

Photo by Dũng Trần Việt on

L'ESR vu par le prisme de la donnée universitaire

Une autre vision de nos organisations

La collection numérique de l'Amue

NOVEMBRE 2019

Edito Par Stéphane Amiard, Vice-Président Numérique & Patrimoine, Université d'Angers

La donnée, sujet de société

La donnée, bien individuel, collectif ou commun ou bien d'autres est aujourd'hui source d'enjeux économiques et sociétaux incontournables. L'émergence des monstres de l'économie numérique en est une preuve sans conteste. Les questions autour de la donnée, sa qualité, sa mise en cohérence avec d'autres données, son traitement, son stockage, son archivage sont pléthoriques. Les usages parfois non maîtrisés de ses données, même à caractère personnel, par les puissances publiques et privées : vol, usurpation, traitement inapproprié, profiling indésirable, contrôle de masse, révélation non désirée, aboutissent à des tensions au sein de nos sociétés.

Nœud gordien

Cette situation est amplifiée par un phénomène bien connu, où les réglementations successives essayant de protéger le citoyen des effets pervers de l'économie libérale appliquée à la donnée, celui-ci dans le même temps par ses usages en divulgue davantage dans la sphère publique : c'est un véritable nœud gordien.

Une opportunité, un défi

Pour autant, et c'est de mon point de vue à la fois une opportunité unique et un défi majeur pour les universités : tranchons ce nœud. Devenir une organisation exemplaire en termes de production, valorisation et mise à disposition

de données : pour servir l'open science, pour assurer la réussite éducative et professionnelle de tous, pour garantir la souveraineté des actifs numériques du citoyen.

A l'heure de l'accélération de la transformation numérique par l'usage massif des algorithmes, les universités sont au pied de la montagne. Soit elles passent la main comme dans la transformation 2.0 de l'internet, ou elles relèvent le défi avec la mise en place à court terme de plans de gouvernance de la donnée, de schémas directeurs, de recrutements de CDO (Chief Data Officer), la création des services de transformation numérique ... et des résultats rapides.

Arme politique

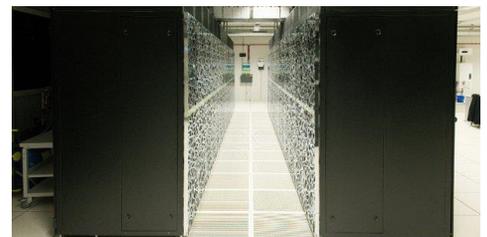
La donnée doit sortir des silos des SI et devenir une arme politique pour les établissements ; une forme de diplomatie de la donnée : oui l'université est une organisation qui fournit des données qualifiées pour l'avenir de nos sociétés et non elle ne fera pas la promotion des actifs numériques de ses usagers et des citoyens dans la ligne droite des "actifs corporels" du livre d'anticipation de Bernard Mourad.

DANS CE NUMERO



Photo Chatawat Pawapowadon de Pixabay

Mipnes : Vers une culture de la donnée Page #3



Source CINES

CINES : Comment transmettre aux futures générations la mémoire des données? Page #13



La "Donnée universitaire", de quoi parlons-nous ?

Par le Département Stratégie et Programmation du Système d'Information de l'Amue

ACQUERIR

La Donnée, les data

Dans toutes les organisations, la **Donnée**, nous utiliserons dans cet article la majuscule pour la définir comme concept, ou les **data** pour reprendre le mot latin pluriel, est un élément stratégique, qu'il s'agisse de données sur les clients, du référentiel des articles pour les entreprises de commerce, de documents administratifs ou de coordonnées d'usagers d'un service publique, etc...

Les données sont affublées de très nombreux qualificatifs : interoperables, de qualité, consolidées, personnelles, saisies, déstructurées, normalisées, urbanisées, archivées, calculées, de référence, sensibles, quantitatives, qualitatives, structurées, capturées, échappées, transitoires, dérivées, méta, fraîches, accessibles, exploitables, découvrables... Quant aux data, on les considère souvent Open, Big, Fair, Smart, Crowd ou associées aux anglicismes tels que centric, driven, hub, linked, mining, scientist, visualization, etc...

Dorénavant principalement numériques, les données d'une organisation existent encore sous d'autres 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 structurés, ...

Le sujet couvre à la fois les données, les informations et les connaissances ; ce dernier point résonnant particulièrement dans le contexte de l'Enseignement Supérieur et de la Recherche.

La Donnée universitaire existe-t-elle ?

Pour le secteur public, la stratégie de l'Etat sur la donnée est portée par [Etalab](#).

Nous le verrons dans ce numéro, la Donnée est, pour nos établissements, un vaste sujet et un sujet d'importance pour le pilotage, la gestion, les activités, les usagers... Il couvre l'ensemble des strates au sens de l'urbanisation à savoir la stratégie, le métier, le fonctionnel, l'applicatif et l'infrastructure.

Mais, quelle est la particularité des établissements de l'Enseignement Supérieur et de la Recherche, dans le domaine de la Donnée? Hormis des contenus spécifiques portés par les cœurs de métier tels que l'enseignement et la recherche, en quoi les "données" des universités et établissements seraient-elles particulières ? Les établissements rencontrent les mêmes problématiques que d'autres

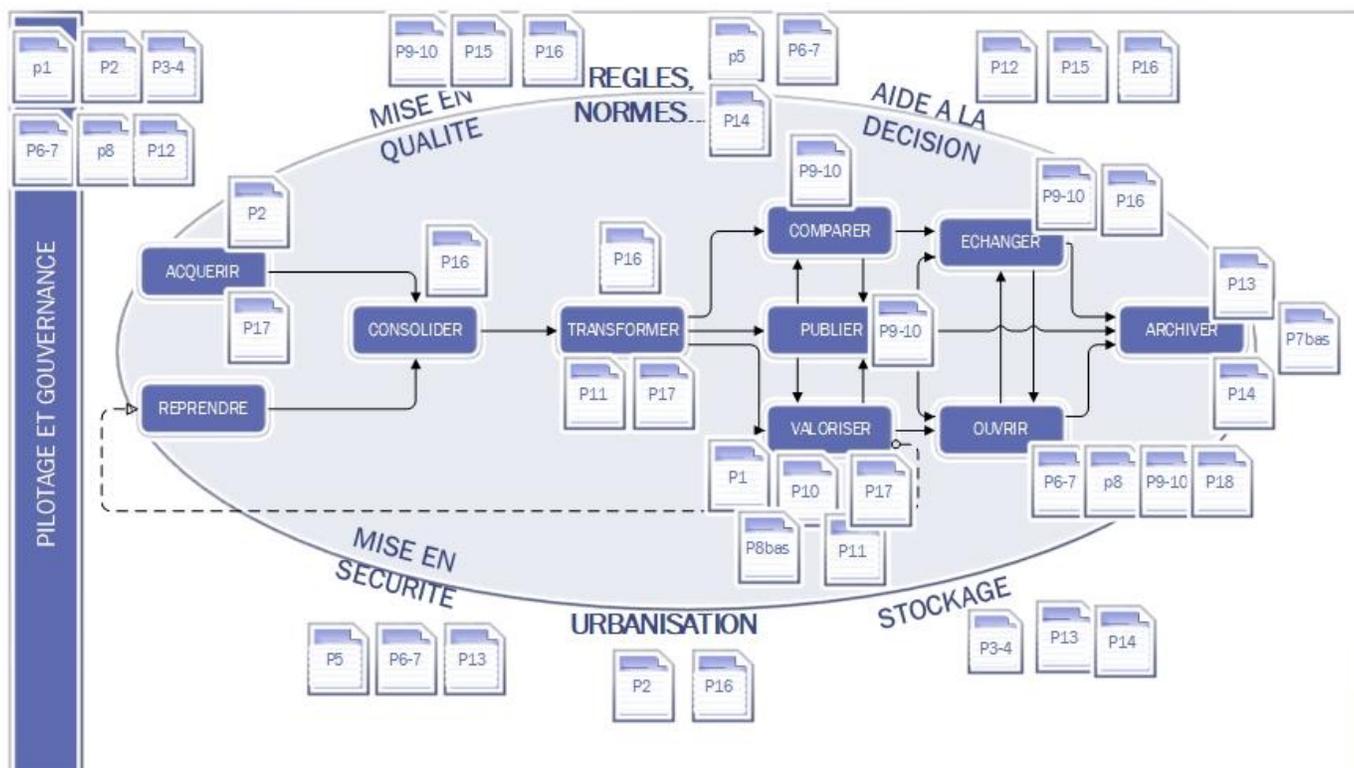
organisations : accroissement des volumes, silos (sujet plutôt organisationnel parfois mis en exergue par les données), qualité, maîtrise des données, protection des données personnelles mais aussi les mêmes succès : indicateurs de performance automatisés, publications accessibles à tous, partage du patrimoine informationnel ... Ce numéro nourrira probablement votre réflexion sur ce sujet. Peut-être une réponse dans un prochain numéro ?

Cycle de vie

Nous vous proposons de représenter le sujet de la donnée par le prisme de son cycle de vie. De son acquisition à son archivage définitif, ce schéma positionne différentes activités autour de la donnée (ex "transformer") et différents sujets associés (ex "urbanisation").

Cette "vue macroscopique du cycle de vie de la donnée" sert également de sommaire, vous y trouverez les N° de page des articles associés aux principales activités ou notions présentées. Il sera repris dans un prochain numéro de la collection numérique sur ce sujet de la donnée.

Ce schéma est soumis à vos commentaires, propositions, avis en vous adressant par mail à numerique@amue.fr.



