La gouvernance des services numériques avec AGIMUS

David Reymond, Khadija Dib

▶ **To cite this version:**

David Reymond, Khadija Dib. La gouvernance des services numériques avec AGIMUS. 2nd International Symposium ISKO-Maghreb'2012, IRIT; Université de Lorraine; Loria; Université Paul Sabatier; ISKO Maghreb; Université Mohammed V; ESCT (Ecole Supérieur de Commerce de Tunisie), Nov 2012, Yasmine Hammamet, Tunisie. sic_01809857

HAL Id: sic_01809857

https://archivesic.ccsd.cnrs.fr/sic_01809857

Submitted on 11 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La gouvernance des services numériques avec AGIMUS[1]

D. Reymond[2], K. Dib[3]

***Abstract*— 98% des étudiants disposent aujourd'hui d'un environnement numérique de travail (ENT) avec un bouquet de services numériques. Ces services sont dispensés par les établissements via leur portail ENT. Pendante des coûts d'implantation, de suivi et de maintenance, se pose la question de l'utilité de la panoplie des services déployés ou, en d'autres termes, de l'évaluation du degré d'utilisation de chacun de ces services. Le degré d'utilisation peut prendre des formes variées avant d'entrer sur les formes d'usages chères aux études qualitatives. On le pose simplement ici en termes quantitatifs comme par exemple le taux de pénétration auprès des différentes populations, de fréquence d'accès ou de durée d'utilisation. Du point de vue d'une entité dispensant des services numériques à l'échelle de la gouvernance il peut s'agir aussi du nombre de services offerts et leur taux d'utilisation respectif.**

Les indicateurs de mesure d'activité se posent en évidence pour la construction de connaissances du comportement statistique des générations internet. On peut réaliser par l'appui sur ces données factuelles d'utilisation : l'amélioration du service offert, ou encore la mise en critique des diverses solutions technologiques. Ces indicateurs sont ainsi utiles pour la gouvernance des services et servent dans notre contexte à la création de connaissances à l'échelle territoriale : analyser les modes d'accompagnement aux services, les évolutions de degré d'usage afin d'établir, par exemple, les meilleures pratiques.

Nous proposons un dispositif de génération d'indicateurs enrichis, ouvert et libre, fort d'une flexibilité de mise en œuvre qui s'appuie sur des référentiels normatifs de description de profils utilisateurs et sur une catégorisation des services en gamme applicative. Nous montrerons comment le dispositif ouvre la voie dans une démarche qualité d'amélioration des services numériques mis à disposition par une ou plusieurs organisations, à la co-construction de connaissances, entre autres, sur le comportement statistique des usagers clients. Les exemples d'indicateurs mis en œuvre seront pris sur le terrain : les établissements de l'enseignement supérieur et de la recherche.

***Index Terms*— Environnement Numérique de Travail, Services numériques, analyses de logs, comportement des usagers**

***Keywords:* Cyber-space, web services, log analysis, usages, Consumer behavior**

I. La mesure des utilisations du web

Par l'analyse des actions et interactions des usagers avec le site et les services numériques, il est possible d'inférer une connaissance des pratiques et modes d'usage [1][2] de ces services. Plus généralement, en amont d'enquêtes plus approfondies, les éléments de réponse qu'offre le champ du Web Analytique (*Web Analytics*) permettent de critiquer les sites, leur agencement ou encore d'estimer leur utilité par la mesure des retours sur investissement. Ces études dans le contexte universitaire portent en général sur les ressources et documents électroniques [3].

La *Web Analytic Association* (WAA)[4] définit le Web analytique comme « la mesure, la collecte, l'analyse, le rapport de données internet dans le but de comprendre et d'optimiser les usages du Web ». C'est dans ce cadre générique que se situent nos travaux, incluant la difficulté d'échelle [4](le territoire français) et l'interopérabilité nécessaire des mesures[5], les impératifs techniques (hétérogénéité, solutions non bloquantes), et la délicatesse du sujet [6] (mesure comportementale sur données des usagers du SI). Nous prenons un raccourci en considérant que les services numériques, définis par la suite, s'appuyant sur des transactions http ou https sont équivalents à des sites web de contenus.

De nombreux travaux [7][8] depuis l'origine du web s'attachent à développer des algorithmes d'analyse de données de traces d'activité des serveurs web pour construire des connaissances sur les comportements usagers [9]. Jansen [1] offre un panorama global des différentes méthodes du web analytique positionnant l'objectif du web analytique au cœur d'un processus itératif d'amélioration du site : la collecte, le prétraitement des données, la production d'indicateur et leur interprétation. En sélectionnant quelques références proches de nos contraintes, nous distinguons notamment Mohd Helmy Abd Wahab, et al. [11] qui discutent les types de données de traces (*Transfer*, *Agent*, *Error Log* et *Referer Log*), types de formats de trace (le format *W3C* étendu, le format *Microsoft IIS*, et le format *NCSA Common*). Les auteurs présentent un algorithme de lecture et de nettoyage des données, de transfert vers une base de données pour le stockage, et l'utilisent pour les données d'un serveur ASP.

Doru Tanasa, et al. [12] proposent une méthode d'analyse de traces issues de différents serveurs appartenant à la même organisation. Les organisations de grande taille ayant recours à plusieurs serveurs pour leurs sites web, même si les usagers naviguent de façon transparente entre ces derniers de par leur

[1] Manuscript received 14 septembre 2012. This work was supported in part by the Ministère de l'enseignement supérieur et de la recherché (MESR) à la mission numérique pour l'enseignement supérieur (MINES).

[2] David Reymond est maitre de conférences à l'Université du Sud, Toulon-Var, rattaché au laboratoire I3M (e-mail: david.reymond@univ-tln.fr).

