

Les collections numériques : Données et indicateurs en bibliothèque universitaire

Thomas JOUINEAU

Université de Lorraine : Bibliothèque numérique

Consortium COUPERIN : Groupe de travail Indicateurs



Les normes ISO, COUNTER...

I. ÉTAT DES LIEUX

L'édition scientifique et technique

2015 : 25,2 milliards de dollars de revenus au niveau mondial dont 10M pour les seules revues STM (Sci, Tech., Med.) anglophones*

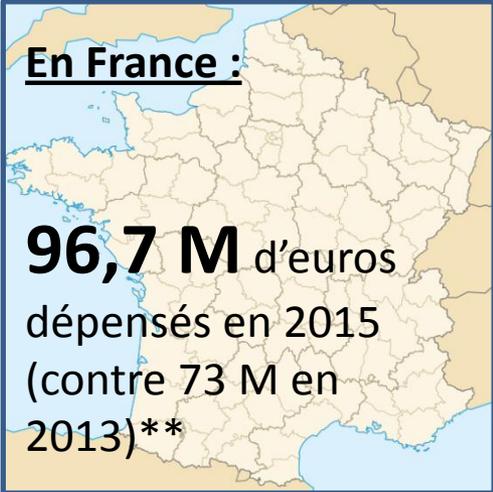
1.5 milliards d'articles sont téléchargés chaque année par 10 millions de lecteurs et usagers.*

Le revenu mondial des revues en sciences, technologie et médecine a augmenté de 8 à 10 milliards entre 2008 et 2013*

Depuis de nombreuses années, le prix de la plupart des revues augmente de 5 à 10% par an*

Les quatre plus gros éditeurs se partagent plus de la moitié du marché mondial*

En France :



96,7 M d'euros
dépensés en 2015
(contre 73 M en
2013)**

A l'Université de
Lorraine :

1,8 M d'euros
dépensés
en 2015



Sources :

* *The STM report : an overview of scientific and scholarly publishing.*, STM, 2015.

http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

** *Enquête sur les ressources électroniques*, Couperin, 2015

La mesure de l'utilisation numérique : normes et recommandations



Un jeu de normes internationales, traduites en français par l'AFNOR :

ISO 2789 : définition des données

ISO 11620 : méthode de création des indicateurs

ISO 16439 : mesure de l'impact



Une recommandation, créée en 2003, actuellement en version 4 (la 5 est en préparation pour 2017).

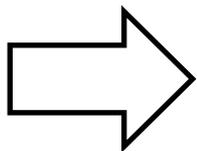
Définit des données d'accès et la façon de les calculer et de les présenter.

S'accompagne d'une obligation de certification pour les éditeurs compatibles.

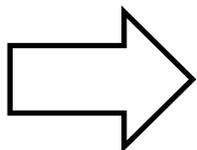
Les normes 2789 et 11620 (concernant le numérique)

- Des documents fort utiles... très peu connus et pratiqués (barrière payante)
- Cycles de révisions entre 3 et 6 ans : nécessité d'adapter régulièrement les normes aux pratiques.
- ISO 2789, de nombreux points à revoir. Ex.:
 - Les ebooks ne sont ajoutés que par achat
 - Décompte des refus d'accès uniquement pour les défauts d'accès simultanés
 - ...
 - Des décomptes fort bien gérés par le code COUNTER, qui gagnerait à mieux « infuser » dans ces documents
 - Une proposition (cycle de révision courant) pour inclure COUNTER en annexe de la norme ISO 2789

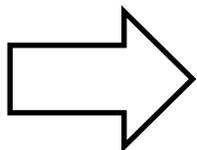
COUNTER, qu'est-ce que c'est?



Ce sont des rapports par type de ressource : revues, livres, bases de données, éléments multimédia, avec des rapports plus généraux (par plate-forme, ou les rapports « titre » tous supports confondus). On trouve des rapports sur l'utilisation (requêtes réussies) mais aussi sur les requêtes rejetées (par absence de licence)



Ce sont des règles de décompte : par ex., un article téléchargé 2 fois en PDF dans un intervalle de temps inférieur à 30 secondes, n'est compté qu'une fois.