Vers une culture de la donnée

Par Franck Estay – Conseiller DGESIP-MIPNES (MESRI)

Le développement du numérique dans l'enseignement supérieur a intégré progressivement l'exploitation et la gestion des données dans ses activités. Dans le domaine de la recherche, les exigences de la Commission Européenne dans le programme Horizon 2020 ont, par exemple, donné lieu au développement des plans de gestion de données. L'initiative de la Research Data Alliance, lancée en 2013 par la Commission Européenne, avec un ensemble de recommandations dans une approche collective, traduit aussi cet intérêt.

Avec les évolutions législatives et réglementaires dans le domaine du pilotage, le ministère de l'enseignement supérieur a façonné une partie des exigences sur les logiciels métiers. Il a souhaité exploiter des données, que ce soit dans la gestion des ressources humaines avec le suivi du plafond d'emploi ou de la gestion prévisionnelle des emplois et compétences (GPEC), dans la gestion financière et comptable avec la construction des tableaux de bords budgétaires ou la mise en place de la gestion budgétaire et comptable publique (GBCP), ou encore dans la gestion des formations avec la collecte des effectifs d'étudiants et de diplômés ou enfin avec la plateforme d'admission dans l'enseignement supérieur. Enfin, l'autonomisation progressive des établissements a amplifié les besoins de pilotage. Elle a aussi fait apparaître de nouveaux champs d'investigation, autour de la formation ou de l'immobilier par exemple, outillés via des dispositifs numériques, producteurs, eux aussi, de données.

Parallèlement, les performances des dispositifs informatiques ont considérablement progressé, aussi bien dans les capacités de stockage que dans les capacités de traitement (**infographie 1**), l'ensemble étant accompagné d'une baisse substantielle des coûts (**infographie 2**).

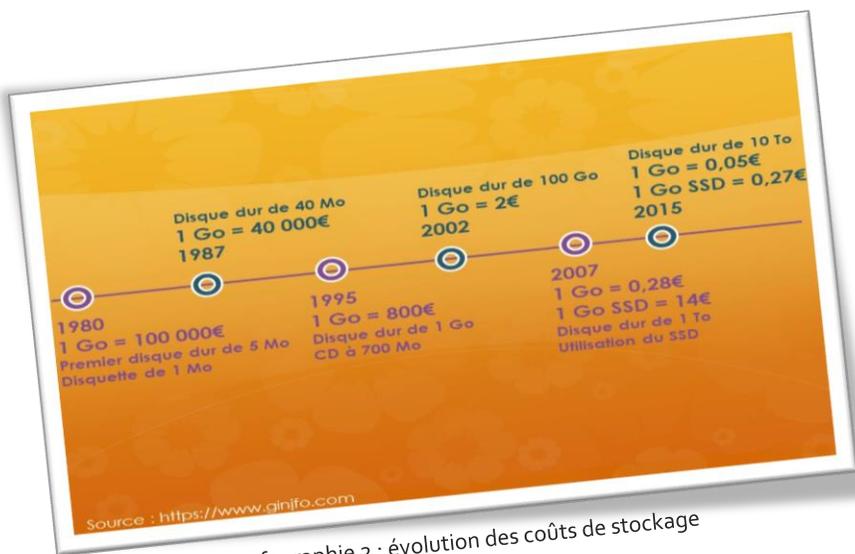


Infographie 1 : évolution des capacités de traitement

Cette accessibilité accrue à des dispositifs de stockage et de traitement a fait évoluer la consommation des ressources et a ouvert de nouvelles opportunités. La quantité et la durée de conservation des données s'en sont trouvées modifiées, la construction de modèles autour de systèmes variés (météorologiques, financiers, biologiques, technologiques) et l'étude de leur comportement ont pu intégrer des composantes plus complexes. L'accélération de l'automatisation et de la dématérialisation de processus ont aussi permis de simplifier certaines activités. La réduction des temps de traitement, l'accès à des volumes importants de données et le décloisonnement lié au web entre l'homme et le dispositif numérique utilisé ont modifié notre rapport au temps (les traitements sont plus rapides et les dispositifs numériques

accessibles de manière continue) et à l'espace (l'étendue des accès réseau permet de se déplacer avec notre environnement numérique).

Comme on vient de l'esquisser, la percolation des outils numériques et la production des données qui en découlent ont diffusé dans les différents domaines d'activités des établissements d'enseignement supérieur. Le cercle des spécialistes s'est rompu et la focale s'est déplacée de la relation homme/machine avec l'information comme support à la relation homme/données avec la machine comme support. A l'aune de ce panorama, on peut alors faire quelques observations :



Infographie 2 : évolution des coûts de stockage

- Les pratiques et les usages sont en constante évolution à travers des dispositifs sociotechniques sans que soient formalisées les trajectoires d'usage (mais il convient de souligner que les cinématiques des normes sociales et techniques ne sont pas identiques),

- La démocratisation des outils numériques amène chacun à être producteur et consommateur de données, sans que ce soit les mêmes et avec des exigences pas toujours explicites sur ces deux facettes,

- Le triptyque donnée/information/connaissance doit être porteur de sens, afin d'éviter de laisser l'individu face à ses propres représentations concernant ces notions, mais aussi pour l'accompagner à comprendre l'articulation entre ces trois notions et les dispositifs numériques qui l'environnent,

- La nécessité d'une formation organisée et systématique, dépassant la simple maîtrise technique des outils numériques (plutôt orientée usages). Cette formation devrait aborder les processus de traitement permettant de passer de la donnée à l'information puis à la connaissance, mais aussi éclairer la compréhension des modèles sous-jacents et des intentionnalités, afin de faciliter l'évolution des compétences de chacun. En effet, "la notion de compétences touche aussi bien les conceptions de l'apprentissage que celles de l'organisation du travail" (Ferrary M., Pesqueux Y, 2006),

- L'inflation des données semble stérile tant leur exploitation, quand elle existe, apparaît relever de la description de l'état d'une organisation. Une approche plus systémique et plus dynamique pourrait être envisagée, par exemple dans une logique de description puis de compréhension des transformations des organisations, mais surtout pour passer du stade de l'expérimentation au stade de l'industrialisation,

- Le glissement, qui se dessine, concernant les processus décisionnels, semble chercher à répondre à notre rationalité limitée (notre incapacité à embrasser toute la complexité d'un problème et donc à se limiter à quelques variables pour le résoudre) et à surmonter l'incertitude des effets (est-ce que la décision va bien produire les résultats escomptés?) en proposant une décision augmentée, c'est-à-dire une décision dont la construction est



Schéma sur la base de l'image de Gerd Altmann sur Pixabay

accompagnée par des algorithmes d'aide à la décision. Cette aide à la décision serait déployée à tous les échelons, avec la promesse d'une meilleure coordination des organisations (des décisions plus sûres conduiraient à des organisations mieux coordonnées dans leur fonctionnement). Ce souci d'une meilleure coordination ne devrait alors pas seulement se focaliser sur les outils numériques mais aussi appréhender la synergie que permet l'intelligence collective.

Culture de la donnée et stratégie de la donnée

L'ensemble de ces observations ne fait pas démonstration, mais suggère, avec un parti pris, qu'une culture de la donnée serait un terreau propice à l'évolution de nos organisations telles qu'elles se dessinent (c'est-à-dire avec une part croissante du numérique), et offrirait des pistes pour répondre aux questions que soulèvent les observations précédentes.

Évoquer la culture de la donnée nécessiterait d'en définir les enjeux, le périmètre et les problématiques, ce qui n'est pas l'objet de cet article. Notons que le choix du mot culture n'est pas anodin. Dans sa dimension sociologique, la culture renvoie à l'homme ancré dans son milieu. On pourrait reprendre la définition de la culture proposée par (Rocher, 1992) : " un ensemble lié de manières de penser, de sentir et d'agir plus ou moins formalisées qui, étant apprises et partagées par une pluralité de personnes, servent, d'une manière à la fois objective et symbolique, à constituer ces personnes en une collectivité particulière et distincte."

Cette culture de la donnée à construire dans sa dimension numérique, prémisses d'une connaissance éclairée, dépasserait alors le simple cadre de nos établissements, pour offrir à cette "collectivité particulière et distincte" des clés pour comprendre la transformation que le numérique induit. Elle permettrait alors d'élaborer, sur cette culture partagée, une stratégie de la donnée.

BIBLIOGRAPHIE

- Ferrary M., Pesqueux Y. (2006), Management de la connaissance, Ed. Economica, 2006, p.19,
- Henry A. (2017). Big Thick Data pour l'intelligence collective, une nouvelle voie pour les organisations de type industriel: le cas de GRDF. Colloque Communication, Organisation, Société du Savoir et Information (COSSI) 2017, May 2017, Montréal, Canada. hal-02054309,
- Kremser W., Brunauer R. (2019) Do we have a Data Culture?. In: Haber P., Lampoltshammer T., Mayr M. (eds) Data Science – Analytics and Applications. Springer Vieweg, Wiesbaden,
- Lehman, A. (2018). L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données. Revue COSSI, n°4-2018,
- Meyer P., Dardayrol JM., Trassoudaine F. (2015), rapport "Meilleures pratiques pour le big data et l'analytique dans l'administration : une nouvelle étape", (2015, p54)
- Rocher G. (1992), Introduction à la sociologie générale, Ed Hurtubise, Montréal, 1992
- Schöpfel J. (2018). Vers une culture de la donnée en SHS : Une étude à l'Université de Lille. [Rapport de recherche] Université de Lille. 2018. hal-01846849,
- Administrateur général des données (AGD). (2016-2017). La donnée comme infrastructure essentielle. Rapport annuel au Premier ministre sur la gouvernance de la donnée 2016-2017,
- Administrateur général des données (AGD). (2015). Les données au service de la transformation de l'action publique. Rapport annuel au Premier ministre sur la gouvernance de la donnée 2015.