[3] Khadija Dib, docteur en informatique, pilote les universités numériques en région (UNR) au Ministère de l'Enseignement Supérieur et de la Recherche, à la mission numérique pour l'enseignement supérieur (MINES).

[4] Cf. http://www.digitalanalyticsassociation.org, visité en sept 2012.

importante inter liaison, il est nécessaire de pouvoir utiliser dans la même analyse les données d'activité issues de chaque serveur. L'analyste doit pouvoir rassembler les chemins empruntés par les usagers au travers des différents serveurs de l'organisation afin de mener une analyse centrée usager (*User Centric*). Les auteurs proposent une solution pour fusionner les fichiers de log et reconstruire les visites en trois étapes : addition et anonymisation des traces, purge des données inutiles, puis structuration dans le format voulu (identification usager, session, visites, périodes, etc.). Les auteurs rajoutent une nouvelle étape à ce procédé de prétraitement classique qui permet à l'analyste de ne sélectionner que les informations d'intérêt pour le stockage et la généralisation.

K. R. Suneetha, et al. [13] visent l'analyse des erreurs principales d'un site afin d'extraire celles provenant des visiteurs potentiels, pour aider les administrateurs systèmes et les designers de sites à leur amélioration en déterminant les liens perdus à l'aide de ces erreurs. Les auteurs agrègent les données depuis différents types de données de trace (coté serveurs web, serveurs proxy et côté client).

Plus avant, Fang Yuan, et al. [14] extraient des traces des traits caractéristiques d'utilisation principale que l'on peut classer en trois catégories : les groupes d'usagers équivalents, les groupes de pages d'intérêt et les chemins fréquemment parcourus. Ces traits saillants sont utilisés pour améliorer la performance des serveurs web et la qualité de services aux utilisateurs finaux. Dans le but de mettre en œuvre ce procédé, les auteurs détaillent les algorithmes pour connaître la topologie structurelle du site, pour le nettoyage des données, l'identification des usagers et des sessions, ainsi que les chemins parcourus. Une étape de prétraitement des données est utilisée pour la consistance de l'opération de fouille de données.

Plus récemment, Honest et al [15] dans un contexte d'analyse de site web universitaire, soulignant l'hétérogénéité et la volumétrie des données, proposent une méthode pour agencer une analyse analytique des données d'usage avec la réalité organisationnelle afin de produire des métriques en phase avec les problématiques communicationnelle de ces institutions. Les auteurs combinent les traces serveurs, la topologie du site et le calendrier universitaire dans le cadre d'un serveur web disposant de pages dynamiques construites par un CMS. Leur approche permet de recomposer l'architecture du site et de s'abstraire de la fragmentation provoquée par ce genre d'édition.

Nos travaux s'inscrivent dans cette lignée, cependant les services numériques sont accessibles en mode authentifié (les usagers sont connus du système d'information), nous devons produire des indicateurs à l'échelle du territoire national et non de quelques serveurs en garantissant les impératifs de cohérence des indicateurs. Nous détaillons ci-après les quelques points clés contextuels et contraintes du projet.

II. Services numeriques et Enseignement Superieur

A. *Historique et situation actuelle*

Dans les établissements d'enseignement supérieur, de nombreux services numériques (messagerie électronique, intranet, forums, publication de site Web, agendas, services liés à la scolarité, services de documentation, gestion budgétaire et comptable, cours en ligne, et d'autres encore) sont aujourd'hui à la disposition des étudiants, des enseignants et des personnels administratifs. Ils sont pour la plupart accessibles à travers les réseaux, via une authentification unique (*Single Sign On* - SSO). Historiquement, la direction de la technologie, du ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche a décidé de favoriser le déploiement de ces services numériques à travers plusieurs actions notamment les universités numériques en région (UNR) pour répondre aux objectifs stratégiques :

a) le développement par les établissements de services numériques pour offrir à la communauté universitaire (étudiants, enseignants, chercheurs et personnels administratifs) un environnement numérique de travail (ENT)
b) la prise en compte de l'accès de l'ensemble des étudiants à ces services (équipements collectifs et individuels, réseaux…)
c) l'accompagnement pour les usages du numérique.

Depuis, nous avons pu identifier via une enquête auprès des services informatiques [4] des établissements qu'une gamme applicative et une variété de composition de services très hétérogène caractérisent les ENT au plan national. La maturité de ces portails de service (portails effectifs ou virtuellement établis via le SSO) d'établissement est cependant atteinte.
Aujourd'hui, la mission numérique pour l'enseignement supérieur (MINES), au sein du ministère de l'enseignement supérieur, assure le suivi et la coordination nationale des UNR. Actuellement 17 UNR couvrent totalement le territoire national. Les principaux services numériques déployés au sein des UNR s'articulent essentiellement autour de la scolarité, l'e-pédagogie, les outils collaboratifs et de communication, la documentation, la vie universitaire et les relations avec les entreprises. D'autres services peuvent être déployés de façon singulière par les établissements. Ainsi, selon les dernières estimations, 98% des étudiants français disposent d'un bouquet de services numériques.