Ce sont des garanties que d'un éditeur à l'autre, les données seront produites de la même manière et donc comparables, agrégeables, etc...

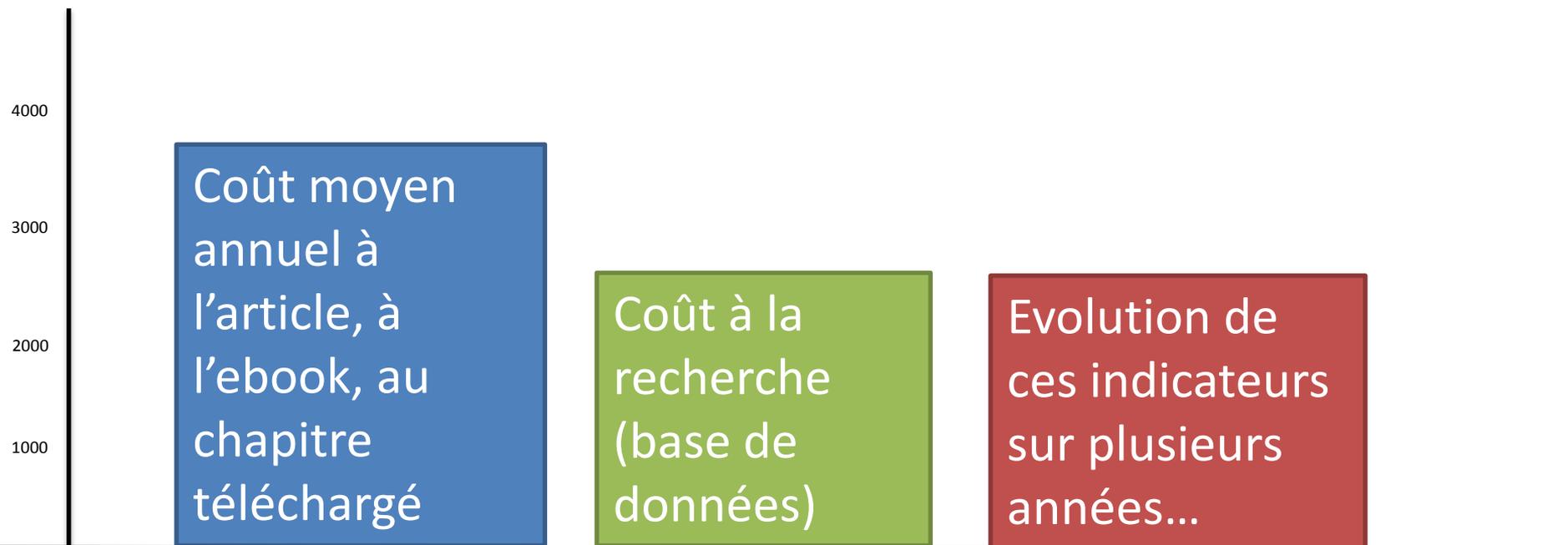
De nombreuses limites : sur le décompte des ebooks, sur l'évaluation circonscrite à la part souscrite ou possédée.

Un éditeur « certifié COUNTER » doit se faire auditer tous les ans (tous les 2 ans pour les petits). En France, trois éditeurs seulement ont fait cette démarche (EDP Sciences, Le Robert, Techniques de l'Ingénieur).

<https://www.projectcounter.org/about/register/>

Les données COUNTER, qu'en fait-on (généralement)?

On en retire des indicateurs et des graphes relativement simples

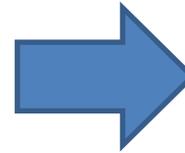


Il y a toutefois quelques limitations...

Exemple 1 : Daniel veut connaître le coût à l'article téléchargé d'un bouquet de revues électroniques (abonnement annuel).
Il dispose de rapports COUNTER.



Journal Report 1 :
Nombre d'articles
téléchargés par
mois et par revue



Montant de l'abonnement
/ Total JR1 = coût d'un
téléchargement pour un
bouquet.

Quid des articles en Open
Access (dit « gold ») de la
plate-forme?

Certaines revues ont été
achetées en licence
nationale...

Comment retirer les usages
non pertinents (formation
etc.)?

