvel accord-cadre sur le **Système d'Information du patrimoine Immobilier (SIPI)** devant 159 participants. Cette offre de service est l'aboutissement d'un processus décisionnel et d'une co-construction avec les établissements le MESRI et l'association Artîes entre 2019 et 2021. Elle fait naître une nouvelle communauté métier et de nouveaux besoins en termes d'accompagnement. Lors du sondage en ligne durant la conférence, sur 89 votants, 54% ont déclaré ne pas être équipés d'un SI immobilier et 71% ont déclaré vouloir s'équiper entre 2022 et 2025.

Rappelons que l'auto-évaluation HCERES et l'évaluation socio-économique des projets immobiliers s'appuient sur les données immobilières. Les processus de labélisations dans lesquelles les établissements s'engagent demandent eux aussi une évaluation par la donnée notamment en matière de DDRS. <https://www.label-ddrs.org/>

La dynamique des données immobilières a été mise en perspective par les travaux conduits par le MESRI sur le cadre de cohérence du système d'information de l'immobilier en ESR. Cette dynamique répond aux enjeux du secteur de l'immobilier qui irrigue l'ensemble de l'ESR de par la multiplicité des acteurs qu'il interroge aussi bien en interne qu'en externe à travers ses processus métier. <https://services.dgesip.fr/T844/S372/immobilier>

QU'EST-CE QU'UNE DATA IMMOBILIÈRE ?

La data, au sens large, représente tout ce qui concerne les données, qu'il s'agisse de leur collecte, de leur interprétation ou de leur transformation.



Assurer la fiabilisation des données immobilières correspond à la capacité de garantir l'exhaustivité, la précision, l'exactitude et la validité des données durant tout leur cycle de vie. La notion de cycle de vie de la donnée immobilière est en lien avec une temporalité et l'interopérabilité qu'elle induit en interne à l'établissement, avec les infocentres ministériels, les plateformes interministérielles et les prestataires et fournisseurs extérieurs à l'établissement.

Dès lors la donnée immobilière couvre un périmètre large, inscrit dans une temporalité qui va de sa gestion opérationnelle à sa gestion stratégique. Ceci implique une organisation métier interne au service immobilier, une montée en compétence des équipes et une approche transversale. La donnée immobilière devient l'expression de la connaissance du parc immobilier et de son évolution. Elle rend compte des usages des sites et des bâtiments. Mise en perspective, la donnée devient un outil stratégique de pilotage de programmation et de négociation.

DATA DESIGN

La complémentarité des approches design et data est de plus en plus utilisée dans les phases de diagnostic et de définition des besoins. Data design signifie « le design des données » et design peut être défini comme le fait de placer l'utilisateur au cœur de la donnée.

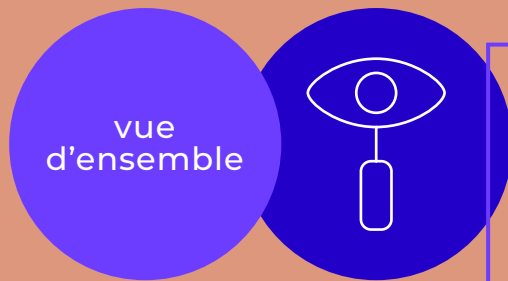
Le « DATA design » permet de fournir les éléments-clés aux usagers et aux décideurs publics par l'appropriation des données utilisées.

Dans le cadre de la mise en œuvre du décret tertiaire par exemple, pouvoir accompagner les utilisateurs qui sont les acteurs de la transformation, implique de leur donner une visibilité de la transformation de l'usage et de la performance de l'immobilier. Il y a alors nécessité de communiquer les données aux usagers pour qu'ils puissent devenir acteurs de la transformation.

Le smart building de l'usage aux usagers

Au départ, le smart building ou le bâtiment connecté adresse le confort d'usage, puis la gestion de l'énergie et maintenant le pilotage c'est-à-dire les données prédictives. Se projeter en termes de gestion c'est le smart, c'est l'évolution de la data. A l'heure actuelle l'entrée de lecture de la donnée immobilière se concentre sur l'usage; connaître les flux dans un bâtiment, l'occupation des locaux et les consommations d'énergie ce qui concerne directement les usagers. Le cycle de vie des données immobilières, la DATA, est un vrai vecteur de transformation des modes de travail, des organisations plus largement des modes de vie

Modéliser et analyser les données immobilières pour communiquer, c'est réintroduire de l'intelligence dans la lecture de la performance d'usage ce qui peut être un argument de communication et de négociation envers les décideurs.



auteur **Jean-Robert Petit**, chargé de domaine pilotage, département développement et accompagnement des compétences, Amue

Intendant de données, c'est nouveau !

Gérer les données produites quotidiennement constitue une condition indispensable de leur qualité. Se dessinent alors les contours d'un nouveau métier dans l'ESRI, puis d'une communauté... on fait la lumière sur l'intendant de données

Le métier d'intendant de donnée souvent appelé data steward dans le secteur privé est un métier nouveau dans l'ESRI.

Il voit le jour avec le déploiement de stratégies d'intégration et de « désilotage » du système d'information et la diffusion de nouveaux outils comme le référentiel de données SINAPS de l'Amue.

LE PORTRAIT DE L'INTENDANT

L'intendant est un acteur important dans la gouvernance de la donnée puisqu'il a la charge de :

- Gérer ou superviser la création, la modification et la suppression de données de références
- Traiter les anomalies et réaliser les dédoublemnages en liaison avec les services
- Concourir à l'optimisation des processus d'acquisition et de circulation des données
- Documenter, former et préparer les reportings.

Il est par ailleurs le correspondant fonctionnel de l'outil qui dans la majorité des cas est SINAPS.

A la marge peuvent être utilisés d'autres applicatifs soit de type MDM soit « faisant fonction » de référentiel des données comme le LDAP ou un logiciel de gestion des services ; enfin une dernière situation, les collègues sans outil.

L'intendant n'a pas ou peu suivi de formation à la gouvernance de la donnée à l'exception des séquences proposées par l'Amue dans le cadre du déploiement de SINAPS.



Un métier ou un rôle ?

Le métier d'intendant n'apparaît pas dans le référentiel des emplois-types de l'ESRI.

Par ailleurs, rares sont les collègues en établissement ayant le titre d'intendant de données et tout aussi rares sont ceux qui assurent les fonctions d'intendance à 100 %.

Une enquête réalisée auprès d'une quinzaine d'entre eux il y a un an montrait que seule une moitié passait plus de 50 % de son temps de travail sur des tâches de gestion de la qualité de la donnée. La quotité restante relève souvent du contrôle de gestion.

Intendant et contrôleur de gestion

Ce lien avec le contrôle de gestion est fort. L'intendant est souvent un contrôleur de gestion à qui l'on a greffé des fonctions d'intendance de la qualité de la donnée. Il est généralement rattaché au service de l'aide au pilotage de l'établissement.

UNE NOUVELLE COMMUNAUTÉ À ACCOMPAGNER

La communauté des intendants de données est récente et de taille encore limitée.

Il est hasardeux d'en évaluer le nombre dans l'ESRI, au plus peut-on relever qu'environ 40 collègues ont participé au moins une fois au « café des intendants de données » de l'Amue qui constitue aujourd'hui le principal outil de structuration de la communauté et dont les travaux nourrissent cet article.

Mensuel ou bimestriel, chaque « café » réunit (en distanciel) environ 15 collègues pendant une heure autour de sujets organisationnels, opérationnels, techniques, définis ensemble. Il est l'occasion de se connaître, d'échanger, d'apprendre des expériences des autres et de capitaliser à travers un *padlet*.

L'accompagnement prend aussi la forme d'une conférence annuelle sur le pilotage de la donnée au printemps, d'une liste d'échanges (à venir) ou encore d'un Lab'U, le prochain se tenant le 24 janvier en présentiel à Paris (inscription obligatoire)



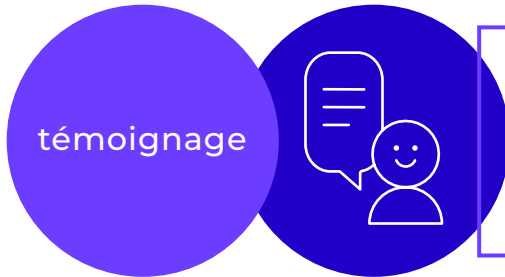
Introduire la gouvernance des données pour des données de qualité

La mise en œuvre de Sinaps est l'opportunité d'une gouvernance des données dans votre établissement : Un levier pour valoriser l'information collectée et exploitée par les métiers <https://www.amue.fr/pilotage/logiciels/sinaps/presentation/gouvernance-des-donnees/>

Ces actions d'accompagnement concourent à la professionnalisation des collègues parfois insuffisamment connus et reconnus dans nos établissements et faciliteront nous le pensons l'installation et la légitimation de cette fonction qui participe de la gouvernance de la donnée et constitue une condition de réussite de toute transformation numérique.



The screenshot shows a web interface for an event titled "Lab'U #3 Contrôle de gestion et gouvernance de la donnée". It includes a "Catalogue des événements" menu with options like "Accueil", "Tous les événements", "Liste de vos inscriptions", and "Conditions d'inscription". The main content area displays the event date "24/01/2022", the organizer "Lab U" with email "hasna.laaguid@amue.fr", and a "Pilotage" tag. A "STATUT" section shows "ACTIF". A "Détails de la session" sidebar lists: "statut ACTIF", "DATE LIMITE D'INSCRIPTION 20/01/2022", "dates Du : 24/01/2022 9h30 Au : 24/01/2022 16h45", "durée totale 1 JOUR(S)", and "lieu AMUE - 103 boulevard Saint-michel 75005 PARIS". At the bottom, there are tabs for "OBJECTIFS", "PROGRAMME", and "PUBLICS CIBLES", with a list of roles: "Contrôleurs de gestion, DGSA ou directeurs du pilotage, intendants de données, auditeurs". A "s'inscrire" button is visible at the bottom right.



auteure
Isabelle Blanc, administratrice ministérielle des données, des algorithmes et des codes sources, conseillère auprès de la DGEIP et de la DGRI

Ouvrir des données de la recherche – Recherche Data Gouv

C'est au printemps 2022 que la plateforme nationale fédérée des données de la recherche ouvrira. On s'intéresse à ses futurs services.

Pour répondre aux enjeux d'ouverture des données de la recherche, la stratégie est de s'appuyer sur les forces françaises des infrastructures de recherche disciplinaires (Huma-Num-Nakala, Data Terra, PNDB, CDS, France Cohortes...) et celles fortement inscrites à l'échelle internationale (GBIF, EBI, GEO)... et de les compléter par une offre pluridisciplinaire.

Recherche Data Gouv proposera ainsi, dès le printemps 2022, une offre souveraine de publication et de signalement des données de recherche ainsi qu'une offre d'accompagnement aux équipes de recherche sur toute question relative à la donnée.

Il s'agit également d'inscrire la France dans un paysage international en évolution en contribuant aux services de l'EOSC (European Open Science Cloud) et en rendant visible les données de la recherche française.