B. *Présentation du projet AGIMUS*

Dans un souci de fournir un service de qualité à l'utilisateur final, la sous-direction des technologies de l'information et de la communication pour l'éducation (SDTICE) au ministère de l'enseignement supérieur et de la recherche souhaite fournir aux UNR un dispositif permettant la remontée automatique d'indicateurs des usages des services numériques déployés dans les établissements. L'enjeu immédiat est de pouvoir faire évoluer et adapter ces services numériques qui sont le pendant « virtuel » des campus universitaires sur la base factuelle des données d'usage qui en sont faits. Si chaque service dispose le plus souvent de données de traces d'activité, celles-ci sont fortement liées à la technologie applicative du service, ces données sont coûteuses et fastidieuses à analyser, et généralement non interopérables : par l'utilisation simple de ces données de trace, il demeure très difficile d'analyser la navigation des usagers entre les services) car le codage des

identifiants utilisateurs peut être différent, et de fait la comparaison des traces (logs) sur des services similaires mais de technologie différente n'est pas immédiate. D'autres limites bien connues surgissent dès lors qu'il s'agit de traiter la volumétrie très conséquente [15] de ces fichiers de trace.

1) Objectif

L'objectif du projet se résume au plan opérationnel en l'uniformisation des données de trace (log) pour la production d'indicateurs indépendants des technologies mises en place, puis de permettre une lecture synthétique par la création d'indicateurs homogènes. Au plan fonctionnel, il s'agit de pouvoir comparer les usages de populations variées (étudiants, personnels) à différents niveaux (année d'inscription, discipline ; etc.) sur des technologies différentes, de comprendre les modes d'usages et de disposer de tableaux de bords permettant à la gouvernance des établissements le pilotage des services dispensés. L'hétérogénéité applicative nationale induit une difficulté immédiate à la lecture comparative des traces d'activité avant même la production d'indicateurs interopérables. L'enjeu de production de ces indicateurs s'inscrit dans la démarche qualité de mise en œuvre des services numériques par les établissements, afin de pouvoir participer au processus de décision des améliorations, des évolutions et des adaptations qui favoriseraient l'adoption de ces ENT par les usagers. De surcroît il s'agit aussi de pouvoir discuter de leur efficacité et de les adapter aux utilisations qui en sont faites.

En comparant ces indicateurs entre différents établissements (ou solutions de services numériques), il s'agit de capitaliser la connaissance des utilisations effectives transversalement entre les différentes structures, même concurrentielles afin de détecter des divergences ou des singularités. Enfin, l'agrégation inter-établissement, de différents indicateurs aux différentes échelles hiérarchiques (que constituent les établissements, les UNR ou regroupements de structures, et enfin le ministère) offrira une source de données d'usages utile *in fine* à la construction de « bonnes pratiques » de mise en œuvre, mais aussi de projection des tendances. Cette base permettrait ainsi une répartition nationale ou régionale des coûts expérimentaux de la mise en œuvre de nouveaux services ou de validation/estimation d'une amélioration de services existant.

2) Contraintes et préconisations

Les préconisations d'élaboration de la solution se résument dans les termes de collecte de données, traitement puis agrégation récursive à différentes échelles, précisés par les points suivant :

1. La définition d'une architecture de collecte des données, non intrusive et contextuellement paramétrable pour intégrer les diverses modalités de dispense des services ;
2. La collecte automatisée de ces données d'activité et leur homogénéisation, qui soit indépendante des technologies, et qui réponde à critères précis utiles à leur interopérabilité pour la production des indicateurs d'usage. Un service spécifique constitue l'unité élémentaire sur laquelle est produite un ensemble d'indicateurs. En outre, les données doivent être préparées pour l'agrégation nécessaire à la production d'indicateurs à différentes échelles. L'agrégation de premier niveau se construit ainsi à partir de données issues d'une collection de services. Le niveau suivant se construit à partir d'une collection de données de collecteurs (cf. Figure 1).
3. Cette construction récursive permet la mise en œuvre de mode de collecte répondant à toute organisation hiérarchique ou virtuelle. Ainsi la construction (établissement, UNR, Ministère) présentée précédemment est immédiate tout en laissant libre à d'autres arbres de collecte issues de regroupements aussi variés que le paysage universitaire.
4. Disposer d'une interface permettant aisément la lecture comparative des données (historiques, évolution, et agrégation d'indicateurs entre établissements d'une UNR) mais aussi de pouvoir scinder les usages selon diverses catégories de population des établissements (étudiants, personnels administratifs, enseignants chercheurs, etc.). Idéalement cette scission sera effectuée pour la catégorie étudiant à une granularité plus fine : celle des filières principales auxquelles ils sont rattachés.

Une première expérimentation à l'échelle d'une UNR a été réalisée en 2008 en Aquitaine [4]. Cette action expérimentale menée a permis d'exploiter la technologie des marqueurs afin de produire les données permettant d'alimenter les indicateurs, d'en montrer l'intérêt sur le plan décisionnel à chacun des niveaux et, aussi, de constater la faisabilité opérationnelle [5].

C. Application de Gestion des Indicateurs de Mesure des Utilisations de Service (AGIMUS)

Suite à un appel d'offre lancé en octobre 2009, l'application AGIMUS a été développée sous maitrise d'ouvrage du Ministère et est aujourd'hui fonctionnelle. La solution proposée s'abstrait de la difficulté technique de construction des indicateurs sur la variété des gammes applicatives et des architectures de dispense des services ; elle permet aux établissements d'inclure leur contexte de déploiement et d'usage des services et, pour en favoriser l'appropriation, offre aux établissements la possibilité de construire ses propres indicateurs. Le même dispositif est utilisé pour produire des agrégations et construire des arborescences de collecteurs de données d'utilisation des services aussi variés que le paysage de l'ES (Instituts, Ecoles d'ingénieurs, Universités, UNR, PRES, etc., sont autant de paradigmes de mutualisations de moyens ou d'objectifs pouvant nécessiter d'autres indicateurs de mesure d'utilisation de services numériques).