Il y a un rapport dédié
(JR1GOA)

Il faut enlever les titres et
les années achetés en
licence nationale...

Pas vraiment de solution
méthodologiquement
fiable.



Oui mais!

Exemple 2: Poulpy le documentaliste veut évaluer l'utilisation de ses **ebooks**.
Il dispose tout comme Daniel de rapports COUNTER.

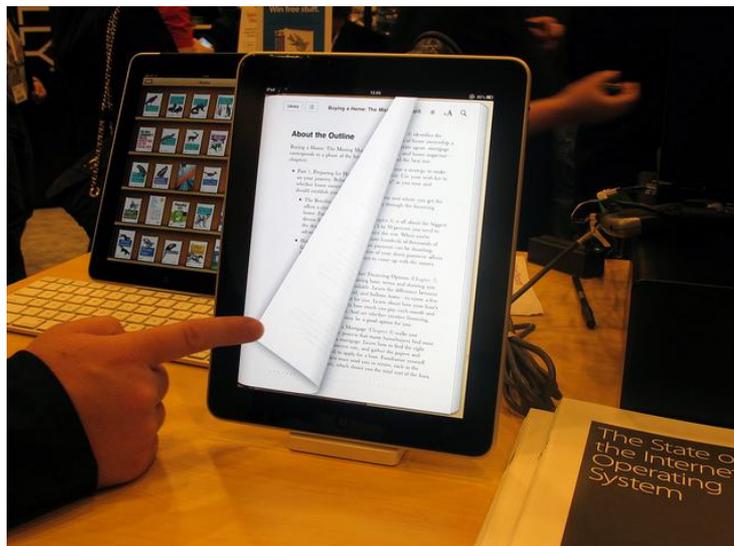
Deux rapports (COUNTER v4)



BR1 : nb total de
demandes
fructueuses de
titres par mois et
par titre



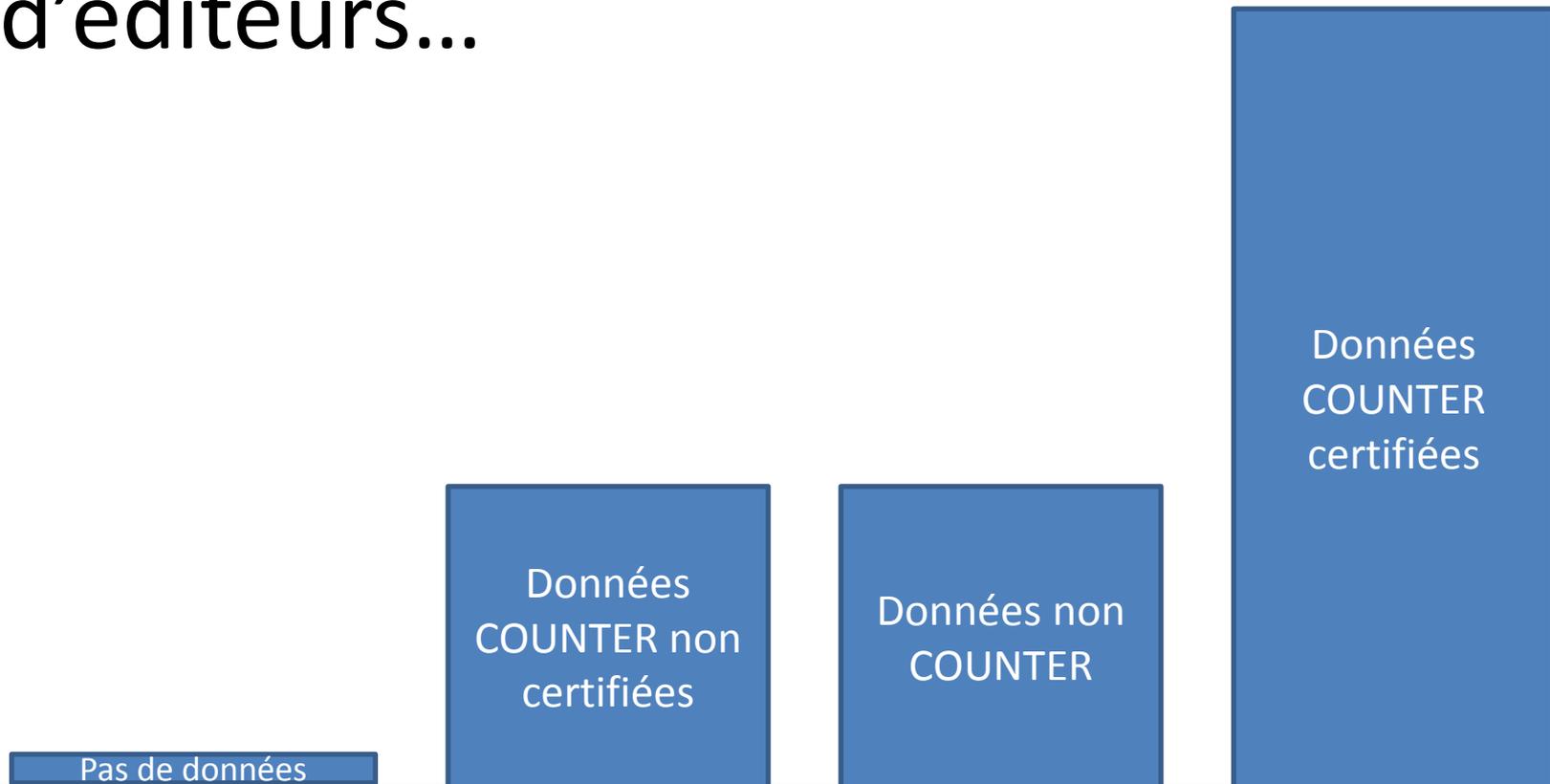
BR2 : nombre total
de demandes
fructueuses de
parties de livres par
mois et par titre