Recherche Data Gouv: 5 modules d'accompagnement et de services pour favoriser l'ouverture des données de recherche

Trois modules complémentaires pour accompagner les équipes de recherche

→ **Ateliers de la donnée** : en proximité des équipes de recherche, ils apportent un premier niveau d'expertise et des services généralistes en matière de gestion, préparation et diffusion des données ;

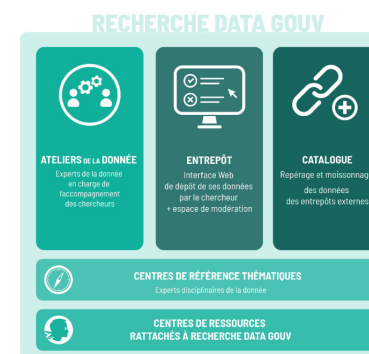
→ **Centres de référence thématiques** : spécialisés par domaines thématiques et adossés aux infrastructures disciplinaires, ils interviennent en soutien aux ateliers de la donnée auxquels ils apportent une expertise scientifique au travers des référentiels de pratiques disciplinaires (vocabulaires de description des données, période d'embargo...) en cohérence avec l'écosystème international ;

→ **Centres de ressources** : assurent le support des modules entrepôt et catalogue des données et mettent en commun les pratiques, guides, formations déployées par les ateliers de la donnée.

Deux modules pour la publication et la découverte des données de la recherche française

→ **Entrepôt de données** : un service de dépôt et de publication des données de la recherche pour lesquelles aucun des entrepôts disciplinaires existants ne constituerait une solution adaptée ;

→ **Catalogue pluridisciplinaire des données de la recherche française** : il fédère les données publiées ou signalées dans Recherche Data Gouv. Les données publiées dans des entrepôts nationaux ou internationaux thématiques et disciplinaires seront repérées et leurs métadonnées de description moissonnées par Recherche Data Gouv.



DES SERVICES DISPONIBLES DÈS LE PRINTEMPS 2022

La phase projet durera jusqu'en 2023, mais dès le printemps 2022 Recherche Data Gouv permettra de :

- Publier ou signaler des données de recherche grâce au module « entrepôt » ;
- Consulter les données ou les méta-données décrites dans Recherche Data Gouv
- Trouver un entrepôt thématique reconnu, de l'écosystème français
- Identifier les premiers ateliers de la donnée labellisés offrant un accompagnement sur toute question relative aux données

Après ce premier lancement, la labellisation des ateliers de la donnée sur tout le territoire se poursuivra et le module catalogue des données permettant de signaler les jeux de données présents dans des entrepôts externes nationaux et internationaux thématiques sera conçu et déployé.

RECHERCHE DATA GOUV DÉVELOPPÉ PAR ET POUR LA COMMUNAUTÉ DE RECHERCHE

Les services d'accompagnement se développent actuellement sur tout le territoire grâce aux initiatives déjà existantes dans les établissements. Ces derniers sont appelés à mettre en commun leurs compétences et leurs expertises afin d'offrir une offre visible et structurée à tous les laboratoires de leur territoire.

Le développement des modules entrepôt et catalogue a été confié à INRAE, l'acteur français ayant le plus d'expérience dans le domaine des entrepôts pluridisciplinaires. Sept établissements ayant également une expertise en terme de publication ou d'accompagnement à l'ouverture des données de la recherche ont rejoint INRAE pour ce projet.



Le projet est entièrement financé par le Fonds national pour la science ouverte de 2021 à 2023 et est piloté par le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.



auteurs

Maud Lê Hung (directrice DRSU), **Hervé Luga** (VP Numérique), **Yaël Salama** (coordinatrice DRSU), **Fanny Lalleman** (responsable Observatoire), **Michel Cazaëntre** (responsable SID, Service Numérique), Université fédérale Toulouse Midi-Pyrénées

DataVUE, une plateforme d'aide au pilotage pour les Villes Universitaires d'Équilibre

Quand les établissements portés par l'université fédérale Toulouse Midi-Pyrénées mutualisent leurs outils et compétences, les indicateurs s'envolent vers la performance et l'analyse nationale. On vous dévoile le projet.

La répartition territoriale de l'académie de Toulouse présente une double originalité : une métropolisation très forte (86% des étudiants sont à Toulouse) et un maillage territorial important constitué de dix Villes Universitaires d'Équilibre (VUE), villes de taille moyenne ayant des implantations universitaires. Le développement de ces VUE fait partie des axes stratégiques de l'Université Fédérale Toulouse Midi-Pyrénées (UFT). Le projet DataVUE a été initié pour cartographier l'activité de ces villes afin d'en améliorer leur structuration et leur pilotage. Il fournit des indicateurs territoriaux selon de multiples axes d'analyse sur l'ensemble des domaines ESRI des VUE : constitution et structuration ESRI des VUE, offre



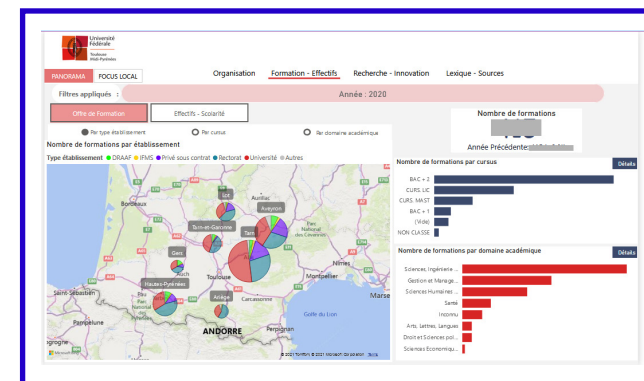
de formation et effectifs, vie étudiante, recherche et innovation, impacts économiques, etc. Avec un accès réservé, la plate-forme s'adresse aux établissements de l'UFT comme à ses partenaires (Collectivités territoriales, EPLE, etc) aux attentes qui peuvent être différentes. Afin de rester au plus près des besoins de ses utilisateurs finaux, ce projet est réalisé en co-construction avec les acteurs métiers concernés et piloté avec les partenaires du projet (Rectorat de région académique, Région Occitanie, DRAAF). Le projet a bénéficié d'un fond d'amorçage de la Région Occitanie et d'un soutien pluriannuel du MESRI dans le cadre du Dialogue Stratégique de Gestion.

DataVUE est alimenté, automatiquement pour l'essentiel, à partir des systèmes d'informations des établissements de l'UFT (notamment l'offre de formation et le guide de la recherche), de la DRAAF, du Rectorat et de l'INSEE. Les données collectées localement sont projetées dans des référentiels nationaux, structurant les périmètres d'analyse et permettant des comparaisons inter et intra-régionales. Des données collectées manuellement sont ajoutées pour compléter la présentation des VUE (équipe d'animation, éléments d'activité, etc.).

Depuis 2007 et la création des projets d'Université Numérique Régionale, l'UFT joue un rôle majeur en matière de développement de projets mutualisés.

D'abord essentiellement orientés services (environnement collaboratif, carte, portail d'offre de formation) et basés sur un annuaire commun, ces projets se sont ensuite diversifiés permettant la création d'un observatoire germe du projet DataVUE et structurés autour d'un cloud à haute disponibilité opéré de manière mutualisée. En accompagnement de l'hybridation se sont ainsi un service de distribution de vidéos et maintenant d'accès distant qui ont vu le jour. D'un point de vue structurel le site porte pour la partie Ouest le développement du projet de Datacenter mutualisé DROCC et développe avec ses partenaires établissements et organismes l'offre de services de son centre opérationnel Ouest.

Le soutien à la recherche, que cela soit autour du calcul, de la donnée, ou des services de science ouverte est un aspect important des projets communs impliquant à la fois les acteurs reconnus comme le MesoCentre CALMIP (UAR CNRS 3667) ou la plateforme de bioinformatique GENOTOUL et plus largement l'ensemble des laboratoires et personnels du site.



DataVUE est développé sur la plateforme de dataviz Microsoft PowerBI, plateforme également utilisée par les établissements de l'UFT pour construire leur Système d'Information Décisionnel. Cela permet un partage de compétences et une mutualisation des investissements en formation. DataVUE fournit ainsi des graphiques, histogrammes et cartes interactifs. L'utilisateur peut alors de manière dynamique affiner ou élargir son périmètre d'analyse en interagissant avec les filtres proposés ou en sélectionnant une partie d'un histogramme. Il est également possible d'explorer les indicateurs à différents niveaux, hiérarchisés, allant du panorama régional jusqu'à des niveaux très détaillés (par exemple : par VUE, par secteur d'activité, etc). De plus, afin de maximiser l'adéquation de l'IHM de DataVUE, un UX designer a accompagné l'équipe projet pour la réalisation de l'interface.

DataVUE est un outil qui existe grâce à la mutualisation et la coordination entre les établissements portés par l'UFT: la connaissance des VUE du Département Réseau des Sites Universitaires, les données de l'Observatoire, les compétences de montage et de gestion de projet numérique du Service Numérique, les compétences d'analyse de données. Ce projet s'appuie également sur des systèmes d'information déjà mis en place par l'UFT, également issus d'une politique d'animation du site. Les enjeux sont à présent d'étendre les domaines et d'enrichir les indicateurs déjà présents dans DataVUE afin de contribuer à l'élaboration de projets communs et conforter l'ESR dans les territoires.



auteure
Sylvie Haouy,
architecte
d'entreprise,
Université Côte
d'Azur

Vers une Gouvernance de la donnée universitaire

A l'université Côte d'Azur le dénominateur commun de la stratégie établissement et de la gouvernance transverse, c'est la donnée.

Dans les établissements de l'ESRI, structures et projets autour de la donnée se multiplient à tout niveau que ce soit pour la scolarité, la pédagogie, l'administration, la recherche ou encore le pilotage. A l'université Côte d'Azur, ce foisonnement d'usages et la volonté générale d'un pilotage par les données nous conduisent à la mise en place d'une stratégie établissement de la donnée et d'une gouvernance transverse pour assurer sa mise en œuvre.

UNE DÉMARCHÉ D'ARCHITECTURE D'ÉTABLISSEMENT

L'architecture d'établissement, permet de travailler avec les métiers, la direction générale et la Présidence sur des trajectoires et des cibles de transformation numérique. L'objectif d'Université Côte d'Azur est de devenir un établissement avec un pilotage par la donnée :

→ La première étape de notre trajectoire a été de mettre en place un outil de Master Data Management pour la mise sous contrôle des données de référence (aujourd'hui structures, individus personnels et apprenants) avec un travail de (ré)écriture des différents processus métiers impliqués. D'autres ensembles de données de référence vont suivre.

→ La seconde étape, actuellement en cours, consiste à formaliser une gouvernance de la donnée. Les livrables sont :

- Un cadre - ensemble de principes à respecter - permettant une diffusion de la culture de la donnée dans tout l'établissement ;
- Un modèle de dictionnaire des données indiquant à chaque métier comment décrire ses objets métiers et ses données ;
- Et des processus de gouvernance pour le domaine du pilotage à décliner et appliquer dans chaque autre domaine métier du SI.

→ La troisième étape est l'appropriation de ces processus de gestion de la donnée par chaque responsable de domaine fonctionnel visera à l'établissement des procédures applicables pour assurer la garantie d'une donnée complète, conforme, sécurisée et fiable, conditions nécessaires à sa valorisation.

DES ACTEURS DATA DE DIFFÉRENTS MÉTIERS

Un travail collaboratif entre l'architecte d'établissement et trois types d'acteurs est en cours pour parvenir à une gouvernance et une animation de la communauté data.