1) Définition des services numériques des ENT

Loin d'être unanime, la notion de service numérique varie selon les points de vue. Suite à une enquête lancée au niveau national auprès des informaticiens en charge de services numériques nous avons pu collecter des points de vue pragmatiques de cette notion. En phase avec les objectifs de définition du périmètre de l'application de mesure des utilisations effectives des services, nous retenons une définition technique :

Un service numérique de l'ENT se définit en trois points :

1. Une application en mode client/serveur Web (que les usagers utilisent à l'aide d'un navigateur embarqué ou pas)
2. L'accès à l'application nécessite une authentification qui est centralisée par un service d'authentification unique (*Single Sign On* - SSO) s'appuyant sur un annuaire d'identification ;
3. L'identification est réalisée en appuie sur un annuaire des membres de l'organisation ou un annuaire fédéré.

Par cette définition est englobé l'ensemble des services numériques d'un établissement relevant d'un accès web en mode authentifié. Si toutefois le mode d'architecture de l'ENT suivait un schéma différent (par ex : l'ENT est constitué par la plateforme pédagogique), le caractère ouvert du mode de

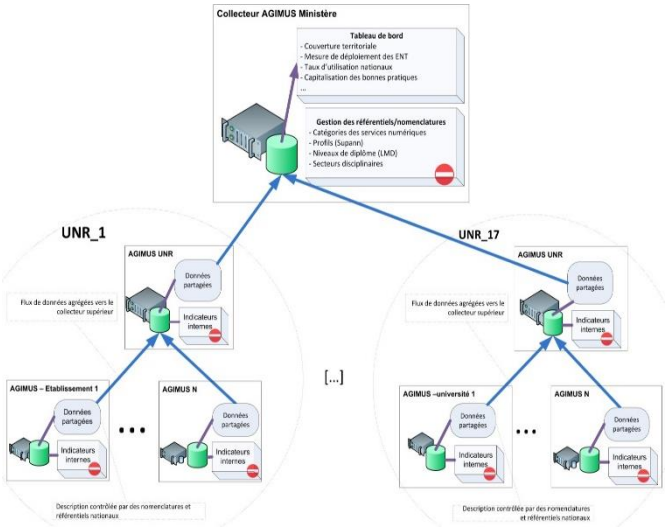

Figure 1 : Architecture d'implantation récursive interétablissement, flux de données et nomenclature d'échange

recueil des données d'utilisation laisse possible une adaptation *in situ* du dispositif.

2) Collecte et prétraitement

AGIMUS s'implante dans le système d'information des établissements par sa connexion à l'annuaire LDAP et l'ouverture d'une zone de dépôt pour les fichiers de trace des serveurs web, frontaux aux applications qui supportent, dans la majorité des cas, l'ensemble des interactions des usagers avec les applications web. Un second point clé est de tracer l'identification initiale de l'usager pour lui émettre un cookie qui sera suivi dans l'ensemble de ses transactions avec les différents applicatifs. L'association cookie-Id de l'usager sera utilisée pour l'enrichissement des données usagers dans la phase traitement-anonymisation par hachage [17] à l'aide de l'annuaire et permettre ainsi la caractérisation des profils. La caractérisation est à deux niveaux : en premier il s'agit des types de membres. On distingue ici un enseignant d'un étudiant par exemple. Pour pouvoir être agrégés correctement il convient de s'appuyer sur une nomenclature. L'entrée *eduPerson* du schéma Supann a été retenue. Pour les profils étudiants, l'annuaire de l'établissement est interrogé dans cette phase afin d'augmenter la caractérisation des numéros anonymes : niveau de formation (LMD), secteurs disciplinaires, type de formation (Régime d'inscription) et affiliation (composante) constituent les dimensions majeures de cette caractérisation. Des dimensions libres sont présentes dans la base de données AGIMUS au choix de l'établissement d'associer d'autres axes d'analyse (par ex. formation, sexe, groupes). Les dimensions majeures sont normalisées par les référentiels nationaux existants (base de données SISE[5] et annuaire Supann[6]) comme indiqué dans la Figure 2.

III. Indicateurs produits par AGIMUS

Les indicateurs initiaux sont relativement simples (cf. annexe). Ils se distinguent en trois catégories selon leurs objectifs d'utilisation : suivi et maintenance, pilotage et indicateurs d'exploitation propres au système de collecte.

A. *Les indicateurs de base d'AGIMUS*

1) Indicateurs de suivi

Les indicateurs de suivi informent sur les services opérationnels au quotidien et leur degré d'utilisation. Ils sont utiles à une planification de mise à jour des applications, de modifications pour permettre de choisir des périodes creuses afin de gêner le moins possible les usagers. Au plan N+1, ces indicateurs détaillent des comparatifs de degré d'utilisation effectif par niveau N. Ils permettent d'estimer a priori un degré d'utilisation de chacun des services. Cette gamme d'indicateurs couvre par ex. le nombre de connexions, les utilisateurs uniques quotidiens et le nombre de services opérationnels dans les ENT.

2) Indicateurs de gouvernance

Les indicateurs de gouvernance décrivent les services et les gammes de service (par dénombrement), le comportement moyen des usagers exprimés par la durée de session, ainsi que la fréquence de ces utilisations. Qu'ils soient produits au niveau d'un établissement ou par agrégation ces indicateurs informent sur l'utilisation effective des services. Les établissements disposent du taux de connexion brut pour compléter les analyses de ces utilisations.

Bien entendu, chacun des indicateurs précédemment cités peut se décliner selon les catégories de populations et, pour la population étudiante, selon les dimensions qui la caractérise (cf. Figure 2).

[5] Cf la Base Centrale des Nomenclatures et le Système d'Information sur le Suivi de l'Étudiant – SISE. Cf. http://infocentre.pleiade.education.fr/bcn.

[6] Cf. http://www.cru.fr/documentation/supann/2009/index.