La fourniture ou non de ces rapports dépend de la présentation des livres sur la plate-forme (livre entier ou présentation au chapitre, plus courante).

Il est méthodologiquement délicat de les regrouper, même s'il est possible de suivre l'évolution d'un bouquet ou d'un titre précis dans le temps.

Si on devait faire quatre groupes d'éditeurs...



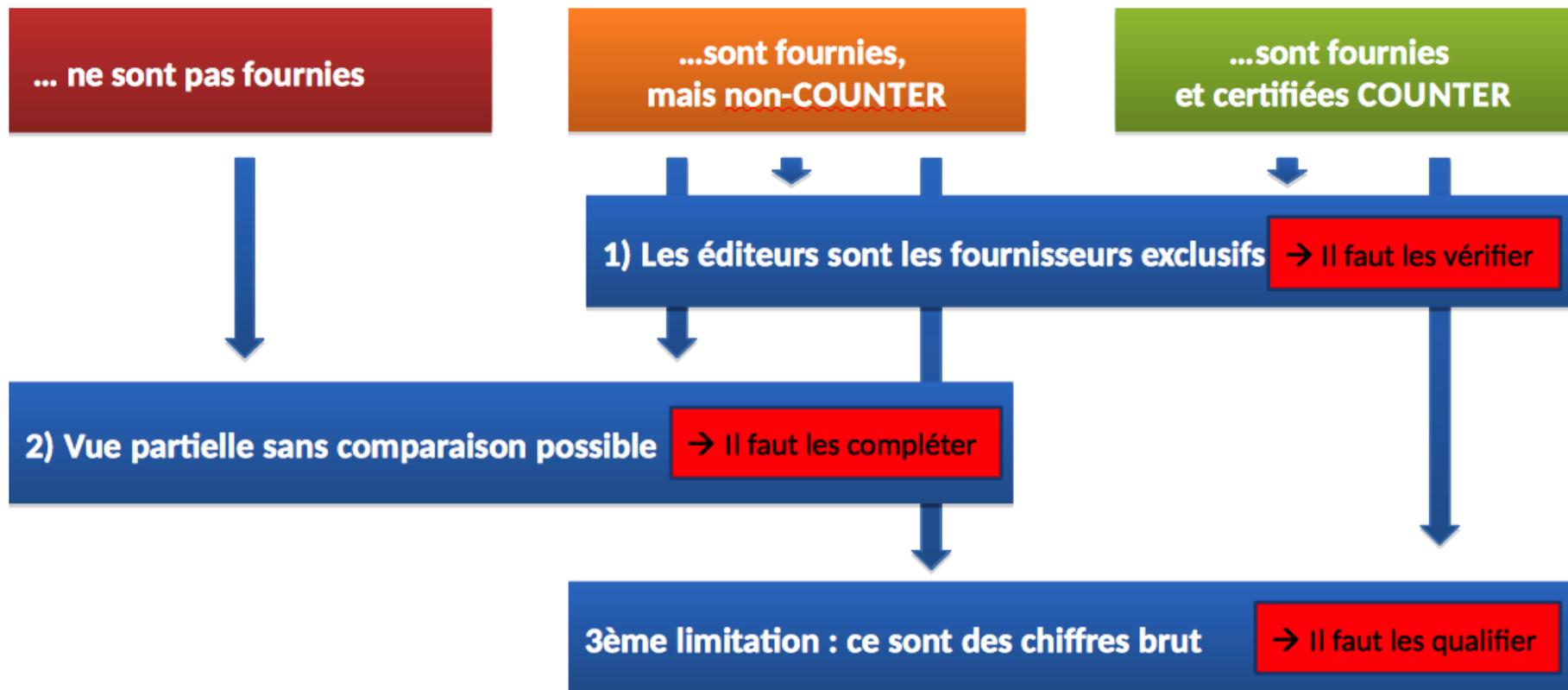
(concerne très peu de ressources aujourd'hui)