Les DPO au service des usagers et des établissements

Par Guy Bisiaux, Rafia Berenguier, Victor Larger, Marion Lehmans et Jean-Luc Tessier, pour le réseau SupDPO

REGLES,
NORMES...

Règlement général à la protection des données (RGPD)

La fièvre de l'entrée en application du RGPD retombe. Le grand public est plus sensible à la protection de sa vie privée (voir [baromètre](#)) (+30% de plaintes auprès de la CNIL (voir [article](#)), plusieurs actions collectives,...), le droit européen et national sur la protection des données est stabilisé, les autorités de contrôle affirment les nouvelles règles du jeu à coups de sanctions (voir [site](#) de suivi) , priorisent l'accompagnement sectoriel, soutiennent le rythme de publication des [guidelines](#) et affinent leur doctrine. Les experts, quant à eux, ont délaissé les colloques de sensibilisation sur les grands principes pour discuter de sujets de plus en plus pointus, prospectifs voire géopolitiques. Et les débats juridiques se tournent désormais volontiers vers l'innovation et l'anticipation.

La mise en conformité RGPD se poursuit, avec un enjeu considérable, à la mesure de la gigantesque masse de données qui est traitée dans l'enseignement supérieur : celle de 2,7 millions d'étudiants ([source](#) ESR en chiffres) et de plus de 150 000 personnels, enseignants et administratifs ([source](#) Repères et références statistiques).

Data Protection Officer (DPO)

Les établissements ont désigné leur délégué à la protection des données (communément appelé DPO - Data Protection Officer- ou DPD), nouvelle fonction pour piloter cette "conformité RGPD". Véritable chef d'orchestre de l'application du Règlement dans les organismes, il informe et conseille chacun des membres de la communauté mettant en œuvre des traitements de données personnelles, qu'il soit personnel de l'établissement, administratif, enseignant, chercheur, ou bien étudiant.

Le DPO aide à construire par défaut et dès la conception des services numériques protégés à tous les niveaux (sécurité informatique, juridique et physique des données). Au quotidien, il :

- veille à ce que les formalités et les mesures de sécurité juridique soient remplies (e.g. dans les marchés publics ou l'information des personnes);

- analyse les processus de gestion des données, renforce les procédures techniques et organisationnelles et impulse les bonnes pratiques ;
- est chargé de contrôler le bon respect du RGPD dans l'établissement.

Le DPO promeut surtout une stratégie sur les données par l'amélioration des processus de gestion auprès de la gouvernance, relayant les recommandations des autorités concernées, accompagnant les audits de sous-traitance et sollicitant les services compétents pour faciliter l'exercice des droits des personnes.

Dynamique vertueuse

Les procédures de conformité -parfois incomprises en interne- créent ainsi une dynamique vertueuse : gage de transparence vis-à-vis des intéressés, elles permettent par exemple la minimisation du stockage des données dans le temps (dont l'impact économique et écologique pourrait être étudié), la diminution des risques d'atteinte à l'image institutionnelle et la clarification des responsabilités -entre l'institution et ses partenaires sous-traitants, personnels administratifs, académiques et scientifiques, étudiants-.

Aussi, le fait d'intégrer la protection des données dès la phase d'élaboration des projets est garant d'une bonne démarche méthodologique à même de produire une gestion des données de qualité.

C'est donc avec deux ambitions naturellement compatibles que les DPO assurent leurs missions : garantir prioritairement la vie privée des usagers dont ils traitent les données, tout en sécurisant l'établissement dans l'exercice de ses missions.

Et c'est là toute la difficulté de la mission : dans un secteur qui est innovant par nature et collecte des données pour exercer ses missions (plus de 280 000 chercheurs ([source](#) Etat de l'Enseignement Supérieur et Recherche 2019) , être la personne référente, souvent unique par ses compétences -voire isolée- et ne disposant parfois que d'un temps partiel dans une structure structurellement complexe, du fait de son organisation et ses usages, est un défi de tous les instants.

LE RESAU SUPDPO

<https://reseau.supdpo.fr/>



SupDpo

Dans ce contexte, le réseau SupDPO a été créé pour répondre à un besoin d'accompagnement et de mutualisation "entre pairs" du secteur à part de l'Enseignement supérieur, de la Recherche et de l'Innovation ("ESRI").

Il permet de se nourrir des bonnes pratiques mises en place, de mutualiser l'instruction de dossiers communs et, le cas échéant, de peser dans une négociation face à un même interlocuteur et de concevoir des recommandations sectorielles. Le réseau SupDPO a ceci de tout à fait exceptionnel : il est le regroupement de profils très variés à l'image des origines professionnelles diverses de ses membres : direction générale, informatique, juridique, conformité, documentation, archives,...

Reconnu "tête de réseau" par la CNIL, SupDPO regroupe les structures de l'ESRI (universités, grandes écoles, organismes de recherche, GIP (Groupement d'Intérêt Public) ...) et fédère uniquement les DPO "internes" désignés auprès de la Commission.

Dans un monde de plus en plus organisé par la maîtrise des données, il est à parier que les établissements relevant le défi de la protection des données seront les mieux armés dans la compétition internationale universitaire et sauront en tirer parti rapidement. Le temps utile qu'ils consacrent à cette transformation est fondamental.

Les établissements doivent faire écho aux ambitions européennes (déclinaisons des valeurs fondamentales de l'Union européenne, ouverture des données de la recherche, protection de la vie privée,...). SupDPO continuera de les accompagner en ce sens.



RGPD – open data – innovation : quelles interrelations ?



Par Anne Vanet Vice-présidente de l'association des VP-Num de l'Enseignement Supérieur et Recherche, Pr de biologie moléculaire et computationnelle à Université de Paris



Les participants à la table ronde

Bordeaux 21 et 22 Novembre 2019

Le BootCamp des VP-Num a souhaité cette année, mettre en perspective les interfaces entre open science – RGPD et valorisation autour d'intervenants professionnels (De gauche à droite : François Pellegrini, Pr Université de Bordeaux, membre CNIL, David Bessot, associé, cabinet Inphotep, Ekateritina Kirakossian, juriste, SAAT Aquitaine, Kenneth Maussang, MdC Université de Montpellier, membre du CoSO, David Rongeat, Responsable numérique, Amue). En effet, de nouvelles lois et règlements voient le jour : la [loi](#) pour une république numérique avec notamment l'obligation de l'ouverture de données (Open Data), [le plan S](#) et le règlement général sur la protection des données ou [RGPD](#). La valorisation de la recherche bien que faisant partie de la seconde mission des enseignant.es chercheur.es depuis 1984 est peu pratiquée sur le terrain.

Injonctions contradictoires

Ces sujets sont nouveaux pour les chercheur.es qui les voient comme autant de nouvelles tâches qui s'ajoutent à leur "todo" liste déjà longue. Pris ensemble ces sujets créent parfois des injonctions contradictoires puisque par exemple, si la valorisation par la communication des résultats de recherche participe bien à l'ouverture de la science, la valorisation financière via oblige à ne pas divulguer ces résultats. Des recherches en SHS utilisent souvent des données personnelles qu'il sera nécessaire

d'anonymiser ce qui parfois est un tel casse-tête que science ouverte et recherche deviennent orthogonales. La valorisation financière par dépôt de brevet oblige à ne pas ouvrir les données, en parallèle certains pensent que la science ouverte pourrait aussi générer de la valeur à travers l'open innovation. L'avènement des big data ne simplifie pas le problème : certaines données anonymes peuvent rapidement être désanonymisées lorsqu'elles sont suffisamment nombreuses. Par exemple connaître vos données de déplacement (de votre adresse personnelle à votre lieu de travail) permet la plupart du temps de vous identifier.

L'ouverture des données vers les citoyens, les chercheur.es, les étudiants ou l'industrie ne recouvre pas les mêmes tâches. En effet le citoyen souhaitera avoir une information synthétisée, alors que les chercheur.es souhaiteront avoir accès à la totalité des données (raw data), les étudiants des données recentrées sur le sujet étudié, et l'industriel, les données pouvant être valorisées. En outre il est parfois compliqué de savoir si les données que vous venez de générer doivent être ouvertes ou bien brevetées, de connaître leur propriétaire, alors que l'information est un bien immatériel qui n'a donc pas de propriétaire, de savoir si, en la croisant avec d'autres données, elle générera une nouvelle donnée brisant les bases du RGPD.

Bien qu'appliquer ces lois semble chronophage, il est évident que la non-conformité a aussi un coût. Alors comment faire pour tenter de s'y retrouver, pour être au

plus proche de la loi et peut-être même, y trouver un intérêt personnel ?

Quelques pistes

Le RGPD permet de compléter la loi "Informatique et Liberté" datant de 1978, ces notions ne sont donc pas nouvelles et peuvent plus aisément être appliquées si l'on fait preuve, à en croire les intervenants, de bon sens et du sens des responsabilités. Dans un monde se structurant autour des données, il est désormais important de se poser la question de leur devenir dès le début d'un projet, en tentant d'anticiper le risque au moyen par exemple d'une analyse d'impact relative à la protection des données (AIPD). Lorsque l'on collecte des données, il faut se poser la question de leurs différentes finalités et tenter de trouver une base légale pour chacune. Une fois le choix fait de l'ouverture ou non, de son anonymisation ou non, une licence ouverte du type "[creative common](#)" par exemple vous permet de partager ces données selon les critères que vous souhaitez. Il s'agit finalement de définir une trajectoire de mise en conformité et tenter d'être au plus près de la loi en utilisant en premier lieu le bon sens.