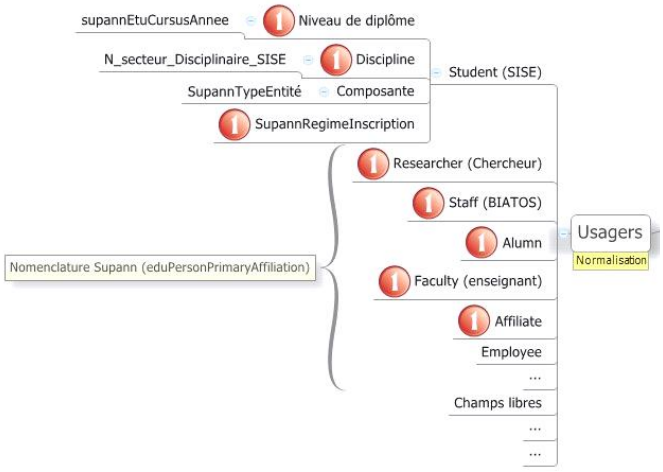

Figure 2 : Caractérisation des usagers dans AGIMUS, les symboles marquent les dimensions agrégées par le dispositif.

3) Indicateurs d'exploitation

Ces indicateurs ont pour vocation d'aider à sa mise en place et à la validation de la procédure d'audit des services numériques. Au plan N+1 ils sont reconstitués par une agrégation sommative. On distingue à ce niveau les dénombrements de population, la liste des services tracés et les URL détectées comme potentielle application à tracer.

Pour plus de détails sur ces indicateurs, le lecteur est renvoyé à [6][18].

B. Indicateurs composites

Au plan de la gouvernance une vue plus synthétique sera proposée pour obtenir une description factuelle de l'ENT de façon bi annuelle. A l'échelle de l'établissement est ainsi proposée une panoplie d'indicateurs décrivant la richesse de l'ENT et le degré d'utilisation.

1) La richesse de l'ENT

Cette dernière est définie comme le nombre de services offert par catégorie fonctionnelle. A la gamme de services dispensés, nous mettons en exergue dans cet indicateur les applications les plus utilisées.

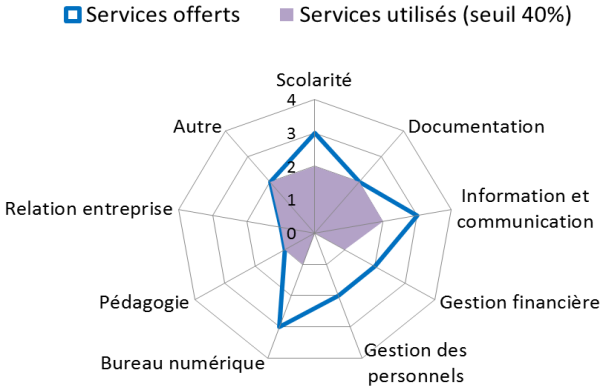

Figure 3 : La richesse fonctionnelle de l'ENT et degré d'utilisation des services

Un seuil (arbitraire de 40% dans l'exemple de la Figure 3) est appliqué pour afficher les services utilisés par plus de 40% de la population. Ainsi, dans l'exemple ci-dessus, nous pouvons voir que la gamme applicative est couverte par l'ENT de cet établissement à l'exception de la Gestion des Personnels. Pour la gamme Scolarité, trois services numériques sont offerts. Deux de ces services sont utilisés par plus de 40% de la population. Cet indicateur développe une vision synthétique tant de la couverture des services numériques (les applications déployées) que des usages effectifs qui sont faits de ces applications.

2) Les hit-parades

Une seconde vue synthétique est proposée par l'ensemble des éléments phares de l'ENT : sous la forme de hit-parade ces indicateurs délivrent les applications les plus utilisées ou encore les profils les plus consommateurs.

Par exemple, la Figure 4 below montre pour un ENT donné le taux d'utilisation pour l'ensemble de la population des trois principaux services. Cet indicateur met en exergue les services les plus utilisés pour lesquels une attention particulière doit être portée (disponibilité, fiabilité) ou encore une zone de communication médiatique (à recouper par la fréquence d'utilisation).

Un autre exemple en ce sens est d'extraire les disciplines les plus consommatrices des services de l'ENT. La Figure 5 below montre une répartition d'usage par discipline pour un établissement pluridisciplinaire.

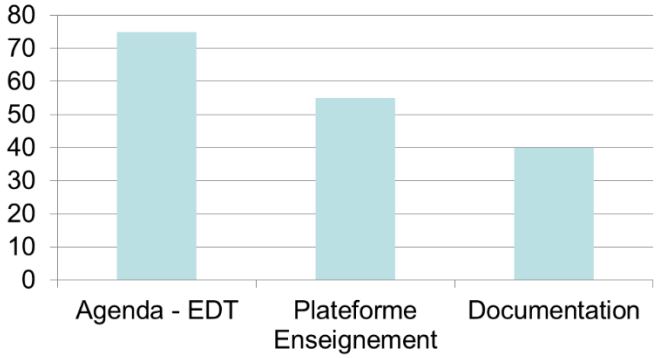

Figure 4 : Hit parade des applications les plus utilisées.

Cet indicateur permettrait d'interroger plus avant de l'intérêt des différents services offerts pour les branches les moins consommatrices (ici Langues et Droit) et initier des campagnes d'information, d'accompagnement ou d'étude de quels services seraient nécessaires.