L'**intendant de données**, avec sa vision transversale de l'université, rend les données partageables entre les SI et les transforme en données de référence. Son talent est de trouver un équilibre entre la qualité et l'opérationnel, c'est à dire ne pas être intransigent sur les saisies tout en ayant un réel souci de la qualité des données. Il travaille quotidiennement à créer du lien entre les différents propriétaires et usagers des données et à faire évoluer leurs pratiques en accord avec la Gouvernance des données.

Le **data officer**, avec sa forte expertise en analyse des données, fournit des outils d'aide à la décision pour les approches transversales ou croisées de plusieurs domaines et ce toujours dans un souci de contextualisation. Son talent est de chercher l'harmonisation des méthodes utilisées par les équipes de pilotage des composantes pour arriver à une même valeur de vérité au sein de l'établissement.

Le **data Scientist**, « fouineur » de données sur un périmètre plus ou moins restreint, vise à créer de nouvelles données souvent de type comportementales (learning analytics) dans le respect du contexte RGPD. Son activité est assez proche du data officer mais avec un aspect exploratoire et son talent est la créativité.

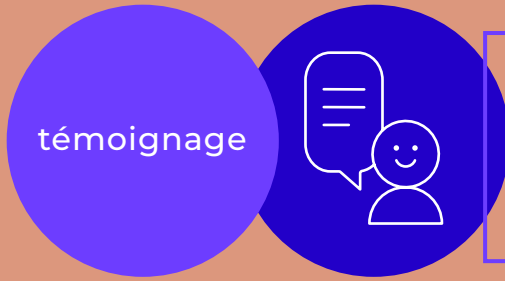
UN CHANGEMENT DE PARADIGME

La mise en place d'une gouvernance de la donnée universitaire est un travail collaboratif de cadrage et d'outillage mais aussi de sensibilisation et d'acculturation. Nous appuyer sur une démarche méthodologique éprouvée permettra d'installer cette responsabilité de la donnée dans chaque macro processus métier, support et pilotage.



Vu aux assises du CSIESR
Echanges data-centric : table ronde autour de la donnée, comment elle est perçue et traitée à l' Université de Côte d'Azur
Replay bientôt disponible [ici](#) →





témoignage

auteur
Fériel Goulamhousen,
directeur général délégué
de l'Université Gustave Eiffel en
charge du Centre d'Innovation
Pédagogique et Numérique

Gustavizz,
la datavisualisation
de l'offre de formation
par les compétences

Parce que l'innovation est au cœur
des savoir-faire de l'université Gustave
Eiffel et que pour une vue d'ensemble
on n'a pas trouvé mieux que le graphique,
un outil rend intelligible l'offre de
formation, pour informer toujours mieux.

Avec le projet D.Clic (ANR Nouveaux Coursus à l'Université NCU), ayant notamment pour objet
de déployer l'approche compétences dans les licences de l'Université Gustave Eiffel, le Centre
d'Innovation Pédagogique et Numérique (CIPEN), a eu l'idée de développer une application
permettant de visualiser l'offre de formation de manière graphique.

L'idée a germé dans le bureau des chefs de projets. Après avoir réalisé l'outil de gestion de
l'affichage de l'offre de formation de l'Université Gustave Eiffel, Ludovic Collin, Lou Hamonic
et Sébastien Tersou ont eu l'idée de visualiser les liens entre les compétences, les UE et les
matières, formation par formation, dans un premier temps, simplement par curiosité.

En effet, le CIPEN dispose de la maîtrise de l'affichage de l'offre de formation, des liens entre
les formations, blocs de compétences, compétences, UE, matières etc. : des relations n-m ou
relations Many-to-many dans tous les sens !

Cette visualisation, assez complexe dans un premier temps, a fait réaliser à l'équipe du CIPEN
qu'à la fois les formations ont organisé leurs compétences de manière différentes, avec des intitulés
variables mais également que des matières et des cours mobilisés dans les compétences
ont une lisibilité limitée.

UN OUTIL AU SERVICE DE LA FORMATION

Grâce à Claudie Meyer, enseignante-chercheuse au laboratoire DICEN-IDF et responsable du
département Management de l'Innovation et Ingénierie des Services de l'IFIS, le CIPEN a pu
tester grandeur nature l'outil Gustavizz dans le but de rendre le plus lisible possible les formations
des L2 et Master MITIC et MIPI : mettre en lien les compétences, les matières, les années
et les situations de travail en entreprise.

L'outil a été testé avec les professionnels des formations (en apprentissage), les enseignant.
es, les chargés d'enseignement vacataires et les étudiant.es : cela améliore la construction du
projet professionnel des étudiants et les situations de travail en entreprise.



Cela permet « aux entreprises de savoir si elles confient une mission
en adéquation avec la formation ».

LE PILOTAGE DES FORMATIONS
ET DES COMPÉTENCES PAR LA DATAVIZ

L'intérêt de l'outil est également de voir rapidement si certaines
formations, qui souhaitent entrer dans la démarche compétences, ont
déjà entamé le processus : identification des intitulés des compétences
par des verbes d'action, regroupement de matières, modalités
d'évaluation etc.

De même, d'un point de vue macro, il est également possible d'identifier
si les composantes disposent d'un continuum des compétences
dans leurs formations (en L et en M) via les regroupements graphiques
qui sont rendus possible grâce à la datavisualisation.

L'AMÉLIORATION DE LA COMMUNICATION
ET DES INTITULÉS DES COMPÉTENCES

Maintenant, il est possible d'imaginer améliorer la communication
sur les formations, à la fois auprès des étudiants, des entreprises, des
recruteurs et même pour les futur.es enseignant.e.s/chercheur.e.s qui
souhaiteraient s'investir au sein de l'université.

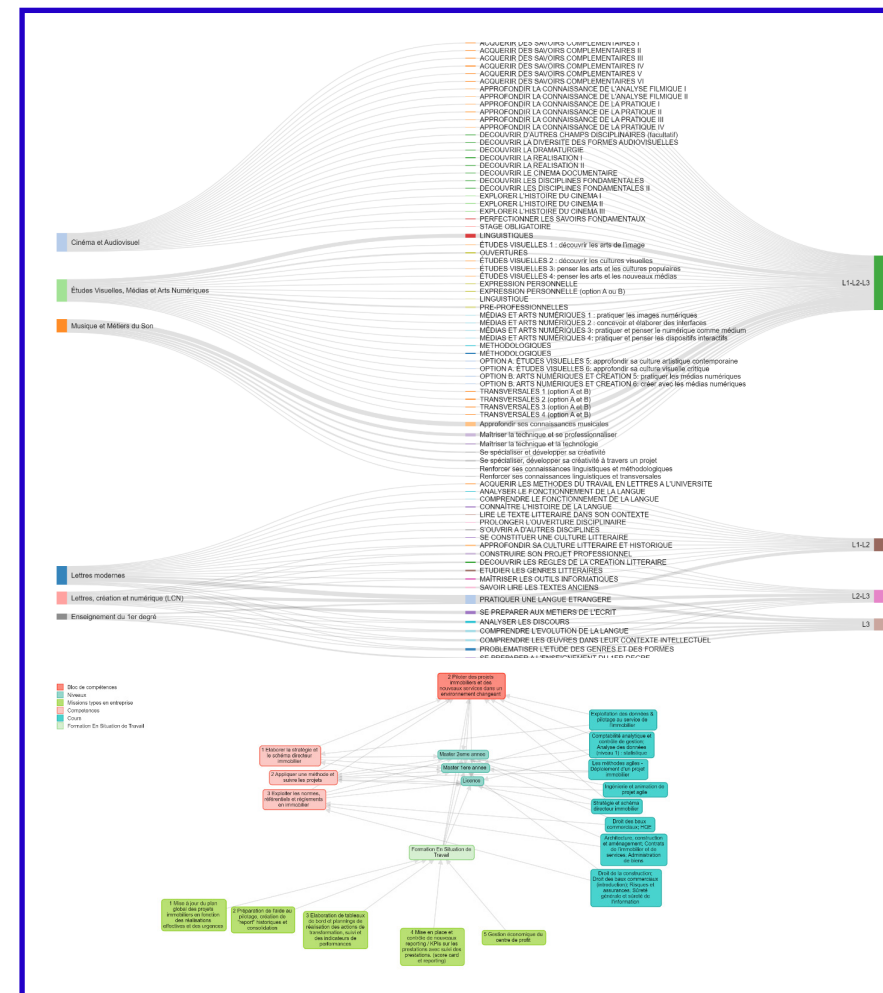
Sans évoquer la question de l'uniformisation des compétences, il apparaît
également que certaines compétences reviennent plus régulièrement

ment, par exemple les compétences autour
de la communication, de la littératie numérique,
encore des langues ou de l'expression écrite.
Il est possible aussi d'envisager
et de proposer des dispositifs pédagogiques
mutualisés afin de pouvoir faire en sorte
de ne pas multiplier les efforts dans toutes
les formations et les composantes,
et c'est le rôle du CIPEN de faire cela.

LA PROCHAINE ÉTAPE ?

Il s'agit maintenant de pouvoir améliorer l'outil,
du point de vue du design et de la navigation.
Si toutes les composantes et les formations
sont représentées, elles ne le sont pas toutes
au même niveau (de celle des formations
gérées par notre collègue Claudie Meyer).

De même, il sera nécessaire de pouvoir expliquer
le sens des données visualisées car vu de
manière brute, de multiples interprétations
sont possible. Elles seront à contextualiser,
discipline par discipline, formation par
formation, composante par composante.



https://
cipen.univ-
gustave-
eiffel.fr/
dataviz/#/

Projet D.Clic :
accompagnement
à la mise en œuvre
d'une approche
par compétences :
https://dclic.univ-
gustave-eiffel.fr/
La datavisualisation
à l'épreuve de
l'intelligibilité
de compétences :
une attention
nécessaire dès
la conception du
projet - Meyer,C.
2020 - Journée
d'études AIPU
France (5 et 6
novembre 2020) -
Pages 83-85





Quels métiers & compétences de la donnée dans l'ESRI ?

La donnée est partout jusque dans les fiches de poste, et fondamentale dans un profil. On zoom sur son positionnement dans le secteur RH de l'ESRI ?

À tous les niveaux et dans tous les métiers, la maîtrise de la donnée constitue un enjeu croissant qui est moins une affaire de spécialistes que l'affaire de tous : tel est le postulat de base d'une recherche¹ que nous avons menée en nous basant sur un double matériau : d'une part, en exploitant 14 entretiens semi-directifs réalisés au sein de deux universités issues d'un processus de fusion et, d'autre part, par l'analyse des 242 emplois-types BIATSS du référentiel RéFérens².

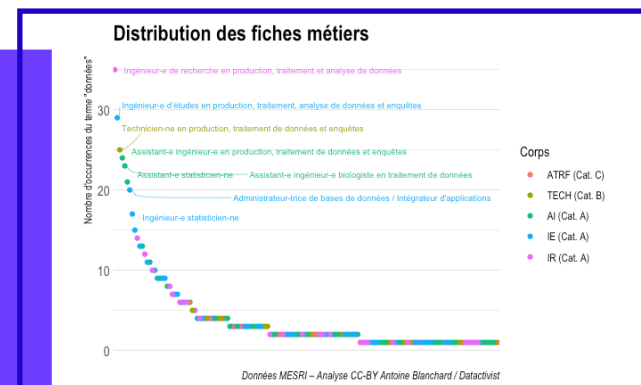
Il ressort de cette dernière que 154 fiches métiers contiennent le terme « données », une forte présence qui confirme l'importance que revêt désormais cet enjeu pour

les ingénieurs et personnels techniques des universités. En particulier, ce terme apparaît le plus dans les deux champs qui décrivent les compétences opérationnelles et les activités principales des fiches métiers : la donnée est donc une préoccupation immédiate de ces métiers. En outre, c'est le troisième concept (après « développement » et « gestion ») le plus fréquent dans le champ "Facteurs d'évolution à moyen terme" des fiches métiers.