Nouveau chemin

Nous ne sommes qu'au début du chemin, il va falloir ensemble s'acculturer à ces nouveaux principes et recréer un sentiment de responsabilité. N'ayant pas dans nos établissements le personnel pour traiter toutes les données tant administratives que de formation et de recherche, il est important de se doter d'outils qui puissent permettre de



Un plan de gestion des données pour la recherche

Retour sur le modèle de plan de gestion des données de l'Agence Nationale de la Recherche – DSPSI, Amue

Dans le cadre de sa politique science ouverte, l'Agence nationale de la recherche (ANR) demande l'élaboration d'un Plan de Gestion des Données (PGD) pour les projets financés à partir de 2019.

Anticiper pour mieux partager

Ce plan de gestion des données issues de la recherche est un document qui vise "à *décrire la façon dont les données de recherche seront produites, collectées, réutilisées, stockées, protégées, disséminées ou encore conservées à long terme*".

Ce plan réalisé en concertation avec l'Institut de l'Information Scientifique et Technique (INIST) suit les recommandations du Comité pour la Science Ouverte.

Un modèle comme guide

Ce plan est présenté comme un document modèle, à suivre et à compléter dans les six mois qui suivent le démarrage scientifique du

projet. Il propose de porter l'attention sur six points de la donnée pour la recherche, qui correspondent au cycle de vie, et à la stratégie que l'on souhaite autour des projets lauréats.

1. description des données et collecte ou réutilisation de données existantes,
2. documentation et qualité des données,
3. stockage et sauvegarde pendant le processus de recherche,
4. exigences légales et éthiques, codes de conduite,
5. partage des données et conservation à long terme,
6. responsabilités et ressources en matière de gestion des données

EN SAVOIR PLUS

Détail sur les recommandations : ANR. (2019). L'ANR met en place un plan de gestion des données pour les projets financés dès 2019. Consulté le 18 novembre 2019, à l'adresse Agence nationale de la recherche website: <https://anr.fr/fr/actualites-de-lanr/details/news/lanr-met-en-place-un-plan-de-gestion-des-donnees-pour-les-projets-finances-des-2019/>

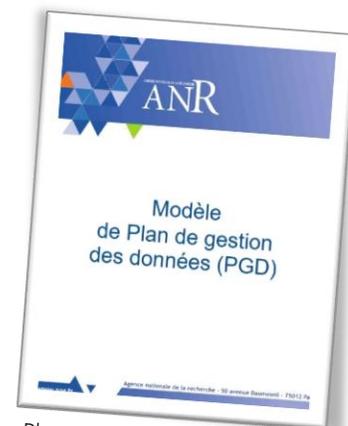


Photo ANR

A l'étranger, des Chief Data Officer (CDO) d'université

Retour sur un article paru au sein d'Educause – DSPSI, Amue.

VALORISER



Université de Caroline du Sud - CC BY-SA 3.0

Le chief data officer, ou directeur des données, est un nouveau métier de la donnée. C'est un cadre dirigeant qui participe au pilotage de la stratégie globale de la donnée de l'organisation : un administrateur

général des données en quelque sorte.

Dans un article d'EDUCAUSE, Mike Kelly, Chief Data Officer de l'université de Caroline du Sud, témoigne de sa nouvelle fonction.

Le rôle du CDO répond "à deux besoins urgents sur le campus :

- *diriger les efforts d'administration des données,*
- *renforcer la capacité d'analyse afin de prendre des décisions en fonction des données."*

Comme pour tous les actifs précieux, les données nécessitent une gestion active, intentionnelle et cohérente, mieux connue sous le nom de gouvernance des données.

Le CDO peut et doit évangéliser la valeur des données pour l'analyse et l'aide à la décision, tandis que la gouvernance des données garantit la disponibilité des meilleures données pour répondre aux besoins tactiques et aux objectifs stratégiques d'une institution.

EN SAVOIR PLUS

L'article en ligne

Kelly, M. (2015). The Chief Data Officer in Higher Education. Consulté le 18 novembre 2019, à l'adresse <https://er.educause.edu/articles/2015/6/the-chief-data-officer-in-higher-education>



Journée Open Data et suite des travaux Amue

Résumé par David Rongeat, DSPSI, Amue

La 1^{ère} étape de sensibilisation au sujet de l'Open Data a été réalisée par la publication du [numéro](#) intitulé "Open Data et ESR, opportunité de créer de nouveaux services" de la collection numérique daté de septembre 2019. Au même moment, l'Amue, en tant qu'opérateur public, initiait sa propre démarche OpenData avec ses 1^{ers} jeux de données publiés sur data.gouv.fr et catalogués sur une page dédiée data.amue.fr.

La journée OpenData organisée le 14 novembre par l'Amue à la maison des universités avait un double objectif. Le matin consistait à poursuivre, sous une autre forme, la sensibilisation et l'acculturation à l'Open Data à destination des établissements. L'après-midi des ateliers en parallèle ont été organisés pour travailler sur des outils et méthodes permettant l'identification de jeux de données à ouvrir dans le contexte de l'Enseignement Supérieur et Recherche.

Découvrez ou revivez le séminaire

Une matinée organisée en 2 séquences : tout d'abord une vue d'ensemble sur l'Open Data portée par [Etablab](#) en charge de data.gouv.fr et de la politique d'état sur l'OpenData, la coopérative [Dataactivist](#) et l'Amue sur les notions clés à retenir, l'association [OpenDataFrance](#) au sujet de l'Open Data dans les territoires et un témoignage de ré-utilisateur de données ouvertes dans le domaine des transports par [Kisio Digital](#). La seconde partie portait sur un focus Enseignement Supérieur et Recherche : intervention de Marin Dacos, conseiller du Directeur Général de la Recherche et de l'Innovation au MESRI pour [la science ouverte](#), du SIES porteur de [#dataESR](#) au MESRI, une intervention Amue sur la qualité de données et enfin un retour d'expérience de l'université d'Angers sur l'intégration du sujet de la Donnée dans son Schéma Directeur du Numérique. Que les intervenants et participants soient remerciés.

Vous pouvez trouver les liens vers les supports de présentation, visionner la matinée en vidéo et (re)lire le LiveTweet [#OpenDataAmue](#) sur la page Open Data de l'Amue : data.amue.fr.

Quelques points clés

De manière non exhaustive, retenons ces quelques éléments principaux. Un positionnement dual, producteur d'Open Data mais aussi consommateur de données ouvertes par d'autres est mis en avant par plusieurs intervenants, en positionnant également son établissement comme consommateur de ses propres données ouvertes. Inscrire une démarche Open Data dans un plan stratégique sur les données ou dans un Schéma Directeur du Numérique (SDN) peut être un élément pertinent pour la soutenir et faciliter l'implication de porteurs politiques ou métiers dans celle-ci. Le forum [#TeamOpenData](#) est un bon moyen de résoudre les questions que vous vous posez sur ce sujet. Afin de hiérarchiser les jeux de données à ouvrir, il est proposé de les prioriser selon leurs usages potentiels et leur capacité à produire de nouveaux services pour les usagers ou de la valeur pour les consommateurs. C'est cette qualification qui a été menée dans les ateliers de l'après-midi.

Pour ce qui concerne l'Open Science, sujet important autour des données ouvertes dans l'Enseignement Supérieur et la Recherche, deux recommandations de lecture dans l'encart ci-après.

PUBLIER

OUVRIR

DEUX LECTURES COMPLEMENTAIRES SUR L'OPEN SCIENCE

Au sein de ouvrirlascience.fr] le Cadre juridique pour l'ouverture des données de recherche : [ici](#)

Le dossier de la [revue](#) politique et parlementaire intitulé « Des nains sur les épaules de géants : ouvrir la science en France » : l'article [ici](#)

Ateliers de travail

L'objectif premier de ces ateliers était d'éprouver une démarche et des outils permettant aux établissements de construire une feuille de route sur l'Open Data et répondre à la question "par quels jeux de données commencer ?"

Sous forme de sessions parallèles, ces ateliers étaient organisés en deux temps. Le premier temps d'idéation consistait à lister, sans réserve, des jeux de données susceptibles d'être ouverts et décrire des utilisations potentielles. Cette liste était complétée par l'identification de jeux de données ouverts par d'autres et intéressants pour l'établissement.

Dans un second temps, ces différents jeux de données ou objets métiers ont été passés au crible d'une analyse permettant d'évaluer différents critères tels que le délai de mise à



Journée Open Data 14 nov.2019 Photo Amue



disposition, effort de mise en qualité, potentiel de création de service, Ainsi évalués, ces jeux de données sont positionnés dans une matrice sur deux dimensions : faisabilité et valeur d'usages.

Enfin, les participants ont travaillé sur une matrice de décision permettant d'identifier des freins et des moteurs à l'ouverture des données au sein d'un établissement.

Vous pouvez découvrir la synthèse de ces travaux sur la page data.amue.fr.



Matrice de prise de décision

Suite des travaux

Les outils et résultats de ces ateliers du 14 novembre 2019 seront réutilisés dans la suite des travaux mutualisés sur l'OpenData.

Un appel à participation à un groupe de travail "Open Data", animé par l'Amue, est toujours d'actualité. Ce groupe vise à mettre en musique la suite des travaux pour les établissements : poursuivre la sensibilisation et les retours d'expériences des établissements, construire opérationnellement des schémas de données ouvertes par l'Enseignement Supérieur et Recherche, travailler sur des outils méthodologiques, des formations pour les établissements et mutualiser des solutions.