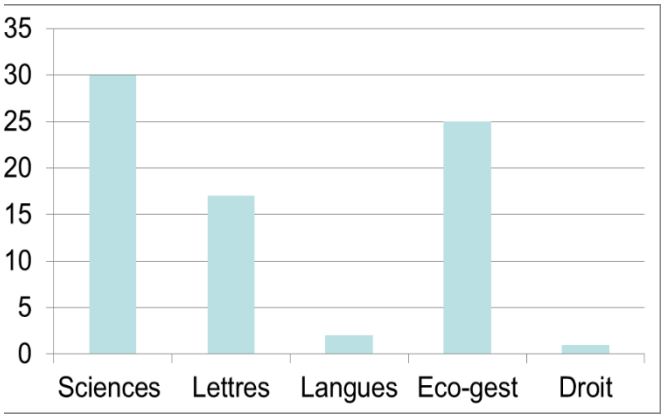

Figure 5 : Comparaison (fictive) de la consommation de service par discipline

Enfin, par la combinaison des deux indicateurs précédents AGIMUS propose la création d'indicateurs composites tel que présenté à la Figure 6. Il s'agit de la répartition comparative à l'échelle nationale (sur une UNR) présentant la consommation du service webmail par niveau de diplôme et par année (format LMD suivi du numéro de l'année).

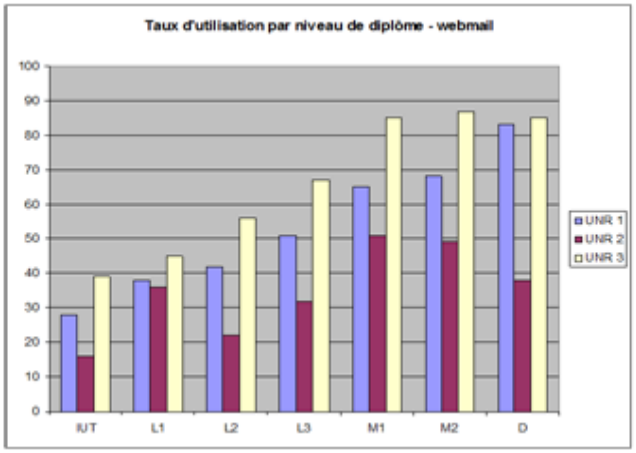

Figure 6 : Répartition de la consommation du service webmail selon les niveaux de diplôme. Les taux sont donnés comparativement à une échelle régionnale.

Outre l'aspect comparatif permettant de souligner les outils les plus adoptés par les populations, il peut s'agir aussi de mettre en évidence les meilleures pratiques d'information, d'accompagnement, ou de formation aux outils. C'est sur cette dimension comparative possible grâce à la normalisation des données que pourra se construire une base de connaissance à l'échelle du territoire

C. Base de connaissances nationales

L'utilisation des référentiels nationaux de description des populations et de caractérisation des services (affectation de services) va permettre de créer au niveau de l'instance nationale une base de connaissances co-alimentée automatiquement par les différentes instances AGIMUS.

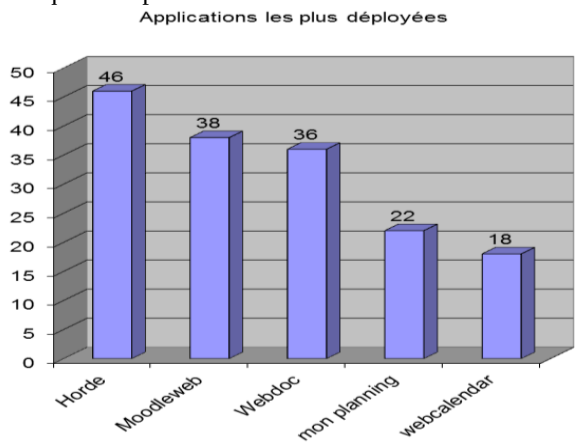

Figure 7 : Applications les plus déployées (sept 2012)

Au-delà d'indicateurs informatifs tel que les applications les plus déployées (Figure 7), les plus utilisées au plan national il s'agit de délivrer des informations sur les usages les plus communs et d'établir des niveaux de référence utile à la gouvernance des services numériques. En effet, à ce stade, nous ne pouvons établir encore quels sont les « bons » niveaux de références d'usage des services numériques :

- Quel est le niveau moyen de connexion à un service numérique ?
- Quel est sa fréquence d'usage moyenne ?
- Etc.

Pour compléter ces analyses il conviendra d'établir des niveaux de référence selon les différentes catégories afin d'établir les éventuelles spécificités déclinées selon les niveaux d'étude et les disciplines.

Grâce à l'établissement de cette connaissance collective du comportement des étudiants (population majoritaire cliente des services numériques des ENT d'établissement), il sera alors possible d'amorcer, en interne, des accompagnements spécifiques, des campagnes de formation ou de communication (en complément des études qualitatives sur les usages ou non usages constatés) pour favoriser l'appropriation et aider au pilotage et à la mise en œuvre des services numériques.

En suivant les mêmes principes il sera alors possible de répartir des éventuels coûts de mise à disposition des services.

IV. CONCLUSION

La méthode de production d'indicateurs associée à une méthode de pilotage est sous tendue par des recherches actuellement très actives au plan international [20][21][24].

La solution technique mise en œuvre, développée dans le domaine du libre, est adaptable aux situations variées de mise à disposition des services et reste évolutive en laissant place à l'ajout de nombreux services numériques dont la gamme s'étend au quotidien. AGIMUS est actuellement en cours de déploiement sur le territoire national français.

L'interopérabilité des données est assurée par l'uniformisation réalisée par les dispositifs AGIMUS opérant à la base de la collecte. La fiabilité des mesures effectuées est garantie par la mise en œuvre d'AGIMUS par les détenteurs de l'information sur les spécificités locales [5][6] : les gestionnaires desdits services qui activent ou pas l'agrégation des données au collecteur supérieur pour certains services.