De plus, l'étude des branches d'activités professionnelles (BAP) montre que les BAP E (Informatique, statistiques et calcul scientifique) et D (Sciences humaines et sociales) possèdent en proportion la plus grande part de métiers liés aux données. Par ailleurs, les BAP A, B et C qui concernent directement le soutien à la recherche sont moins concernées par les données que les BAP F et J qui sont davantage liées à l'administration de l'université.

1 | https://dataactivist.coop/je_docs_donnees/
 2 | https://github.com/dataactivist/je_docs_donnees

Figure – Nombre d'occurrences du terme « données » dans les emplois-types RéFérens selon les corps de la filière ITRF



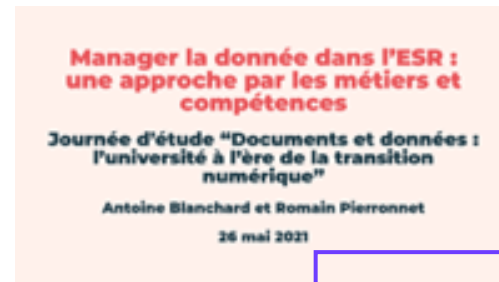
Enfin, l'étude par corps montre que la proportion de métiers comprenant le terme « données » progresse avec le niveau d'emploi (des corps ATRF, TECH et AI aux corps IE et IR).

Mais comment ces constats se traduisent-ils sur le terrain et dans les pratiques professionnelles ?

Notre recherche montre que l'échelon central des universités s'empare des enjeux relatifs à la donnée, autour d'un triptyque équilibré (mettre à disposition des outils) / former (développer les compétences) / accompagner (les initiatives et actions de terrain). Tandis que cette logique peut se heurter à des résistances locales (par exemple par crainte de nouvelles injonctions bureaucratiques peu utiles à la recherche), d'autres contextes sont plus favorables à la prise d'initiatives de la part des communautés, notamment lorsque la donnée constitue un enjeu dont la maîtrise collective est consubstantielle de l'activité de recherche (par exemple, en astrophysique où la donnée est rare et implique un partage organisé). Aussi, pour agir au plus tôt et contourner les résistances, c'est prioritairement pour les jeunes

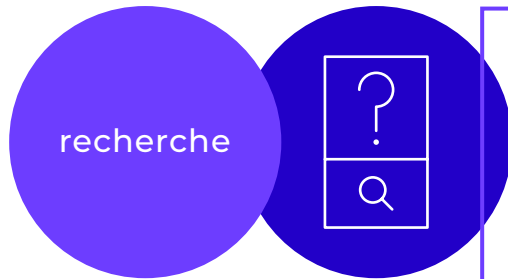
chercheurs, notamment les doctorants, que sont organisées les formations.

Alors que la donnée semble désormais très présente dans les emplois-types RéFérens, on constate sur le terrain l'inflation des data-métiers (« data scientist », « data analyst », « administrateur des données », « data steward », « data librarian »...), qui montre (au moins dans le vocabulaire) l'ampleur de la transformation. Les universités soulignent la difficulté à recruter sur ces profils, pour lesquels les pratiques de gestion des ressources humaines ne permettent pas aux établissements d'être suffisamment attractifs, notamment vis-à-vis des carrières que propose le secteur privé. L'État a justement engagé il y a quelques années une réflexion sur la politique RH de sa filière numérique (voir par exemple le rapport « Évaluation des besoins de l'État en compétences et expertises en matière de donnée », DINUM-Insee, septembre 2021). Reste à voir si et comment cette démarche nationale se prolongera à l'avenir dans les politiques des universités, à l'aune de leurs stratégies et de leurs moyens.



L'intégralité de la communication à la Journée d'étude "Documents et données : l'université à l'ère de la transition numérique" https://dataactivist.coop/je_docs_donnees/





auteurs **Pauline Boyer**, doctorante et **Julien Gossa**, maître de conférences, Laboratoires SAGE et ICube, Université de Strasbourg, CNRS¹

Évaluer les réformes ESRI, c'est possible grâce aux données

Les données ouvertes constituent une source importante d'information pour l'évaluation des politiques publiques. Mais aujourd'hui, nous avons encore besoin d'une méthodologie rigoureuse pour les exploiter.

L'adoption en 2016 de la Loi pour une République numérique concerne directement les établissements français d'enseignement supérieur et de recherche, qui contribuent à une importante production de données ouverte. On recense ainsi 184 jeux de données associés au MESRI sur la plateforme nationale data.gouv.fr et la plateforme du MESRI #dataESR.

➤ DONNÉES OUVERTES ET ACTION PUBLIQUE : POUR QUI, POUR QUOI ?

Les données ouvertes peuvent servir aux acteurs locaux pour la gestion et le pilotage stratégique de leur institution, en aidant à la situer dans le temps et dans l'espace national.

Elles peuvent également servir aux acteurs nationaux pour observer la transformation du système dans sa globalité : regroupements (PRES, COMUE, EPE, ...), élargissement des compétences (RCE), différenciations, etc. Dans l'idéal, elles devraient même aider au pilotage national et à la comparaison internationale.

Mais les données ouvertes pourraient aussi permettre d'évaluer l'action publique : Quels ont été les effets des réformes ? Ont-elles permis d'atteindre les objectifs en vue desquels elles ont été conçues ? Notre hypothèse de travail est que de très nombreux secteurs sont concernés, mais que l'abondance de données et de réformes font de l'ESR un terrain d'étude idéal pour concevoir une nouvelle méthodologie d'exploitation des données ouvertes dans l'évaluation des politiques publiques.

➤ VERS UNE MÉTHODOLOGIE D'EXPLOITATION DES DONNÉES OUVERTES DANS UNE PERSPECTIVE D'ÉVALUATION DES POLITIQUES PUBLIQUES

Notre programme de recherche repose sur une approche informée à la fois par les sciences sociales et la science des données, mobilisant des méthodes de recherche quantitatives et qualitatives. Cette approche pluri-disciplinaire est particulièrement difficile, car même lorsqu'elles étudient le même objet, ces sciences divergent notamment par les personnes, le vocabulaire, les technologies et les bases scientifiques. Élaborer une méthodologie commune à ces deux sciences représente donc un défi indispensable pour se doter d'un outil conceptuel d'analyse des transformations de l'action publique.

1 | Ce programme de recherche est financé par l'appel 80|Prime du CNRS

Après l'élaboration d'un état des lieux des données ouvertes disponibles ainsi que de leurs limites dans notre perspective de travail, nous nous attacherons à remonter la chaîne de production de la donnée en échangeant avec les acteurs clés des données ouvertes de l'ESR. En parallèle, nous tenterons d'évaluer les effets de certaines réformes, afin d'éprouver les conditions de faisabilité d'une évaluation des politiques publiques par les données. Nous espérons ainsi identifier les éventuelles modifications qui permettraient d'améliorer les données ouvertes et leur processus de production/publication, et ainsi contribuer à leur démocratisation.

Sources :

Denis, Jérôme, et Pontille, David. 2012. « Travailleurs de l'écrit, matières de l'information ». *Revue d'anthropologie des connaissances* 61 (1): 1-20.

Plateforme nationale data.gouv.fr : <https://www.data.gouv.fr/fr/>

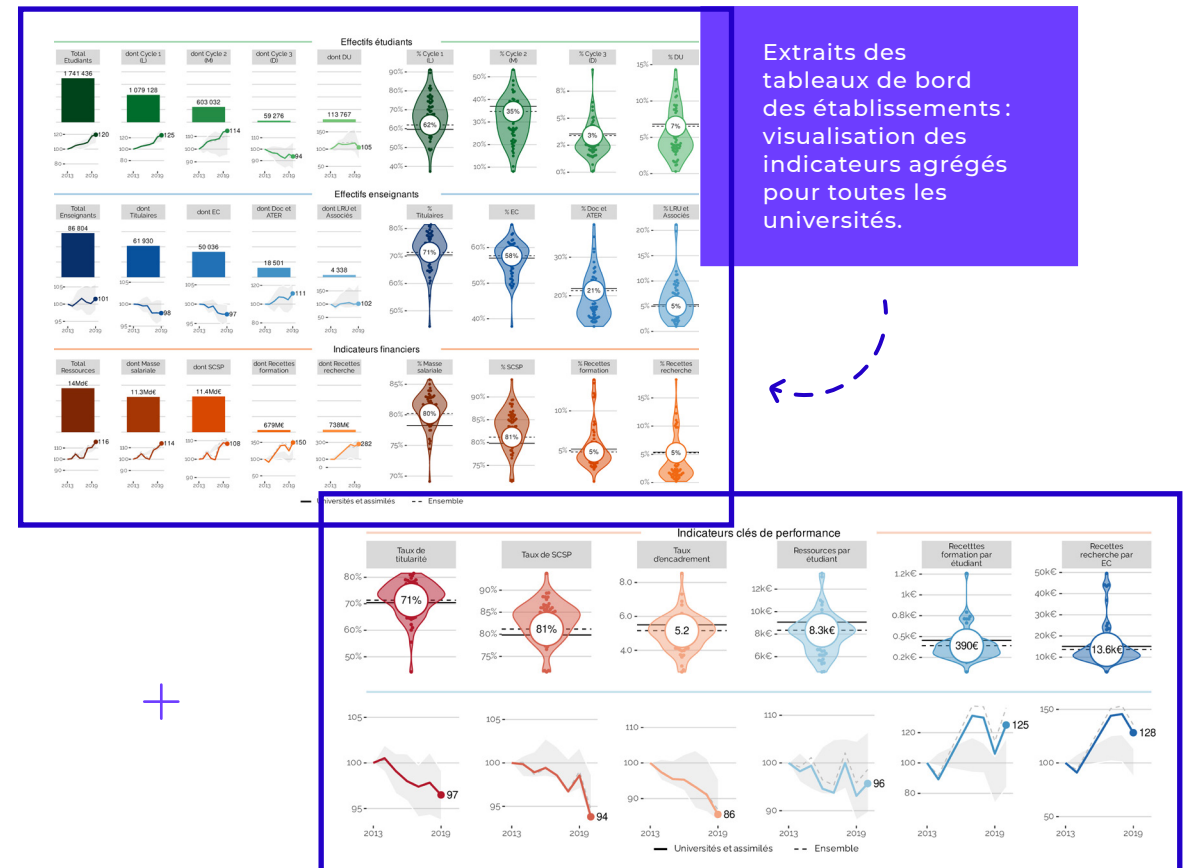
Plateforme du MESRI #dataESR : <https://data.esr.gouv.fr/FR/>

Des données brutes aux tableaux de bord

La construction des tableaux de bord des établissements de l'ESR a nécessité la fusion de 5 jeux de données ouvertes contenant un total de 373 variables et plus de 270 000 observations, pour 423 établissements entre 2006 et 2020.