Ce travail ne peut se faire qu'avec la contribution active des établissements pour co-construire un socle commun OpenData et en ainsi en optimiser la mise en œuvre. Pour participer à ce groupe de travail, contactez l'équipe numérique de l'Amue : numerique@amue.fr

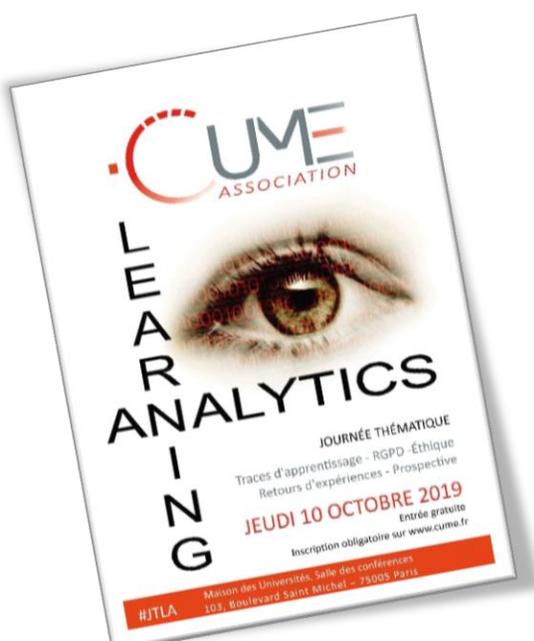
NOUVELLES PUBLICATIONS OPEN DATA

Publié après le bouclage du numéro sur l'Open Data, ce guide Open Data de la Cnil en collaboration avec la Cada : à lire [ici](#) "Publication en ligne et réutilisation des données publiques" ("open data")



Retour sur la journée d'études du CUME

Retour sur la journée d'études par le bureau de l'association CUME



La Journée Thématique du CUME sur les Learning Analytics a eu lieu le 10 octobre 2019, à la Maison des Universités à Paris et portait sur l'analyse des traces d'apprentissage ou "Learning Analytics".

Une dizaine d'interventions se sont succédées mêlant des témoignages de l'état de la recherche par des laboratoires français (Université de Lorraine/ LORIA, Sorbonne Université/ CAPSULE, IUT Laval/Labo.Info.de l'Université du Mans (LIUM), MICA Université Bordeaux Montaigne) une intervention du Cabinet INFHOTEP sur le RGPD mais

aussi des retours d'expériences en cours dans les universités (Paris Saclay/UVSQ, Université du Mans, EOLE de l'Université de Lorraine, Université Rennes 1).

L'ensemble de la journée thématique a été captée et pourra être visionnée ainsi que les supports des interventions en consultant le site <http://www.cume.fr/>

EN SAVOIR PLUS

Le CUME est une association nationale ayant pour objectif de mettre le numérique et l'informatique au service de l'enseignement dans les établissements d'enseignement supérieur et de recherche.

Retrouvez les informations sur cette journée thématique et les résumés sur cette [page](#)



Les *learning analytics* en quelques mots

Par Soufiane Rouissi, MCF, Laboratoire Mica (EA 4426), Université Bordeaux Montaigne

Depuis le début du XXI^e siècle les établissements d'enseignement supérieur ont investi dans l'implémentation et la gestion de plateformes d'apprentissage en ligne, des LMS (Learning Management System) dont Moodle, plateforme diffusée en mode open source, serait la plus utilisée, aussi bien en France qu'un peu partout dans le monde.

Le contexte des apprentissages mobilisant les technologies numériques change

Ce recours généralisé contribue largement à produire de très nombreuses données d'usage : accès aux cours en ligne, progression et avancement, suivi et achèvement d'activités, scores obtenus à des tests de connaissances en ligne, téléchargement de documents, participation aux forums de discussion, visionnage de vidéos, notes et résultats annuels. De plus, les MOOC (Massive Open Online Courses), qui ont fait leur apparition de manière importante, ont accentué la production massive de données alimentant un Big Data pour l'éducation de manière accélérée. Ainsi ce qui semble réalisable avec quelques étudiants nécessite une toute autre approche lorsque le nombre d'étudiants croît considérablement, notamment dans des contextes où la massification devient la norme.

De nouvelles interrogations apparaissent

Les questionnements nombreux pour la recherche en éducation ou en informatique avec les EIAH (Environnements Informatiques pour l'Apprentissage Humain), mais aussi dans les Sciences de l'information et de la communication et en Technologie éducative, s'emparent de ces sujets. Cela était le cas avec ce qui est dénommé "educational data mining" (EDM), et plus récemment, à travers ce qui est regroupé désormais sous l'appellation "learning analytics". A l'instar de ce que nous avons connu avec le e-learning, l'introduction de nouvelles notions et expressions non stabilisées comme celles-ci n'échappent pas à la règle, avec la tentation de proposer diverses définitions en fonction des contextes d'application et de recherche. Nous proposons de retenir, même si nous ne voulons pas les réduire à cela, que les learning analytics regroupent des méthodes et des techniques autour de la mesure, de la collecte, de l'analyse des activités menées par des usagers dans des contextes d'apprentissage, à des fins de compréhension voire d'anticipation des comportements mais aussi avec pour objectif potentiel celui de l'amélioration des méthodes d'enseignement mises en place favorisant ainsi le pilotage des dispositifs numériques éducatifs.

Nous imaginons de les appréhender selon un niveau micro à l'échelle de chaque étudiant en prenant en compte l'ensemble de ses actions menées dans l'environnement numérique, à un niveau meso qui serait celui de la formation et/ou du département dans lequel les enseignements en mode numérique sont mis en place, et enfin à un niveau macro qui serait celui de l'université, voire à l'échelle supérieure. Face aux difficultés de gérer de très grands groupes d'étudiants, dans un contexte de banalisation du recours aux outils et environnements d'apprentissage en ligne, les learning analytics n'ont pas fini de retenir notre attention et d'alimenter le débat.



Photo by [Core Analytics](#)

Exploration autour des *learning analytics*

Par Feriel Goulamhousen (Upem), Bertrand Mocquet (Amue), Soufiane Rouissi (Mica, UBM)

L'Agence de Mutualisation des Universités et Etablissements (Amue), l'Université Paris-Est Marne-la-Vallée et l'équipe de recherche MICA en Sciences de l'information et de la communication de l'Université Bordeaux Montaigne, se sont associés dans un projet de recherche dénommé "Étude exploratoire

EAGLE", pour Étude Adaptée aux Générations du e-Learning En questionnant le recours à des outils d'analyse des données produites par les activités à caractère pédagogique enregistrées dans les plateformes d'apprentissage en ligne ou LMS (Learning Management System), ce projet

de recherche vise à mieux comprendre le processus de transformation des données pédagogiques en données administratives de gestion.

Au-delà de son approche exploratoire, le projet vise à caractériser les enjeux et à identifier des pistes de réflexion ou d'instrumentation.



L'approche par la donnée et sa liaison au facteur humain.

Par Bertrand Mocquet, DSPSI, Amue

L'approche par la donnée possède quand l'on s'y penche un caractère très technique de l'activité d'une organisation universitaire, universitaire ici s'entendant comme concernant les universités et les établissements.

S'appuyant sur ce que les services échangent, produisent ou récoltent, la donnée universitaire vit avec son jeu de métadonnées et ses processus métiers associés : il en résulte parfois un agglomérat de spécificités de forme de ces données.

Le risque est alors grand de faire de la gouvernance de la donnée universitaire un objet purement technique, là où l'humain, nos usagers comme nos membres de l'université sont particulièrement concernés. Des mouvements très récents apparaissant de manière systématique sous l'appellation *user-centric* dans les systèmes d'information. Intéressons-nous ainsi aux usagers.

La donnée et les étudiants

L'Enseignement supérieur et la recherche est parfois considéré comme un marché, celui de la connaissance, l'étudiant ou le stagiaire pourrait être ainsi considéré comme un "client" d'un point de vue système d'information. La donnée liée à l'étudiant est aujourd'hui partout avec la multiplication des canaux (réseaux socio numériques, connexion wifi sur le campus, prêt de livre, traces sur la plateforme de formation...): ce sont autant de points de contacts avec lui. Tout l'enjeu serait d'obtenir une vision 360° de l'étudiant pour lui adresser le bon message au bon moment et personnaliser son expérience tout au long de son parcours de formation ou de sa vie étudiante. On voit ici aussi la précaution à prendre en termes de respect du règlement général sur la protection des données (RGPD).

La donnée et les membres de l'établissement

Avec les membres, le rapport à la donnée est double. Comme les étudiants, l'université détient des données sur chacun des membres, données professionnelles (RH,

financière,...) mais aussi à caractères personnels au sens de la CNIL. Mieux connaître ses membres en respectant le RGPD, pourrait faciliter l'expérience de chacun d'entre nous, nous évitant, par exemple, des multiples saisies lors de nos démarches administratives internes (promotion, frais de mission, déclaration de service...). On pourrait aussi permettre également la personnalisation des outils en fonction du type de statut, du contrat...

Mais la donnée pour les membres c'est aussi le respect de la qualité des données, quelle que soit notre fonction dans l'organisation. Une bonne saisie, un étudiant qui change de filière en respectant le processus, la transmission de tout changement de situation personnelle, l'information de tout changement en cours d'année universitaire sont des exemples courants, qui nous concernent potentiellement tous, et pour lesquels nous devons penser tous et toutes à fiabiliser la donnée dans le système d'information.