La granularité des indicateurs appuyée sur des regroupements naturels de profils des organisations universitaires, une nomenclature standardisée de leurs descriptions offre la possibilité d'un pilotage actif d'accompagnement, de mise en œuvre et de déploiement des services numérique [19]. L'utilisation et la capitalisation en inter-établissement de ces indicateurs ouvre la voie à une vision performative des services numériques tout en permettant une répartition des expérimentations et la capitalisation des bonnes pratiques sur la base de données factuelles. L'adoption généralisée de cette solution mettra les établissements universitaires sur un piédestal égalitaire avec les grands acteurs du web dont les services numériques alléchants font déserter la population étudiante de la plupart des services offerts. Agimus associé à un référentiel national, ouvre la voie à des modes de mise en œuvre, de rajout

fonctionnels, d'amélioration ergonomiques en phase avec les utilisateurs effectifs et potentiels que les établissements se doivent de conserver dans leur pendant virtuel que représente les ENT [25][26][27][28].

REFERENCES

[1] Jansen, B. J. "Understanding User-Web Interactions via Web Analytics", N°6 in *Synthesis Lectures on Information Concepts Retreival and Services*, Morgan & Claypool Publishers, 2009.

[2] Jaideep Srivastava, Robert Cooley, Mukund Deshpande, and Pang-Ning Tan. 2000. Web usage mining: discovery and applications of usage patterns from Web data. *SIGKDD Explor. Newsl.* 1, 2 (January 2000), 12-23. DOI=10.1145/846183.846188

[3] Boukacem-Zeghmouri, C., (éd.). *L'information scientifique et technique dans l'univers numérique : mesures et usages*, ADBS éditions, 319 p. 2010.

[4] Reymond, D. and Dib, K. 2008, « Indicateurs d'usage des services numériques déployés au sein des universités numériques en région », dans *Actes du colloque international "l'Université à l'ère du numérique"* (CIUEN 08), Bordeaux.

[5] Reymond, D. and Dib, K., 2009, « Vers une interopérabilité de la mesure d'usage des ENT : enjeux, objectifs et méthode », Intelligence collective et organisation des connaissances, in *actes du 7e Colloque du Chapitre français de l'International Society for Knowledge Organization* (ISKO), M. Hassoun et M. El-Hachani (Eds), Ensib, Lyon, p. 287–293.

[6] Reymond, D. and Dib, K., 2010, « Mesure d'usage et organisations multi-échelles : indicateurs et méta indicateurs d'usage des ENT.», dans *colloque international « l'Université à l'ère du numérique » (CIUEN)* juin 2010, Strasbourg

[7] Ming-Syan Chen, Jong Soo Park and Philips S. Yu, 1996 "Data mining for path traversal patterns in a web environment". In *Proceedings of the 16th International Conference on Distributed Computing Systems* (ICDCS '96). IEEE Computer Society, Washington, DC, USA, p. 385-392.

[8] R. Cooley , B. Mobasher , J. Srivastava., Web Mining: Information and Pattern Discovery on the World Wide Web. *Proc. IEEE Intl. Conf. Tools with AI*, Newport Beach, CA, pp. 558-567, 1997.

[9] Osmar R. Zaïane, "Web Usage Mining for a Better Web-Based Learning Environment", in Proc. of *Conference on Advanced Technology for Education*, pp 60-64, Banff, Alberta, June 27-28,2001.

[10] C. Shahabi , A. M. Zarkesh , J. Adibi , V. Shah, Knowledge discovery from users Web-page navigation, *Proceedings of the 7th International Workshop on Research Issues in Data Engineering (RIDE '97) High Performance Database Management for Large-Scale Application*s, p.20, April 07-08, 1997

[11] Mohd Helmy Abd Wahab, Mohd Norzali Haji Mohd, et. al, "Data Pre-processing on Web Server Logs for Generalized Association Rules Mining Algorithm" , *World Academy of Science, Engineering and Technology*, 2008

[12] Doru Tanasa and Brigitte Trousse, "Advanced Data Preprocessing for Intersites Web Usage Mining", *IEEE Computer Society*, March/April, 2004.

[13] K. R. Suneetha, Dr. R. Krishnamoorthi, "Identifying User Behavior by Analyzing Web Server Access Log File", *International Journal of Computer Science and Network Security*, Vol. 9 No. 4, April 2009.

[14] Fang Yuan, Li-Juan Wang, Ge Yu, " Study on Data Preprocessing Algorithm in Web Log Mining ", *Proceedings of the Second International Conferences on Machine Learning and Cybernetics*, Xi'an, 2-5 November 2003.

[15] Nirali Honest, Bankim Patel and Atul Patel. Article: Applying Web Usage Mining to a University Website Access Domain. *International Journal of Applied Information Systems* 2(9):7-14, June 2012. Published by Foundation of Computer Science, New York, USA

[16] Reymond, D., Dib, K., and Kromm H, (2012) "Data-mining and indicators in cyberinfrastructure. Zoom out perspective in French universities cyberspace", *Proceedings of the IEEE EDUCON2012 International Conference: Collaborative Learning & New Pedagogic Approaches in Engineering Education*, Morocco 17-20 April 2012, p. 1-10, DOI 10.1109/EDUCON.2012.6201068.

[17] Goy Alain, L'appariement sécurisé de fichiers d'étudiants grâce au « hachage » des identifiants, *Courrier des statistiques* n°113-114, mars-juin 2005, p 27-32

[18] Reymond, D. and Dib, K. 2010, « Mesures d'utilisation : vers un pilotage intelligent des services numériques de l'enseignement supérieur », In *Actes du Colloque International AFRICAMPUS, Les usages intelligents des technologies de l'information et de la communication dans la réorganisation universitaire*, Dakar Sénégal.