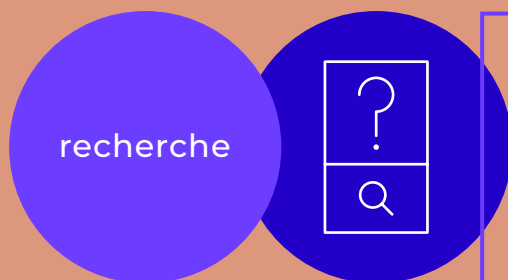
Nous les avons agrégés en 3 indicateurs primaires et 12 indicateurs secondaires, additionnés de 6 indicateurs clés de performance, reconstituant un jeu de données plus synthétique de 27 variables et 2670 observations, chacune concernant un établissement à une année.

Les indicateurs sont présentés sous trois formes permettant leur comparaison temporelle et spatiale. Une phase de bêta-test sur le site de la Conférence des praticiennes et praticiens de l'enseignement supérieur (CPESR) sera ouverte au moment de la publication de ces lignes.



Extraits des tableaux de bord des établissements : visualisation des indicateurs agrégés pour toutes les universités.





auteure
Françoise Genova,
directrice de Recherche
Emérite, Observatoire
Astronomique de
Strasbourg (ObAS),
Université de Strasbourg
– CNRS, UMR7550



Le Centre de Données astronomiques de Strasbourg

Où l'on comprend que le partage de données est un process de longue date et inhérent à toute recherche scientifique et technologique.

L'astronomie a été, bien avant que l'on commence à parler de Science Ouverte, une discipline pionnière pour le partage des données scientifiques. Les astronomes travaillent quotidiennement avec des données qu'ils trouvent en ligne dans les archives des grands observatoires sol et spatiaux et dans des centres thématiques qui mettent à leur disposition des bases de données « à valeur ajoutée ». Celles-ci regroupent, homogénéisent et diffusent des données, extraites en particulier des archives d'observatoires ou des publications académiques de la discipline.

Le Centre de Données astronomiques de Strasbourg (CDS) est l'une des ressources largement utilisées par la communauté internationale. C'est l'un des plus anciens centres de données au niveau international, toutes disciplines confondues : il a été créé en 1972 par l'Institut National d'Astronomie et de Géophysique (INAG) et l'Université Louis Pasteur (ULP), et il reste sous la responsabilité de leurs successeurs, l'Institut National des Sciences de l'Univers du CNRS et l'Université de Strasbourg. La qualité du travail du CDS et son impact du CDS sont pleinement reconnus au niveau national : il est inscrit sur la Feuille de Route Nationale des Infrastructures de Recherche depuis la première version de celle-ci en 2008, sa labellisation ayant été renouvelée depuis à chaque mise à jour de cette Feuille de Route.

La mission du Centre de données est depuis l'origine :

- Collecter les données utiles sur les objets astronomiques, sous forme électronique ;
- Les améliorer en les évaluant de façon critique et en les combinant ;

- Distribuer les résultats à la communauté internationale ;
- Conduire des recherches en utilisant les données.

Cette mission, qui anticipait largement les concepts qui sous-tendent le partage des données dans la Science Ouverte, continue à guider le travail du CDS. Celui-ci est une équipe de l'Observatoire Astronomique de Strasbourg (Université de Strasbourg – CNRS, UMR7550), qui comprend des chercheurs et enseignants chercheurs, des documentalistes spécialistes des données astronomiques et des ingénieurs informaticiens. La présence de ces différents profils qui travaillent en interaction assure que les différents aspects du travail nécessaires au succès du CDS sont pris en compte, les besoins de la communauté scientifique, la qualité du contenu des services et leur mise en œuvre tirant le meilleur parti des possibilités technologiques.

Les services du CDS génèrent plus d'un million et demi de requêtes par jour. Ils comprennent SIMBAD, la base de données de référence pour la nomenclature et la bibliographie des objets astronomiques hors système solaire, VizieR, la base de données de référence pour les grands catalogues astronomiques et les données attachées aux publications académiques, et l'atlas interactif du ciel ALADIN, qui comprend une base de données d'images du ciel gérée par le CDS mais donne aussi accès aux données des archives des observatoires.

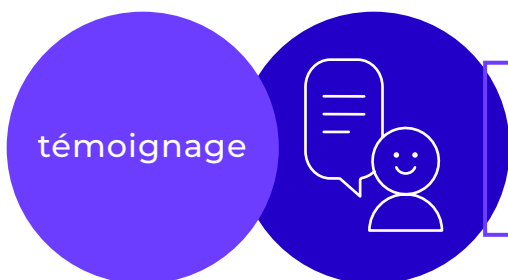
La communauté astronomique a en effet défini, bien avant les principes FAIR¹, un cadre de partage de ses données, l'Observatoire Virtuel (OV) astronomique, dont les standards², qui permettent aux données d'être **F**aciles à trouver, **A**ccessibles et **I**nteropérables, sont mis en œuvre par les producteurs de données et dans les outils d'accès aux données. Un format commun, FITS³, qui contient des informations sur la provenance des données, permet de **R**éutiliser celles-ci. Les données sont ouvertes (éventuellement après une période d'embargo) et leur mise à disposition avec les standards de l'OV permet à l'ensemble des données de la discipline d'apparaître comme une seule infrastructure de données globale, où une seule requête permet de retrouver toutes les données pertinentes quel que soit le service qui les met à disposition.

Le CDS est depuis l'origine de l'Observatoire Virtuel astronomique un des acteurs majeurs du développement de celui-ci. Ses services en sont des éléments incontournables.

1 | FAIR signifie Findable (Facile à trouver), Accessible, Interoperable, Reusable (Wilkinson et al. <https://doi.org/10.1038/sdata.2016.18>)

2 | Les standards de l'Observatoire Virtuel astronomique sont définis par l'International Virtual Observatory Alliance (IVOA), qui a été créée en 2002, <https://www.ivoa.net/>

3 | Flexible Image Transport System, <https://fits.gsfc.nasa.gov/>



témoignage

auteur
Bertrand Monthubert,
Président d'Occitanie
Data/EkitIA

Partager les données en confiance, c'est possible !

L'Open Data est une politique à double tranchant, où se côtoient les termes d'ouverture et de confiance, confidentialité et risques, éthique et avenir. Occitanie Data/Ekitia s'est emparé du sujet et lui donne un cadre.

Depuis plusieurs années de nombreuses initiatives ont lieu pour ouvrir les données : c'est la politique d'Open Data, qui place notre pays au 4ème rang mondial¹. Cette dynamique se retrouve bien entendu dans la science, avec la politique d'Open Science, et la volonté de mettre à disposition le plus largement possible les données issues de la recherche. Si cette politique est importante et doit nous mobiliser, elle a pour autant des limites : il n'est tout simplement pas possible de partager toutes les données, pour des raisons variées mais toutes légitimes : confidentialité des données (notamment personnelles), risques encourus si certaines informations sont connues (par exemple les zones de présence ou de reproduction d'espèces en voie de disparition faisant l'objet de braconnage), pour ne donner que deux exemples.

Malheureusement, très souvent quand une organisation constate que certaines données ne peuvent être ouvertes, elle les laisse fermées. Pourtant, il pourrait être très utile, et même souvent nécessaire, de partager ces données, dans des conditions particulières, afin de réaliser des traitements à partir de ces données tout en respectant la réglementation et en évitant les risques.

Pour permettre d'aller au-delà de la dualité données ouvertes/données fermées, des acteurs publics et privés, dont des acteurs académiques comme les universités de Toulouse et de Montpellier, ainsi que le CNRS, ont créé Occitanie Data/Ekitia. Il s'agit d'une association, en cours de transformation en groupement d'intérêt public d'envergure nationale, EkitIA, dont l'objectif est de créer un cadre de confiance éthique et souverain pour l'économie de la donnée. Sur la base d'une charte éthique que nous avons créée en 2020², nous accompagnons des organisations afin d'étudier si leurs cas d'usage respectent bien les principes éthiques qui légitiment la réalisation des traitements de données. Un label pourra désormais être attribué pour rendre visible le respect de notre charte.

1 | Global Open Data Index <https://index.okfn.org/place/>

2 | <https://www.occitaniedata.fr/la-charte-ethique/>



Charte Ethique Occitanie Data pour un développement responsable de l'économie de la donnée

Nous travaillons aussi sur la mise en place de solutions techniques permettant d'effectuer des croisements de données entre plusieurs organisations sans que les organisations ne puissent voir les données des autres. Enfin nous travaillons à la définition de modèles économiques alternatifs à celui des GAFA.

L'articulation des trois niveaux : cadre éthique, modèle économique, solutions techniques, nous permet de travailler sur des cas d'usages importants, comme la visualisation des zones les plus à risque pour la propagation de la COVID-19, la flexibilité énergétique ou pour améliorer la mobilité en zone urbaine. Il s'agit notamment de donner de nouvelles possibilités aux chercheurs pour conduire leurs travaux de recherche en utilisant des données aujourd'hui inaccessibles, mais aussi à permettre d'utiliser les données des laboratoires de manière plus importante, dans le cadre de collaborations respectueuses. Cela peut même parfois concerner les données à l'intérieur d'un même laboratoire, qui ont été récoltées dans des conditions différentes et ne peuvent pas être transmises.

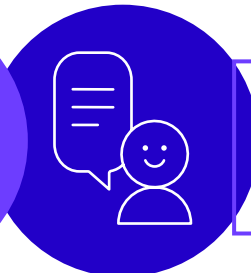
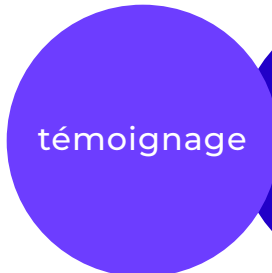
Au-delà de l'Open Data, de nouvelles voies d'utilisation des données de la recherche se présentent, au bénéfice de nos laboratoires et d'une science qui s'ouvre plus largement.

Occitanie Data/Ekitia

Créée en janvier 2019, Occitanie Data/Ekitia est une association de préfiguration d'un pôle d'économie de la donnée. Structurée en Occitanie mais rassemblant des acteurs au-delà, ses membres et partenaires sont des entreprises privées et publiques, des collectivités, des acteurs académiques ou encore des clusters et des pôles de compétitivité. Occitanie Data/Ekitia ambitionne de construire un cadre de confiance, éthique et souverain, destiné à permettre aux acteurs de partager et de croiser leurs données tout en respectant les intérêts des individus et des propriétaires des données.

<https://www.occitaniedata.fr/>





auteur
Joël Gombin,
cofondateur
de Dataactivist

1 | Jérôme Denis, Samuel Goëta. Rawification and the careful generation of open government data. *Social Studies of Science*, SAGE Publications, 2017, 47 (5), pp.604 - 629. doi : 10.1177/0306312717712473

L'Open data c'est Issy !

La ville d'Issy les Moulineaux fait le choix de la transparence, de la réactivité et de l'efficacité en pariant sur l'Open Data et c'est plutôt concluant.