Au sein de la direction de l'établissement

Le pilotage par indicateurs analytiques, par exemple plafond d'emploi en RH, repose sur des seuils et sur des calculs s'appuyant sur la donnée. Dans ce cas la donnée permettrait avant tout d'anticiper les coûts, de prévoir avant de décider : c'est une aide à la décision qui deviendrait indispensable tant les universités et les établissements se complexifient.

Avec les partenaires

Le rôle crucial pris par la donnée pourrait faire changer de valeur. Cela pourrait devenir un actif stratégique en interne, mais aussi pourrait également être valorisée auprès des partenaires, ou bien ouvertes.

En guise de conclusion, nous pensons que la donnée doit être abordée comme un concept global qui touche chacun d'entre nous, les acteurs de l'Enseignement supérieur et de la recherche.

AIDE A LA
DECISION

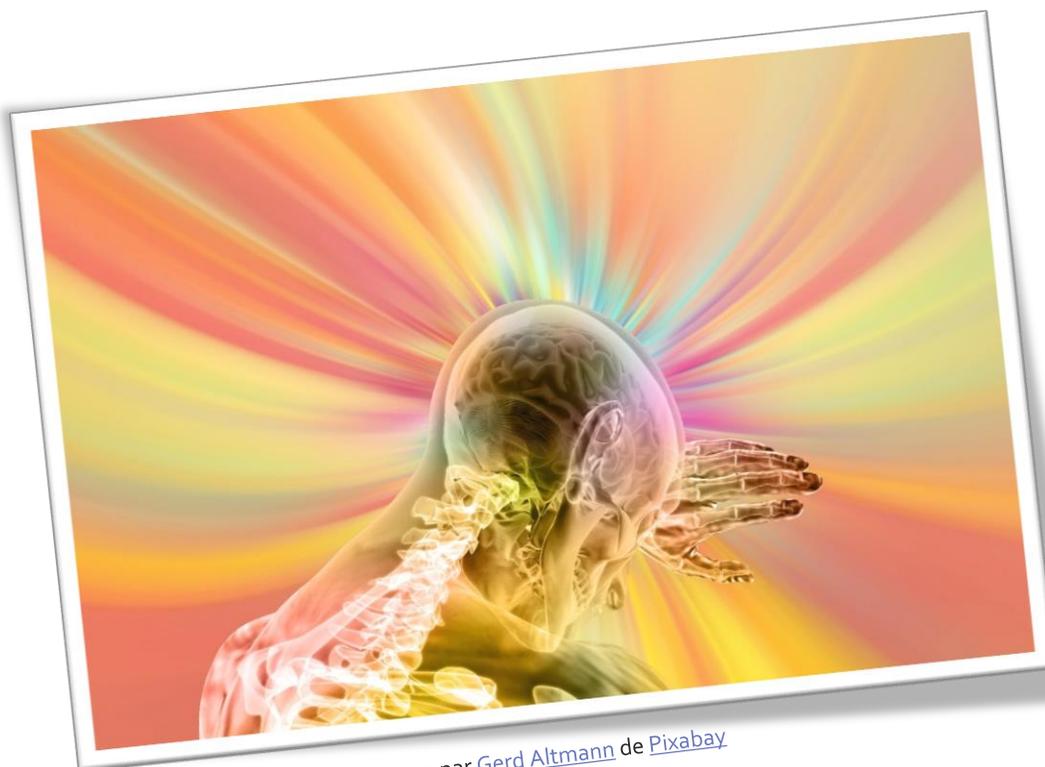


Image par [Gerd Altmann](#) de [Pixabay](#)

Comment transmettre aux futures générations la mémoire des données?

Témoignage de Nicolas Cazenave CINES - Centre Informatique National de l'Enseignement Supérieur

Aujourd'hui le véritable enjeu n'est plus la production des informations mais leur préservation dans le temps. Ce que l'on nomme le "big data" montre bien comment l'outil numérique est devenu rapidement, pour nous tous, un formidable accélérateur de production de données. Ce changement de paradigme dans nos pratiques engendre un certain nombre de problèmes auquel il convient d'apporter des solutions. Et en premier lieu, la question du stockage de ces données de plus en plus volumineuses. Ce qui nous amène à réfléchir sur le choix de garder telle ou telle information. Puis à ce qu'elles vont devenir dans le temps.

Le temps : le facteur de risque principal de perte de ces données

Un certain nombre de causes va empêcher la transmission des informations importantes aux générations futures.

Si les SAE (Systèmes d'Archivage Électronique) apportent des solutions en termes d'outils, la préservation des données numériques engage, sur le long terme, les équipes d'experts métiers, archivistes et informaticiens, ainsi que les institutions qui portent cette mission.

Depuis plusieurs années, le CINES a su évoluer et innover dans cette mission d'archivage pérenne pour le monde universitaire et la recherche française. Aujourd'hui l'implication du CINES est internationale, de par son engagement dans plusieurs projets européens faisant la part belle à cette problématique de préservation des données.

Le SAE Vitam (Valeurs Immatérielles Transmises aux Archives pour Mémoire) est un programme interministériel qui propose un socle d'archivage électronique réutilisable par l'ensemble des administrations françaises pour classer, conserver et sécuriser les documents numériques.

Un projet inter-ministères

Fruit de la collaboration des trois porteurs du projet (Ministère des Armées, Ministère de la Culture, Ministère de l'Europe et des Affaires Étrangères), ce chantier informatique est piloté par la Direction Interministérielle du

Numérique et du Système d'Information et de Communication de l'État (DINSIC aujourd'hui DINUM) et par le Comité Interministériel aux Archives de France (CIAF).

Ce SAE a été testé puis mis en production avec l'aide d'une quinzaine de partenaires conventionnés, dont le CINES (par l'intermédiaire de son directeur M. Boris Dintrans) qui les représente aujourd'hui dans le comité de direction du programme.

C'est la première fois qu'une telle initiative est lancée en France avec la volonté de créer un standard pour les SAE, et le manifeste du programme visible sur son site (www.programmevitam.fr) en témoigne : "Pour les services de l'Administration. Afin de satisfaire les enjeux d'accès dans le temps à leurs informations numériques, Vitam propose une solution logicielle libre d'archivage, évolutive, simple et facilement interfaçable, qui permet la gestion unitaire et sécurisée de milliards d'objets et vise son adoption par le plus grand nombre d'acteurs publics".

La solution logicielle open source Vitam utilise les technologies du big data, seules à même de relever le défi de l'archivage du nombre d'objets numériques qui seront produits ces prochaines années par les administrations de l'État.

Une structuration européenne

Mais depuis plusieurs années, la recherche scientifique s'oriente vers une structuration des communautés scientifiques à l'échelle européenne. Cette évolution facilite le partage des données tout en permettant des économies d'échelle par la mutualisation des outils et des compétences notamment. Dans cette optique, penser l'archivage électronique non seulement à l'échelle nationale mais aussi au niveau européen devient non seulement envisageable mais de plus en plus réalisable.

C'est pourquoi le CINES est engagé dans plusieurs projets européens qui montrent la nécessité de préserver la mémoire des données à l'échelle d'un big data international. Nous participons à créer avec les communautés scientifiques une infrastructure de données fédérée qui pourra gérer l'ensemble du cycle de vie des données,

MISE EN
SECURITE

ARCHIVER



Robot, Photo CINES

de la production à l'archivage pérenne à long terme.

Par exemple, le projet ICEDIG a permis de prouver la capacité de traiter, préserver, et mettre à disposition des millions d'images numérisées d'herbiers issus des collections de différents muséums d'histoire naturelle.

L'Europe a souhaité offrir un portail EOSC-HUB permettant de rassembler dans un grand catalogue tous les services offerts aux infrastructures de recherche et aux chercheurs. La problématique spécifique de la préservation et la réutilisation des données produites est évidemment au cœur des préoccupations européennes.

Le récent concept FAIR (Findable, Accessible, Interoperable and Reusable) est apparu pour concrétiser cette volonté. Il existe une quinzaine de principes FAIR, permettant de s'assurer que les données puissent être plus tard, "trouvables, accessibles, interopérables et réutilisables" dans l'intérêt de la science.

Nous participons à EOSC (European Open Science Cloud) en offrant un service de préservation à l'échelle européenne appelé eTDR (pour Européan Trusted Digital Repository) en partenariat avec DANS (Data Archiving and Networked Services) aux pays Bas.

Bien sûr, en tant qu'institution nationale, le CINES a un rôle important à jouer pour apporter ces services aux chercheurs, mais prévoir dès le début quelles seront les modalités d'accès, de préservation et de collaborations futures autour des données produites revient naturellement au porteur de projet de recherche. De plus cela lui permet de mieux gérer ces aspects tout au long d'un projet de recherche et donc d'en maîtriser les coûts.



Archives numériques : à la confluence de deux métiers

par Lucile Schirr, présidente de la section Aurore, Chargée d'archives à l'Université de Strasbourg



Les premières prises en charge d'archives en université remontent à une vingtaine d'années, avec la création des premiers services d'archives, puis les parutions des circulaires du 2 novembre 2001⁽¹⁾ et instruction de 2005⁽²⁾.

Le domaine de l'Enseignement supérieur et de la Recherche connaît un véritable "Baby-boom archivistique"⁽³⁾ dans les années 2010. Toutefois certaines universités demeurent dépourvues de services d'archives.

Les services d'archives universitaires assurent la gestion des archives administratives, pédagogiques et de la Recherche.

La section Aurore, section des archivistes des universités, rectorats et organismes de recherche

La section Aurore, créée au sein de l'Association des archivistes français en 2011, est la traduction formelle d'un réseau né en 2007, en réponse à la professionnalisation croissante de la fonction archives dans les établissements d'enseignement supérieur.