[19] Gerbod, D. and Paquet. F. *Les clefs de l'e-administration*, Pratiques d'Entreprises, Editions Management et Société, 2001

[20] Nicholas, D, Huntington, P, "Evaluating the Use and Users of Digital Journal Libraries". In: *Digital Libraries*, Ed. Papy, F, 2008, London: ISTE Wiley

[21] Nicholas D, Rowlands I. "Digital consumers: case study virtual scholars. A deep log analysis". In *L'information scientifique et technique dans l'univers numérique : mesures et usages.* ADBS, 2010, pp 27-42

[22] Harley D., "Use and Users of Digital Resources: A survey explored scholar's attitudes about educational technology environments in the humanities". *EDUCAUSE Quarterly* N°4, 2007

[23] Harley, D. "Why Understanding the Use and Users of Open Education Matters". In *Opening Up Education: The Collective Advancement of Education through Open Technology, Open Content, and Open Knowledge*, edited by Toru Iiyoshi and M.S. Vijay Kumar. Cambridge, MA: MIT Press, 2008.

[24] Abdullah, F. (2005) "HEdPERF versus SERVPERF: The quest for ideal measuring instrument of service quality in higher education". *Quality Assurance in Education*, Vol. 13, No. 4, pp 305-327

[25] Jacquinot G. and Fichez E., *L'université et les TIC*, Collection: Perspectives en éducation et formation, De Boeck Université, 2008. 328 pages.

[26] Brown, M, Paewai, S and Suddaby, G. "The VLE as a Trojan Mouse: Policy, Politics and Pragmatism" *Electronic Journal of e-Learning Volume 8 Issue 2*, 2010, pp 63–72, *available online at www.ejel.org*

[27] Dyson, M., Barreto Campello S., 'Evaluating Virtual Learning Envirronment'. *The Electronic Journal of e-Learning (EJEL)*, p. Vol 1, N°1; Feb 2003, p. 11–20, *available online at www.ejel.org*

[28] Paivandi, S. « L'enseignement à distance : un facteur de changement à l'université », in [28], chap. 13, p177-188

[29] Sun-Mi K.and Verrier C., Le plaisir d'apprendre en ligne à l'université. Implication et pédagogie, Bruxelles, De Boeck Université « *Perspectives en éducation et formation* », 2009, 228 pages.

ANNEXE 1 : INDICATEURS DE BASE D'AGIMUS

Le dispositif permet la création d'indicateurs spécifiques internes, propres à l'établissement dont la complexité peut être plus fine (accès à la base SQL). Dans cette première phase seuls quelques indicateurs basiques mais utiles à la compréhension du phénomène services web et leurs utilisations effectives sont produits. Nous donnons ci-après les indicateurs intégrés dans le package de base AGIMUS. Le tableau 1 présente les indicateurs de suivi, utiles à la gestion quotidienne des services en offrant un panorama global sur ces derniers : les connexions pour les périodes de temps à paramétrer, les services opérationnels et les utilisateurs uniques quotidiens. Le niveau N+1 identique au précédent et se calcule par sommation sur les Agimus interconnectés. A ce niveau il est également possible de visualiser ces indicateurs en « catégories de services », concept qui permet aux établissements niveau N d'avoir une composition variée de leur offre de service.

Tableau 1 : Indicateurs de suivi intégrés au dispositif AGIMUS

Niveau N)	Niveau N+1
Nombre de connexions	Nombre de connexions
Nombre de services opérationnels	Nombre de services opérationnels
Utilisateurs effectifs quotidiens	Utilisateurs quotidiens

Le Tableau 2 présente les indicateurs dits de gouvernance. Ils sont utiles au pilotage à court et moyen terme des services numériques. Ces indicateurs informent tant sur les différents services déployés que sur leur usage effectif. Au niveau N+1, les différents indicateurs sont présentés par catégorie de service pour lisser les singularités internes. Le calcul des indicateurs s'effectue par sommation et moyenne sur l'ensemble du service du niveau. La fréquence moyenne d'utilisation exclue certains services (Tchat, visio par ex.).

Tableau 2 : Indicateurs de gouvernance intégrés au dispositif pour les niveaux N et N+1

Niveau N	Niveau N+1
Couverture applicative	Couverture applicative par catégorie de services
Type de navigateurs clients	
Durée moyenne de session	Durée moyenne de session par catégorie de services
Fréquence moyenne d'utilisation	Fréquence moyenne d'utilisation par catégorie de services
Taux de connexion bruts	Taux de connexion bruts par catégorie de services

AGIMUS permet également de décliner ces indicateurs selon les catégories de population renseignées par l'annuaire. Les taux sont calculés sur des dénombrements de populations (un dénombrement pour chacune des dimensions usagers de la Figure 2 multiplié par les entrées possibles de ces dimensions, seule la dimension « profil usager » est déroulée dans ce schéma).

D'autres indicateurs dits de maintenance ou d'exploitation d'AGIMUS sont récapitulés dans le tableau suivant. Le **Erreur ! Source du renvoi introuvable.** décrit les indicateurs d'exploitation propres au dispositif AGIMUS. Ces indicateurs ont pour vocation d'aider à sa mise en place et à la validation de la procédure d'audit des services numériques. Au plan N+1 ils seront reconstitués par une agrégation sommative.

Tableau 3 : Indicateurs d'exploitation du dispositif AGIMUS

Indicateurs	Définition
Liste des URL détectées	Permet d'identifier de nouvelles applications traçables Permet d'approfondir la granularité
Liste des services tracés	La liste des applications effectivement déclarées et analysées
Dénombrements de population	Les dénombrements sont utilisés dans les calculs des taux