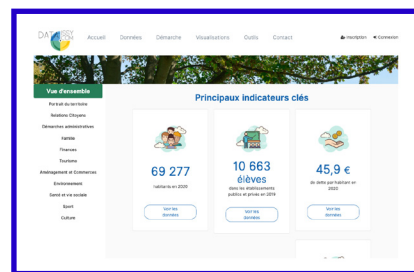
FAIRE DES DONNÉES UN OUTIL DE PILOTAGE DES SERVICES PUBLICS : L'EXEMPLE D'ISSY-LES-MOULINEAUX

Ouvrir des données, c'est bien. En faire un véritable outil de pilotage politique de l'administration de la ville, au service des citoyens, c'est mieux ! Telle pourrait être la devise de la ville d'Issy-les-Moulineaux.

En 2019, cette collectivité, pionnière dans sa démarche d'open data qui remonte à 2012, décide de construire et, surtout, publier sur internet des tableaux de bord sur l'activité des services. Ils devront être alimentés par des jeux de données disponibles sur le portail de données ouvertes de la ville. Cela procure un triple bénéfice : transparence pour le citoyen, qui peut vérifier la présentation qui est faite des données, voire la remettre en cause ; assurance que le tableau de bord prend en compte les données les plus fraîches d'activité des services ; efficacité en évitant des doubles voire triples saisies de données.

En 2019, notre coopérative Dataactivist a accompagné les services de la ville d'Issy-les-Moulineaux dans l'élaboration de ces tableaux de bord. Sept services pilotes ont été identifiés pour participer à ces premiers tableaux de bord. Dans un premier temps, des ateliers avec les agents ont permis de définir les indicateurs qui leur paraissaient à la fois pertinents pour objectiver l'activité de leur service, et réalistes quant à la possibilité d'extraire les données. Cette étape a également permis une réelle acculturation de ces agents sur les enjeux de la donnée et en particulier de l'open data. En effet, la condition *sine qua non* de réussite des projets data est cette capacité des acteurs à s'approprier les enjeux et le vocabulaire – sans cette culture de la donnée, point de salut !

Dans un deuxième temps, et après identification des données qui pourraient effectivement être mobilisées, les agents ont prototypé les tableaux de bord. Cela permet de se concentrer dans un premier temps sur l'objectif poursuivi, le sens – avant de se demander ce qui est effectivement faisable ou pas et à quelles conditions.



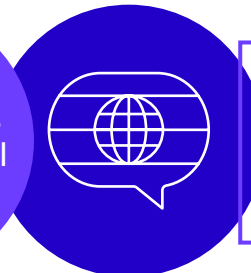
L'identification concrète des données à utiliser, de leur source, l'analyse de leur qualité et leur préparation a constitué l'étape suivante. Un format cible commun a été défini, afin que les jeux de données puissent être ingérés par le portail open data de la ville (fourni par Open-DataSoft), qui est utilisé pour fabriquer les tableaux de bord. Il ne s'agit pas là simplement d'une étape technique, de nettoyage sans intérêt : à travers ce travail de « brutification » des données¹, il s'agit de faire de données une source authentique. C'est lui qui sera actualisé régulièrement (chaque mois, chaque année...), et qui alimente les tableaux de bord et tous les indicateurs, y compris dans les autres services. De ce fait, il devient nécessaire de définir des garanties autour de ce jeu de données : fréquence de mise à jour, référents ou responsables, exigences de qualité, procédures de sauvegarde ou d'historique... Ce sont ainsi les bases d'une véritable gouvernance des données qui sont posées. Grâce à elle, les tableaux de bord et autres visualisations sont actualisés automatiquement, dès que des données fraîches sont disponibles. Les élus, les encadrants comme les citoyens peuvent avoir confiance en ces éléments.

Pour en savoir plus

[Consulter les tableaux de bord ici](#) →

[Lire le billet de blog que nous avons écrit sur cette expérience ici](#) →

[En savoir plus sur Dataactivist ici](#) →



auteure
Valérie Le Strat,
urbaniste SI,
Amue

Quand la donnée se met au service de la transformation digitale de l'établissement.

Extract de l'article figurant dans le rapport 2020 de la délégation Française Educause : L'approche orientée données dans la transformation digitale ou comment les données d'un étudiant peuvent contribuer à améliorer ses chances de succès

Soumis à un fort climat de compétitivité, les établissements d'Enseignement Supérieur aux Etats-Unis doivent impérativement placer attractivité et maîtrise des coûts au cœur de leurs préoccupations. Lors de l'édition 2020 d'Educause, plusieurs établissements ont témoigné de la façon dont la mise en œuvre d'une approche orientée donnée pouvait se mettre au service de ces deux enjeux.

Ainsi, pour Virginia Tech et St John's University, **la maîtrise des coûts et des ressources** passe par l'accès à des données structurées, fiables et de qualité permettant la production d'indicateurs constituant de précieuses aides à la décision.

Pour l'université de Boise, la croissance naturelle des écosystèmes avait entraîné une fragmentation de la donnée, générant des surcoûts de maintenance et des temps de latence élevés lorsqu'il s'agissait d'identifier la "donnée correcte" au sein d'une multitude de sources. Une situation dont l'établissement s'est accommodé jusqu'à ce que le degré de compétitivité entre établissements le pousse vers une démarche d'intégration du SI et de maîtrise de la qualité des données manipulées.

La transformation Digitale (DX) d'un établissement au travers d'une approche orientée donnée peut aussi servir **l'attractivité**, comme l'a illustré la Florida State University. En appréhendant le parcours étudiant d'une manière globale, au sein du campus mais aussi à l'extérieur (logement, restauration, transports, activités, autres études en parallèles ou dans le cadre de la formation tout au long de la vie), l'établissement a pu proposer aux étudiants des services dignes d'Amazon ou de Netflix en termes d'ergonomie ou de réactivité mais aussi un accès facilité à leurs données où qu'ils se trouvent et y compris après leur passage dans l'établissement.





Comme l'a présenté la Pennsylvania State University, l'attractivité passe également par la réussite des étudiants et donc l'identification précoce des décrocheurs. Dans cette perspective, les établissements n'hésitent pas à investir dans des projets de collecte massive de données personnelles sur de très longues périodes afin d'améliorer la prédictivité de leurs modèles mis en œuvre par des Intelligences Artificielles.

La collecte et l'analyse des données collectées nécessitent à la fois un SI parfaitement intégré au sein duquel la donnée est en qualité et une maîtrise des technologies liées à la donnée, que cela soit pour son traitement (analytics, big data, IA etc.) ou son accès (disponibilité, gestion des identités, WIFI, etc.).

En ce qui concerne la stratégie à mettre en œuvre pour une telle démarche, tous les établissements s'accordent sur l'ampleur et la durée d'un tel projet. Il s'agit d'un projet d'établissement qui nécessite l'implication de l'ensemble des acteurs, ainsi qu'un portage politique au plus haut niveau. Une acculturation de l'établissement à la donnée est indispensable à l'acquisition d'une maturité qui permettra d'éviter certains écueils. Parmi ceux-ci, les établissements citent une collecte non maîtrisée de données (parfois au mépris de la confidentialité), une incapacité à exploiter ces données ou encore une compréhension erronée de la notion de responsabilité vis-à-vis d'une donnée.

En 2020, les établissements les plus avancés ont témoigné de la facilité à gérer la pandémie de Covid au sein de leur campus, notamment pour permettre une réouverture des campus au plus tôt, de manière sécurisée.

Pour aller plus loin

Le rapport de la délégation Française Educause 2020 est à lire ici →

Deux articles traitent notamment du sujet de la donnée :

- Des données, des données, partout - L'expansion du Machine Learning par Dominique Verez
- L'approche orientée données dans la transformation digitale ou comment les données d'un étudiant peuvent contribuer à améliorer ses chances de succès par Bruno Urbero & Valérie Le Strat

Bibliographie :

- 1 | Communicating Predictive Analytics Findings Effectively and Ethically - Alejandra Acosta - New America Foundation
- 2 | New Data Architectures Driving Innovation and Data Maturity in Higher Education, Mathew Harrowing (Intenda LLC), Abhay Joshi (Virginia Tech), Sherry Bennett (DLT-Tech Data)
- 3 | Paying Down Technology Debt to Advance Innovation, Anne Rocco Pacione, St. John's University
- 4 | Building a resilient campus experience on a cloud based integration platform, Harnan Londoño, Barry University ; Brian Merrick, Boise State University ; Chris Wessells, Dell Technologies
- 5 | Reimagine Student engagement, Solveig Brown, Florida State University ; Sean Brown, Florida State University ; Rob Curtin, Microsoft Corporation
- 6 | Provost-CIO Strategic Partnerships for Innovation, Ben Vinson III, Case Western Reserve University ; Lev Gonick, Arizona State University ; Sue Workman Case Western Reserve University ; Mark Searle, Arizona State University ; Elias Eldayrie, University of Florida
- 7 | Spinach or Kryptonite? Leveraging Data Governance to Improve Analytics, Cathy Bates Vantage Technology Consulting Group ; Joanna Grama, Vantage Technology Consulting Group ; Augie Freda, University of Notre-Dame ; Rachel Serrano, University of North Carolina at Chapel Hill
- 8 | Data First: Lessons from JCU & EAB's Education Data Hub Collaboration, Jim Burke, John Carroll University ; Danielle Yardy, EAB
- 9 | Leveraging ML to Advance Data Maturity and Innovation in Higher Education, Sherry Bennett, DLT-Tech Data ; Michael Anderson, DLT Solutions ; Yawar Shams, HelioCampus ; Brian Wright, University of Virginia
- 10 | How a Smart Campus Can Build a Better Normal Now, Sasi Pillayn Washington State University ; Sabina Sakensa, Citylife Inc ; Michael Berman, California State University



auteur
Recension d'un article par **Bertrand Mocquet**, expert numérique, Amue

Big Data en clair

Mieux soutenir la recherche sur les Big Data dans un contexte universitaire, c'est l'objet de ce rapport US qui fournit constats et recommandations. On en découvre les grandes lignes.

Préambule

Nous présentons ici une traduction partielle de l'article intitulé « Big Data Infrastructure at the Crossroads: Support Needs and Challenges for Universities »¹ publié le 1^{er} décembre 2021 sur Ithaka S+R, une association new-yorkaise en conseil stratégique et d'accompagnement qui aident les institutions US à améliorer leurs performances et à poursuivre leurs missions.



1 | Ruediger, D., Atwood, T. P., Bharti, N., Briones, B., Campbell, P., Carey, P., Castillo, D., Zickel, L. (2021, December 1). Big Data Infrastructure at the Crossroads: Support Needs and Challenges for Universities. <https://doi.org/10.18665/sr.316121>

Le programme de services de soutien à la recherche de Ithaka S+R explore les tendances actuelles et les besoins de soutien dans la recherche universitaire. Notre projet le plus récent dans ce programme, « Soutenir la recherche sur les BigData », s'est concentré spécifiquement sur l'utilisation émergente et rapide des BigData dans la Recherche de toutes les disciplines et tous les domaines.

Dans le cadre de notre étude, nous nous sommes associés à des bibliothécaires de plus de 20 collèges et universités, qui ont ensuite mené plus de 200 entretiens avec des professeurs. Ces entretiens ont permis de mieux comprendre les méthodologies de recherche et les besoins de soutien des chercheurs travaillant dans un large éventail de disciplines.