Si la section est avant tout un espace d'échanges privilégié pour les personnels en charge des archives, elle est également organisée en groupes de travail et projets ayant vocation à accompagner les organismes producteurs d'archives sur toute question archivistique de gestion et de conservation des documents ou des données⁽⁴⁾.

Elle réunit ses membres une fois par an à l'occasion de ses journées d'études thématiques.

La section Aurore et l'AMUE ont déjà collaboré par le passé en réalisant un guide sur *La gestion des archives au sein d'un établissement d'enseignement supérieur et de recherche*⁽⁵⁾ paru en 2010.

LIEN POUR RETROUVER LA SECTION AURORE

Site de l'Association des Archivistes Français
<https://www.archivistes.org/>



L'archivage du numérique dans les établissements d'enseignement supérieur et de recherche

La révolution du numérique a bouleversé les pratiques professionnelles.

Les archivistes, déjà confrontés à la gestion d'importants arriérés de documents papier repensent et adaptent leurs pratiques pour la prise en charge massive de ces "nouvelles archives"⁽⁶⁾, qu'elles soient natives, structurées dans des logiciels ou applications de gestion, livres (mails, documents bureautiques, pages web, contribution aux réseaux sociaux...) ou issues d'une dématérialisation duplicative.

Dans un contexte de dématérialisation croissante des processus, les archives numériques sont soumises aux mêmes durées de conservation (jusqu'à 50 ans pour les récapitulatifs de scolarité) et sorts finaux (tout n'est pas à conserver) que les archives papier au regard de leur valeur probante ou informationnelle.

L'apport d'un archiviste à un projet d'archivage numérique est d'abord son expertise sur la sélection des données, leur "bonne" conservation dans le temps, leur transfert vers des services d'archives définitives⁽⁷⁾ ou leur destruction finale. Il en connaît les contraintes réglementaires et fonctionnelles (la mise en œuvre du SEDA⁽⁸⁾), peut évaluer des options et orienter les choix concernant les solutions techniques : acquisition d'un SAE⁽⁹⁾, plateformes mutualisées, tiers-archivistes.

Une démarche d'archivage électronique est "l'ensemble des actions, outils et méthodes mis en œuvre pour conserver à moyen et à long terme des informations numériques dans le but de les rendre accessibles et exploitables⁽¹⁰⁾".

Cette démarche induit une collaboration accrue et l'élaboration d'un langage commun entre archivistes et informaticiens, puisque la méthodologie et la démarche archivistique sont aussi importantes que la réponse technologique. Les processus d'archivage électronique doivent répondre à des enjeux :

- archivistiques : conserver la provenance, ainsi que le contexte de production des données, tout en prévoyant la gestion de leur cycle de vie et leur accessibilité grâce à des jeux de métadonnées descriptives...
- technologiques : garantir la conservation des données en tenant compte des contraintes liées à la pérennité des formats et aux obsolescences physiques, technologiques et logicielles des supports d'origines...
- de sécurité : les données doivent être dupliquées sur des sites distants et protégées des attaques informatiques...

Afin de relever au mieux ces défis en lien avec les directions des systèmes d'information, l'archiviste doit être intégré le plus en amont possible dans les processus informatiques et bénéficier d'une visibilité sur l'ensemble du cycle de vie de l'information.

POUR PLUS D'INFORMATIONS

(1) - [Circulaire](#) du Premier ministre du 2 novembre 2001 relative à la gestion des archives dans les services et établissements publics de l'État.

(2) - [Instruction ministérielle DAF/RES/DPACI/2005/003 du 22 février 2005](#), relative au tri et à la conservation des archives reçues et produites par les services et établissements concourant à l'éducation nationale, dont les universités.

(3) - Le Brech Goulven. [Nous ne sommes plus des "fous furieux"](#). In: La Gazette des archives, n°231, 2013-3. Les archives des établissements d'enseignement supérieur et de recherche. pp. 7-10;

(4) - Depuis la loi du 7 juillet 2016, le terme « donnée » est intégré dans la définition des archives. "Les archives sont l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité." (L211-1 du Code du Patrimoine).

(5) [Dossier](#) sur le site de l'Amue

(6) Bécard Lorène. [L'archivage électronique dans l'enseignement supérieur et la recherche : problématique et solutions](#). In: La Gazette des archives, n°231, 2013-3. Les archives des établissements d'enseignement supérieur et de recherche. pp.281-291 ;

(7) Archives nationales pour les établissements parisiens, ou Archives départementales pour les établissements en région.

(8) [Standard d'échanges de données pour l'archivage](#) - 2006 collaboration SIAF - DGME. Il permet de modéliser les transactions entre les acteurs de l'archivage en vue d'une interopérabilité.

(9) SAE : Système d'archivage électronique

(10) Comité interministériel aux archives de France – [Référentiel général de gestion des archives](#) – octobre 2013



La donnée au cœur du processus de décision

AIDE A LA
DECISIONMISE EN
QUALITE

par Clotilde Tonnerre, Responsable Cellule d'aide au pilotage, Université de Nantes

La donnée outil indispensable d'appui à la décision ...

L'enjeu de l'information est un élément essentiel pour piloter un établissement d'enseignement supérieur aujourd'hui et l'université de Nantes place, depuis plusieurs années, la donnée comme un outil indispensable d'appui à la décision.

Depuis 2012, la Présidence anime un dialogue autour de la stratégie et de la gestion avec ses composantes. Cela se concrétise par un contrat pluriannuel d'objectifs et de moyens. Des temps d'échanges donnent ainsi lieu annuellement à des échanges stratégiques, autour des grands projets de l'Etablissement, ainsi que des négociations autour de moyens financiers et humains récurrents et d'enveloppes sur projets. Ces échanges sont nourris par des indicateurs d'aide au pilotage, issus des outils de gestion de l'Etablissement. Ces indicateurs sont le fruit d'un travail de concertation entre les équipes politiques au niveau central et au niveau proximal de l'Université. Des temps dédiés sont organisés entre les Directions de pôle et de composante et l'équipe politique de la Présidence, pour partager une analyse conjointe de la situation des structures, autour de ces données d'activités et de gestion. Les équipes observent et discutent particulièrement de l'évolution de certains indicateurs : le taux de mobilité internationale étudiante entrante et sortante, le taux de réussite en licence, le taux d'encadrement, le développement de ressources propres...

La donnée vient donc nourrir et appuyer la négociation, elle participe à l'objectivation des arbitrages finaux.

D'un point de vue opérationnel, la démarche est animée par la [Cellule d'Aide au Pilotage](#), service rattaché à la Direction générale. La CAP définit les indicateurs prioritaires avec les Directions métiers. S'en suit un long travail de caractérisation, de collecte, de mise en forme, puis d'analyse de la donnée.

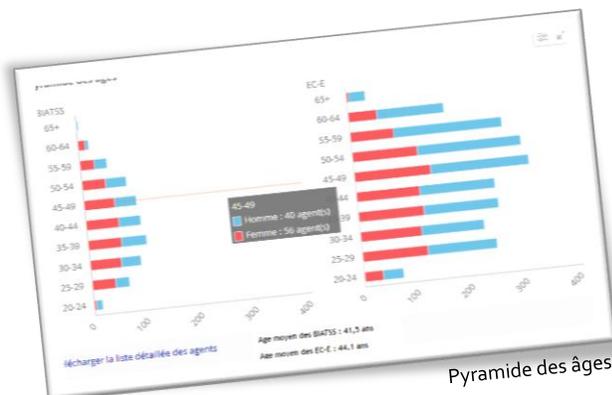
OUTILS DE PILOTAGE

UNDATA

LES DONNÉES POUR PILOTER

... au service d'un projet de gouvernance ambitieux

L'Etablissement souhaite aujourd'hui mener un projet de gouvernance ambitieux autour du regroupement de composantes et d'unités de recherche en pôles, pour d'avantage de



lisibilité et de fluidité de fonctionnement. A l'avenir, le dialogue autour des objectifs stratégiques et de gestion s'opérera entre la Présidence et les pôles, et à l'échelle de chaque pôle, entre structures d'enseignement et de recherche. Certains moyens seront délégués, et les processus associés déconcentrés.

L'accès à la donnée à des fins de pilotage est au cœur de cette nouvelle organisation émergente. Pour répondre à cet enjeu, il est rapidement apparu nécessaire de mieux centraliser les informations pour mieux les partager. C'est de cette manière que notre projet de Système d'Information Décisionnel s'est construit, autour d'une équipe projet réunissant la Cellule d'Aide au Pilotage et la DSIN, fortement soutenue par la Direction Générale des Services, et de choix techniques structurants pour l'Etablissement.

Une démarche de priorisation des données indispensable a été conduite dans l'Etablissement. Sur le plan fonctionnel, le projet a pu capitaliser sur l'existant, notamment sur un travail très fin autour des besoins en outils de pilotage et de la définition et la caractérisation d'indicateurs-clés, et un réseau d'interlocuteurs très impliqués dans les Directions –métiers (DAF, DRH, DEVU). Cela fait partie des pré-requis indispensables. Nous avons conçu un système évolutif, comprenant un extracteur de données, un entrepôt et des

outils de restitution, permettant d'intégrer facilement d'autres domaines.

Système d'Information Décisionnel (SID)

Notre projet s'est également nourri des initiatives d'autres établissements engagés dans la démarche de SID communautaire animée par l'Amue, ce qui a permis de mettre en production nos outils dans des délais très ambitieux, et de contribuer aux démarches nationales en retour.

Nous venons de lancer les premiers tableaux de bords, sur les volets RH, Finances et Formation. Nous constatons déjà que la donnée immédiatement disponible et mise à jour quotidiennement sera un vecteur de transformation de nos pratiques.