Ce rapport fournit un compte rendu détaillé de la manière dont la recherche sur les Big-Data est menée dans des contextes universitaires, en se concentrant sur l'identification des méthodologies, des flux de travail, des résultats et des défis typiques auxquels les chercheurs en mégadonnées sont confrontés. Des détails complets et des recommandations exploitables pour les acteurs sont proposés dans le corps du rapport, qui offre des conseils aux universités, aux bailleurs de fonds et à d'autres personnes intéressées à améliorer les capacités institutionnelles et à favoriser les échanges pour mieux soutenir la recherche sur les Bigdata. Nos principales conclusions sont regroupées dans les domaines suivants :

→ **Tension et interaction entre les perspectives disciplinaires et interdisciplinaires**

La recherche sur les BigData est une entreprise interdisciplinaire menée par des praticiens travaillant dans des cadres institutionnels encore organisés autour de disciplines. Des structures d'incitation divergentes, des cultures et un accès inégal au financement peuvent affecter la participation disciplinaire aux projets de recherche sur les mégadonnées. De plus,



l'utilisation généralisée de méthodologies issues de l'informatique et des sciences des données - surtout une tendance claire vers l'apprentissage automatique - a créé des tensions parmi les chercheurs et soulevé des questions sur l'importance relative des perspectives disciplinaires.

→ **Gestion des données complexes** . À une époque d'abondance relative de données, les chercheurs évitent souvent les dépenses liées à la génération de nouvelles données et choisissent plutôt de travailler avec les données existantes dans la mesure du possible. Le travail d'acquisition, de nettoyage et d'organisation des données est généralement l'aspect le plus laborieux des projets de Big Data.

→ **Structures de collaboration** . La recherche sur les Bigdata est presque toujours une entreprise collective impliquant des étudiants, des professeurs, du personnel, des collègues, des clients et des collaborateurs de l'enseignement supérieur et au-delà. Les laboratoires sont les unités de base de la recherche et, en leur sein, les étudiants (premier cycle et cycles supérieurs) apportent une contribution significative au processus de recherche. Les chercheurs privilégient également souvent les ressources informatiques locales en laboratoire par rapport aux options de stockage et de calcul centralisées du campus, y compris les services de cloud computing.

→ **Partage des connaissances** . Bien que les articles évalués par des pairs restent la forme de communication savante la plus encouragée, les chercheurs sont largement engagés dans le partage ouvert des résultats de la recherche, y compris les

données et le code. Cependant, les pratiques de partage académique reflètent un spectre qui s'étend bien au-delà du partage formel dans des référentiels ouverts qui répondent aux normes FAIR de recherche, d'accessibilité, d'interopérabilité et de réutilisation, englobant de nombreux types de partage informel avec des collègues. Les obstacles au partage formel incluent des perceptions largement répandues selon lesquelles de nombreuses données sont soit dérivées, de faible qualité ou recueillies à partir de sources inappropriées pour un partage ouvert.

→ **Défis éthiques** . Les dimensions éthiques de la recherche sur les Bigdata restent contestées et certains chercheurs sont incertains des meilleures pratiques en matière de conduite de recherche éthique. Bien que les directives de l'International Review Board (IRB) soient appréciées par la mise en place de commissions de contrôle du respect de l'éthique, certains chercheurs ont exprimé des inquiétudes quant au fait que les règlements de l'IRB ne sont pas bien adaptés aux méthodes de recherche nouvelles ou en évolution.

→ **Accompagnement et Formation** . Les chercheurs ont tendance à privilégier les méthodes de formation informelles, telles que les didacticiels sur Internet, à la formation formelle aux méthodes de mégadonnées. Bien que de telles méthodes fonctionnent bien pour résoudre des problèmes immédiats, elles sont moins bien adaptées à l'acquisition de connaissances fondamentales, laissant le potentiel d'angles morts dans la recherche universitaire.



auteur
Recension d'un article par **Bertrand Mocquet**, expert numérique, Amue

7 principes à mettre en œuvre pour assurer la réussite d'une université

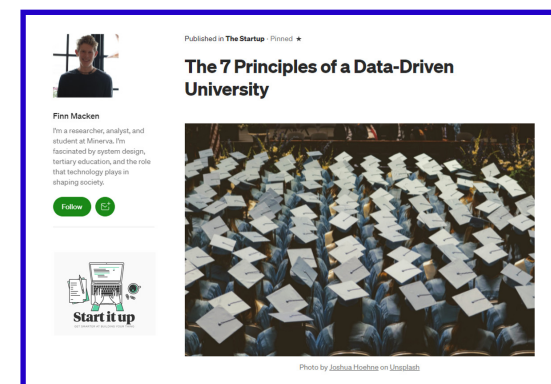
L'université par le prisme d'un étudiant américain qui mesure combien les données sont les éléments clés d'un secteur marchand international, mais évalue aussi le chemin à parcourir pour répondre aux enjeux sociétaux.

Nous présentons ici une traduction de l'article intitulé «The 7 Principles of a Data-Driven University»¹ par Finn Macken, chercheur, analyste et étudiant à l'Université de Minerva (USA), se présentant comme fasciné par la conception des systèmes, l'enseignement supérieur et le rôle que joue la technologie dans le façonnement de la société.

LES UNIVERSITÉS NE S'ADAPTENT PAS MAIS LE DEVRAIENT

L'auteur américain nous rappelle le contexte de l'Enseignement supérieur et comment il est devenu un secteur marchand international, mettant les universités historiques dans une concurrence avec des nouvelles formes de formations et certifications, et cela dans le monde entier grâce à des bourses aujourd'hui disponible

1 | Macken, F. (2020). The 7 Principles of a Data-Driven University. The Startup. Repéré à <https://medium.com/swlh/the-7-principles-of-a-data-driven-university-3525ad9bacc6>





pour les étudiants. La crise n'a pas arrangé cela, on estime depuis que la taille du marché de l'apprentissage en ligne devrait croître à de +10,85 % d'ici 2025

Et pour autant, les universités ne peuvent pas disparaître car elles restent les « gardiens de la raison » dans ce secteur selon l'OCDE et dans une société qui a étendu la définition d'« étudiant » de « un jeune de dix-huit à vingt-cinq ans se préparant à entrer sur le marché du travail » à « tous ceux qui veulent apprendre quoi que ce soit », l'espace de l'enseignement supérieur doit offrir des expériences et des parcours d'apprentissage différenciés, une vision qui devrait inclure des institutions physiques.

↳ L'UNIVERSITÉ DU FUTUR, INVARIABLEMENT, SERA AXÉE SUR LES DONNÉES

Pendant cette période de réinvention et de changement, les données constituent l'un des (sinon le plus) outils puissants à la disposition des institutions. C'est le mécanisme par lequel les universités peuvent devenir réactives et adaptatives, grâce auxquelles elles peuvent répondre aux besoins de leurs parties prenantes. Comme le déclare Kati Haycock, PDG de l'Education Trust, « montrez-moi une institution qui a vraiment amélioré les résultats de ses étudiants, et je vous garantis que les données font partie de leur histoire ».

↳ LES UNIVERSITÉS AXÉES SUR LES DONNÉES NE SONT PAS CELLES QUI COLLECTENT DES DONNÉES ; CE SONT CELLES QUI AGISSENT GRÂCE AUX DONNÉES

Cette définition ne concerne que de manière tangentielle la conception de systèmes de collecte de données, ce qui est la condition préalable typique pour être une « université axée sur les données ». La réalité est que la plupart des institutions ont déjà beaucoup d'informations pertinentes, et collecter ne devrait pas être la première priorité. Les LMS et CMS modernes, les recherches menées par les départements marketing et les rapports sectoriels peuvent fournir des données incroyablement utiles, en particulier concernant les performances académiques et l'engagement ; et il existe des dizaines d'applications qui simplifient le processus de collecte, d'assemblage et de visualisation des données. Au lieu de cela, le problème central est que les institutions, pour diverses raisons, n'intègrent pas les données dont elles disposent dans la prise de décision.

Les universités axées sur les données ne sont pas des institutions qui disposent simplement de systèmes de capture, de nettoyage et de centralisation des données. C'est la partie facile. Ce sont des institutions qui ont créé un cadre d'action fondé sur des preuves, allant de la direction générale au personnel académique, en passant par le marketing, les finances et les affaires étudiantes. Ce sont ces institutions qui ne se contentent pas de collecter des informations, mais agissent en conséquence.

↳ L'UNIVERSITÉ AXÉE SUR LES DONNÉES EN UN ENSEMBLE DE SEPT PRINCIPES LARGEMENT APPLICABLES.

Ces exigences sont loin d'être les seules, mais elles sont à la fois puissantes et généralisables, et toute institution qui y adhérerait pleinement améliorerait radicalement à la fois sa prise de décision et la manière dont elle met en œuvre les initiatives :

2 | Reports Valuates Cision (2020). E-Learning Market Size is Expected to Grow at a CAGR of 10.85% by 2025 | Valuates Reports. Repéré ici

3 | Stein, S., & de Andreotti, V. O. (2016). Cash, competition, or charity : International students and the global imaginary. Higher Education, 72(2), 225-239. <https://doi.org/10.1007/s10734-015-9949-8>



The 7 Principles of a Data-Driven University

Principle #1: Never act without first developing institutional knowledge
Principle #2: Understand what problem you're trying to solve
Principle #3: Construct and test assumptions
Principle #4: Create prototypes
Principle #5: Build and utilize systems for feedback as early as possible
Principle #6: Measure and analyze the right things
Principle #7: Every initiative should be iterative across its lifecycle

Principe n°1

Ne jamais agir sans avoir développé au préalable des connaissances institutionnelles

Les universités axées sur les données comprennent que le changement est nécessaire et qu'il nécessite une compréhension détaillée (et holistique) de la manière dont différents systèmes et forces façonnent le succès des initiatives. Cela est vrai quelle que soit l'institution, et quel que soit le projet. Ils comprennent également que, à cause de cela, le changement est incroyablement difficile.

Les dirigeants axés sur les données prennent en compte la culture, la dynamique du pouvoir et la conception organisationnelle, à de nombreux niveaux différents.



Principe n°2

Comprenez quel problème vous essayez de résoudre

Dans une université axée sur les données, les chefs de projet sont extrêmement clairs sur le problème qu'ils essaient de résoudre, pourquoi il est si important qu'il soit résolu et comment l'initiative le fera. Il doit s'agir d'un processus continu : tout au long du cycle de vie du projet, les décideurs doivent évaluer si la caractérisation de leur problème a changé et si les fonctionnalités mises en œuvre répondent réellement aux problèmes des parties prenantes.

Principe n°3

Construire et tester les hypothèses

Encadrer la gestion du changement en termes d'hypothèses est précieux car il facilite le centrage sur les parties prenantes : les décideurs doivent réfléchir aux aspects du projet qui ont réellement été validés et à ceux qui restent à prouver.

Principe n°4

Créer des prototypes

Les prototypes sont utiles car (lorsqu'ils sont bien conçus) ils fournissent une mesure précise de l'efficacité avec laquelle une initiative résout un problème et des sources de friction susceptibles d'émerger lorsqu'elle est mise en œuvre dans une plus grande partie de l'université.

Principe n°5

Construire et utiliser des systèmes de rétroaction le plus tôt possible

La rétroaction est le mécanisme par lequel les hypothèses sont validées, les initiatives sont réitérées et une université devient axée sur les données.

Principe n°6

Mesurer et analyser les bonnes choses

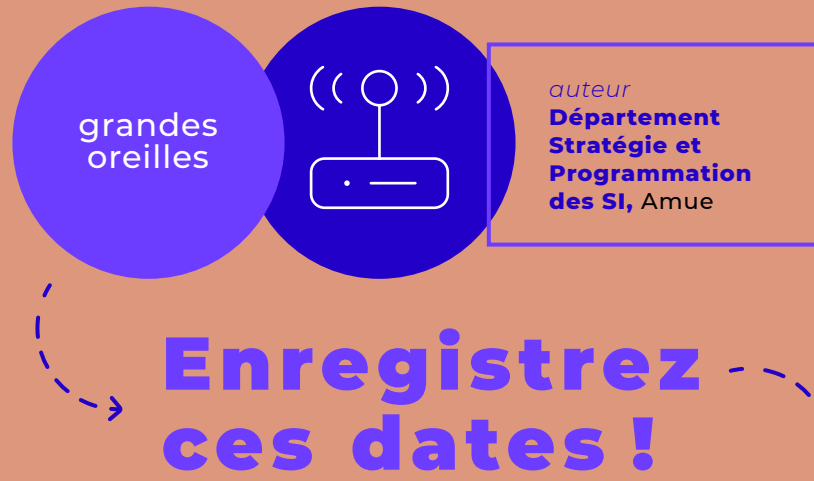
Il est de la responsabilité de ceux qui mettent en œuvre un projet de décider des méthodes qui reflètent la réalité pour mesurer les bonnes choses.

Principe n°7

Chaque initiative doit être itérative, tout au long de son cycle de vie

L'itération est l'outil le plus puissant dont disposent les institutions pour accomplir le changement. Il surmonte les obstacles, satisfait les contraintes, résout les problèmes, renforce les forces, réprime l'opposition. C'est la façon dont les hypothèses sont mises à jour, les commentaires sont intégrés et un projet est amélioré.

L'auteur ajoute en conclusion, qu'il est peut-être impossible de créer une institution véritablement axée sur les données, mais il est possible d'en créer une plus flexible, plus réactive et plus adaptée aux utilisateurs. Il est possible de travailler au sein de systèmes défectueux et de les modifier pour obtenir de meilleurs résultats. Et il est possible de créer une université capable de s'adapter et d'aborder à la fois ses propres problèmes et ceux de la société.



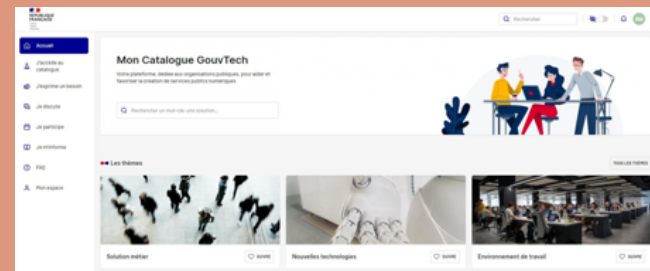
Les Assises du CSIESR 2022,

Elles auront lieu du **8 au 10 juin 2022** au Magic Circus de Marne la Vallée. L'appel à contribution va paraître dans les jours prochains

<https://www.csiesr.eu/sujets/event/assises-nationales>

Journées nationales de l'Innovation Pédagogique dans l'Enseignement supérieur (JIPES)

Elles ont pour vocation de constituer un repère annuel pour marquer les évolutions des pratiques et des politiques pédagogiques dans l'enseignement supérieur. Rendez-vous à la Cité des Sciences et de l'Industrie à Paris les **10 et 11 mars 2022**.



Label gouvtech

Après avoir eu l'honneur d'être bêta-testeur du service, nous vous informons de la sortie de "Mon catalogue Gouvtech" par Sabine Guillaume, de la DINUM. Un catalogue de service destiné aux acteurs publics (description, usage..), une plateforme d'interactions (informations, events, REX et nouveaux besoins...) disponible à <http://mon.catalogue.numerique.gouv.fr>. Les entreprises peuvent s'inscrire et se référencer ici →

Démarches Simplifiées du Sup'

Démarches Simplifiées est un logiciel sous licence libre développé par la Dinum, qui a été conçu pour accélérer la dématérialisation des démarches administratives. Nous vous informons de son ouverture officielle du service de dématérialisation pour les adhérents de l'Agence. La version Sup' est portée par de l'amue et opérée par l'ADULLACT. Plus d'informations sur

<https://www.amue.fr/systeme-dinformation/metier/demarches-simplifiees-du-sup/>



L'Amue labélisée Territoire numérique

Le Territoire Numérique Libre est un label ouvert à toutes les collectivités françaises depuis 2016 : villes, départements, régions, établissements publics, etc. C'est une initiative de l'association ADULLACT, en partenariat avec d'autres associations et institutions reconnues pour leurs actions en faveur du logiciel Libre et la promotion du numérique citoyen dans les collectivités territoriales : l'AFUL, l'April, l'association Déclit, le cluster Naos, la ville de Saint-Martin-d'Uriage et le Conseil National du Logiciel Libre (CNLL).

L'Amue obtient le Label Territoire Numérique Libre niveau 3, valorisant sa démarche d'ouverture du code et l'utilisation de solutions libres pour ses nouvelles solutions

Prix science ouverte du logiciel libre de la recherche

Les prix, science ouverte données de la recherche et science ouverte logiciel libre de la recherche, sont inscrits dans le deuxième Plan national pour la science ouverte annoncé par la ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation en juillet 2021. Ils récompenseront des initiatives emblématiques en la matière. Les prix seront remis début février 2022. Pour plus d'information, lire cette page sur le site « ouvrir la science ».



#DemoES

Saluons nos collègues issus des 17 projets d'établissements ou de réseaux d'établissements identifiés comme démonstrateurs « prêts à mettre en place et à déployer une stratégie de transformation numérique dans l'ensemble de leurs formations » par l'AMI « Démonstrateurs de l'enseignement supérieur » dans la stratégie Enseignement et Numérique porté par l'ANR.

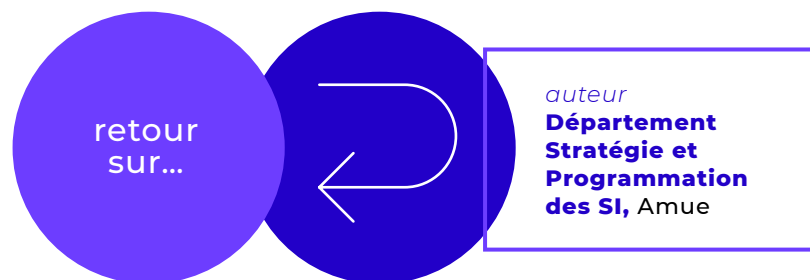
L'Amue en suivra quelques-uns pour évaluer les transformations induites par ces projets sur les solutions du système d'information actuel ou à venir.

<https://www.gouvernement.fr/numerique-educatif-le-premier-ministre-dresse-le-bilan-des-enseignants-de-la-crise-et-presente-des>

Webinaires en Inter-Assos ANSTIA/CSIESR/CUME

Une série d'actions est menée entre trois associations professionnelles l'ANSTIA, le CSIESR et le CUME dont le lancement a eu lieu le 8 Juillet 2021. Il s'agit de répondre au moyen de plusieurs webinaires à la question principale « Comment les transformations sociales et numériques impactent-elles l'articulation entre nos différents métiers. Quelles organisations seront pérennes post crise sanitaire ? ». Plusieurs évènements à suivre durant l'année universitaire 2021-2022. Prochain rdv en ligne le 20 janvier 2022 à 16h.



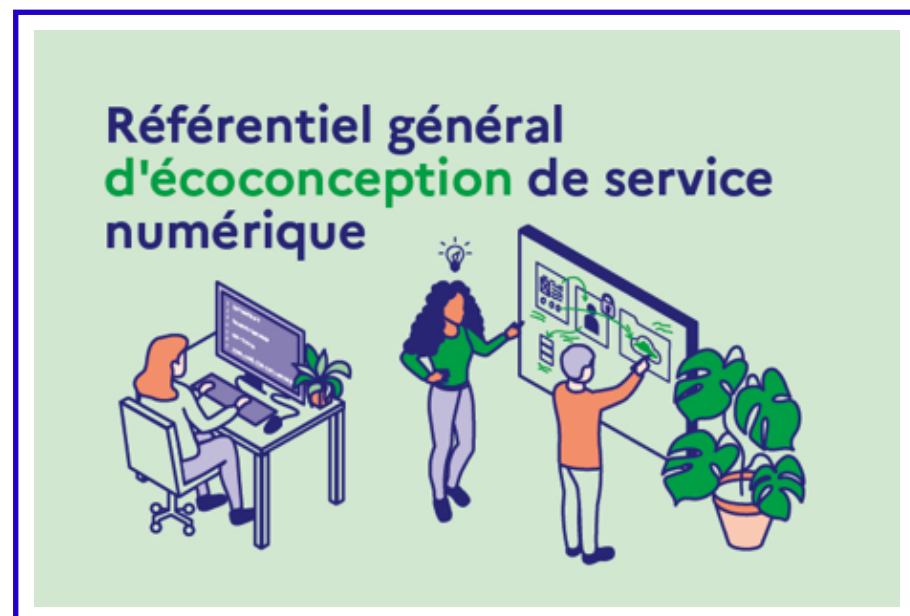


Référentiel Général d'Ecoconception de Services Numériques (RGESN)



On revient sur le numérique durable (Collection numérique N°12 – Numérique responsable, en décembre 2020), en vous proposant l'accès à ce référentiel, mené dans le cadre de la mission interministérielle « Green Tech » et co-pilotée par la Direction interministérielle du numérique (DINUM), le Ministère de la Transition Écologique, mais également l'ADEME et l'Institut du Numérique Responsable.

<https://ecoresponsable.numerique.gouv.fr/publications/referentiel-general-ecoconception/>



Rencontres sur les enjeux du logiciel libre dans la recherche **#FLOSS_ESR**
Ouvrir et promouvoir les codes sources produits par la recherche

ATELIER #1 14 janvier 2022 9h - 12h30, en ligne
Contribution à la production de codes sources en recherche : quelles pratiques, quels enjeux ?

JOURNÉE DE « CONVERGENCE » 3 février 2022 9h - 18h en présentiel (à Paris)
Ouverture et valorisation : pistes d'action au sein de la recherche

ATELIER #2 21 janvier 2022 9h - 12h30, en ligne
Soutenir les logiciels libres de la recherche : de la maturation à la valorisation

Ces rencontres vous sont proposées par :
OUVRIR LA SCIENCE ! **inno** **et alab**



Enjeux du logiciel libre dans la recherche



Le groupe de travail Logiciels libres du comité pour la science ouverte (COSO) et le pôle logiciels libres d'Etalab organisent une série de trois rencontres en janvier et février 2022 sur les enjeux d'ouverture et de valorisation des codes sources de la recherche. L'occasion de feuilleter à nouveau la Collection numérique N°13 – Vive le Numérique libre, de février 2021.

[Inscription et proposition à la facilitation des ateliers ici](#)



L'occasion de vous souhaiter une bonne et numérique (souhaitée pas subie) année 2022. Plein de courage pour les nouveaux projets à déployer et les présents à maintenir.

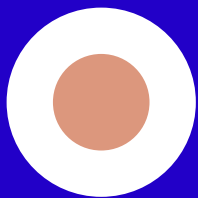


Nos meilleurs vœux pour 2022

décembre 2021



+



amue.fr

+

**prochain
numéro**

Le numéro
de février 2022
sera consacré
aux « Veilles »

+

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