L'Etablissement doit maintenant accompagner ces évolutions et, dès lors, nous posons d'autres questions : comment faire en sorte, sur le plan stratégique, que les fonctions dirigeantes s'acculturent à la donnée, que celle-ci vienne booster les capacités décisionnelles à bon escient ? Au niveau des services, comment restituer la donnée de la manière la plus parlante possible, comment l'analyser de manière la plus objective possible ? Malgré un accès immédiat à une donnée et la régulière nécessité d'arbitrer rapidement, comment se donner collectivement du temps pour l'analyse, l'interprétation et la prise en compte de l'ensemble des éléments participant à une décision juste et éclairée ?



Heures réalisées



Sinaps un outil qui transforme le Système d'Information

par Hermenegildo Carrilho, Responsable du pôle projets du SI à la DNUM, Université de Saint-Etienne

MISE EN
QUALITE
URBANISATION



Image par Wilfried Pohnke de Pixabay

Sinaps a été mis en production sur l'ensemble du périmètre prévu dans la V1 début septembre 2018 à l'Université Jean Monnet de Saint-Etienne.

Au démarrage de ce projet, la DSI avait délibérément validé avec le Directeur Général des Services que nous ne pouvions donner suite sans un engagement fort de la Direction de l'Université

Un projet avant tout Métier vs Informatique

Vu du côté des directions, la mise en œuvre de ce type d'outil ne paraît souvent pas prioritaire sachant que celles-ci travaillent la plupart du temps en silos avec leur propre système et se sentent propriétaires et garantes des données qui y sont produites. Faire comprendre que certaines données doivent être échangées avec d'autres directions dont la fonction et le périmètre sont différents n'est pas chose aisée. Pourtant on peut comprendre individuellement que les données sur les personnes puissent être échangées.

Qui n'a pas un jour maudit les administrations sur la production des mêmes documents d'identité ou références bancaires dans des services différents comme si l'établissement demandeur était un assemblage de containers étanches ?

L'enjeu principal est bien d'urbaniser (simplifier) le Système d'Information et non pas le système informatique. Directions Métiers et services informatiques sont donc bien embarqués dans le même bateau, pour

atteindre les mêmes objectifs de fournir les fonctions nécessaires à la vie de l'établissement.

Un partenariat avec l'Amue

Il faut parler dans cette aventure du lien important avec l'Amue. Partant du postulat que nous étions partenaires, dès la signature de la convention, nous avons eu un rôle important en tant que site Pilote que ce soit en phase de conception ou en phase de réalisation.

Cette relation a permis l'apport d'une confiance dans l'avenir de ce produit et de parvenir à réussir sa mise en production effective.

Le partage de la valeur des données

Identifier, et partager les données dites "maîtres" du Système d'Information permet de valoriser l'emploi de celles-ci en apportant toute la **Qualité** que demandent les exploitants. La véracité de l'information doit être vérifiée au plus proche de la saisie, chez les producteurs de données.

Nous pouvons ainsi nous focaliser sur les données importantes qui sont échangées entre nos différents domaines SI.

Une mise sous contrôle des données par la qualité et le Pilotage

La phase la plus critique pour un fonctionnement quotidien est bien de mettre en musique tout l'orchestre constitué par les

différents domaines métiers et la Direction du Numérique.

En effet, Sinaps transforme la façon de travailler en passant sur un mode horizontal qui implique une forte dimension d'échange entre les partenaires métiers dont le système est connecté à Sinaps. Pour cela il est essentiel de mettre en place une organisation transverse dès le départ dans l'équipe projet mais encore plus en fonctionnement. Nous avons fait le choix de confier la tâche de pilotage de Sinaps au cœur du métier de la Qualité dans une Direction du Pilotage et de la Qualité existante.

Un réel apport au niveau des équipes d'exploitation du Système d'Information

Si l'organisation autour de Sinaps n'est pas facile à mettre en place, elle permet néanmoins de réfléchir à la simplification des processus et à fluidifier la circulation des données. Les équipes métiers et DNUM échangent quotidiennement, permettant ainsi le redressement des données et d'éviter une pollution des systèmes.

Conclusion

Même si nous avons eu à traverser pas mal d'aléas liés à un site Pilote, nous sommes restés malgré tout fidèles à notre engagement et avons cru dans la nécessité de mettre en place cette brique essentielle du Système d'Information.

Sinaps peut-être un pas vers un Système d'Information intégré.



Des données de gestion en lien avec la scolarité, la vie étudiante et les réseaux Alumni



Par Michel Allemand (Directeur du projet) et Jean-François Langlamet (Expert métier) de PC-SCOL

Le projet PC-Scol développe la solution Pégase destinée à remplacer les outils Apogée de l'Amue et ScolarX, SVE de l'Association Cocktail. Même si le spectre fonctionnel attendu pour Pégase est large puisqu'il concerne aussi les aspects vie étudiante comme la gestion des activités sportives, la CVEC, les associations étudiantes et même si les futurs utilisateurs sont également les étudiants et les enseignants, Pégase reste une solution de gestion qui manipulera des données de gestion en lien avec la scolarité, la vie étudiante et les réseaux Alumni.

Pour éviter toutes les ambiguïtés dans les échanges de données avec les outils utilisés dans le domaine de la pédagogie comme les plateformes d'enseignement et les e-portfolios nous devons bien distinguer ce qui est du domaine de la pédagogie et concerne donc les données pédagogiques de ce qui relève de la gestion et concerne donc les données de gestion.

Nous allons illustrer ces différences sur deux exemples caractéristiques.

Construire l'offre de Formation

L'offre de formation est initialement une donnée pédagogique saisie puis consolidée dans le macro-processus "Construire l'Offre de Formation". Elle est conçue par les équipes pédagogiques dans un but purement pédagogique, dans le respect des principes de l'accréditation. Cependant, pour la mettre en œuvre c'est-à-dire avoir la possibilité de faire choisir à des étudiants ces formations, de suivre les résultats des étudiants dans ces formations, de leur délivrer un diplôme ou une certification ou de calculer le coût d'une formation, il faut transformer ces données pédagogiques en données de gestion dans le macro processus "Mettre en œuvre l'Offre de Formation". Le projet PC-Scol facilitera ce processus avec la possibilité que l'on offrira aux équipes pédagogiques de concevoir dans des "bac à sable" proposés par Pégase leur offre de formation qui sera ensuite reprise

dans Pégase par les équipes administratives pour ajouter tous les éléments nécessaires à leur gestion. Un syllabus est donc initialement une donnée pédagogique mais devient aussi une donnée de gestion pour pouvoir être présentée dans différents portails, les éléments permettant d'évaluer les connaissances des étudiants sont des données pédagogiques qui doivent être transformées en modalités de contrôle des connaissances permettant de calculer des résultats avec ou sans compensation par exemple. Ces éléments sont alors des données de gestion.

Evaluer les formés

L'évaluation est ainsi également un bon exemple. D'un côté il y a tout ce qui est du domaine pédagogique dans l'évaluation, que ce soit l'évaluation classique ou plus encore l'évaluation des compétences. Ce processus d'évaluation peut nécessiter des outils numériques, plateforme LMS et/ou eportfolio qui manipulent des données pédagogiques.

D'un autre côté il y a tout ce qui est du domaine de gestion comme la collecte des résultats des évaluations, mais aussi la propagation de ces résultats, l'agrégation et les calculs sur ces résultats ainsi que leur présentation pour des jurys par exemple. Ces éléments sont des éléments de gestion et les données manipulées sont alors des données de gestion. On peut donc en conclure que ces aspects doivent être traités par Pégase. De manière connexe se pose alors la question des référentiels et en particulier des référentiels de compétences qui sont des données de gestion même si elles sont utilisées dans des aspects pédagogiques. Ces données doivent être de base dans les outils de gestion et donc dans Pégase et être utilisées dans les outils "pédagogiques" comme les plateformes LMS ou les e-portfolios.



Intervention aux JRES2019 à Dijon



Création d'un groupe de travail Open Data

Pour construire collectivement un socle Open Data pour l'Enseignement Supérieur et Recherche, un appel à contribution est toujours actif. Les principaux objectifs de ce groupe sont décrits page 9.

Si vous souhaitez y participer, contactez-nous à numerique@amue.fr



L'ESR vu par le prisme de la donnée universitaire Une autre vision de nos organisations

Contactez-nous

Amue-DSPSI
numerique@amue.fr



Novembre 2019

Directeur général de la publication : Stéphane Athanase

Rédacteurs en chef : Bertrand Mocquet et David Rongeat. Secrétaire de rédaction : Mire.

ISSN 2650-8494

La collection numérique est sous Licence Creative Commons CC [BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Ont collaboré à ce numéro : (un grand merci)

Stéphane Amiard, Franck Estay, Guy Bisiaux, Rafia Berenguier, Victor Larger, Marion Lehmans, Anne Vanet, Soufiane Rouissi, Ferial Goulamhousen, Nicolas Cazenave, Lucile Schirr, Clotilde Tonnerre, Hermenegildo Carrilho, Michel Allemand et Jean-François Langlamet

Editeur : Amue – 103 boulevard Saint-Michel - 75005 Paris

Fabriqué en France

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

TOUS LES NUMEROS DE LA COLLECTION SONT EN TELECHARGEMENT

Amue. (2019). La collection numérique, à l'adresse <http://www.amue.fr/systeme-dinformation/metier/la-collection-numerique/>



PROCHAIN NUMERO DE LA COLLECTION NUMERIQUE (JANVIER 2020)

Patrimoine et numérique, vos propositions de témoignage et retour d'expérience à numerique@amue.fr