(beaucoup de Français...)

(beaucoup de produits en s. juridiques, économiques)

(la plupart des éditeurs académiques, part. anglo-saxons)

Et en pratique, ça donne...



Les dispositifs de collecte nationale

ESGBU : Enquête statistique générale auprès des services documentaires de l'enseignement supérieur

- Lancée pour la première fois en 1976
- A fait l'objet de recommandations pour révision (GT ADBU)
- Toujours un peu délicate à remplir, les chiffres collectés supposent d'additionner des chiffres COUNTER et non-COUNTER...
- Données désormais intégrées dans PAPESR

ERE : Enquête ressources électroniques

- Lancée pour la première fois en ??? (fin des années 1990)
- Collecte les données d'abonnement et les données tarifaires
- Gérée directement par le Ministère sur le même outil que l'ESGBU jusque vers 2012 ; Couperin recueille et analyse les chiffres à compter de 2007
- 2014... : reprise complète de la gestion par Couperin, lancement d'une nouvelle interface.

La dispersion de l'usage et le projet DUL

- Les considérations qui précèdent tiennent pour acquis que l'usage des ressources se concentre sur les portails éditeurs, or c'est le cas le plus simple
- Un grand nombre des consultations peut aujourd'hui se déporter, pour un même article, sur d'autres « lieux » : réservoirs en OA, réseaux sociaux de chercheurs
- Projet CrossRef/COUNTER, « Distributed Usage Logging ». <http://blog.crossref.org/2015/12/private-channel-dul.html>
 - Utiliser le DOI comme donnée « pivot » pour permettre un décompte de cet usage dispersé.

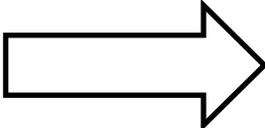
Comment avancer?

Le périmètre des données disponibles est stable depuis de nombreuses années : nombre de consultations d'un document donné, coût des abonnements, niveau d'utilisation...

Les indicateurs qui en dérivent (coût à l'article etc.) sont donc eux-mêmes en nombre limités et laissent peu de place à l'innovation

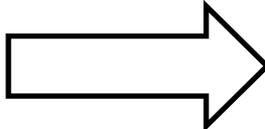


Des **données** peu exploitées en lien avec celles-ci : celles issues des systèmes d'information des établissements (annuaires, données d'utilisation des services numériques, enseignements des webmetrics, statistiques de réussite étudiante...)



Des **périmètres** à (re)conquérir :

- De **collection** : évaluer réellement une collection revient à délimiter le plus précisément possible le corpus réellement payé
- **D'abonnés, d'utilisateurs** : évaluer l'usage à l'échelle nationale, ou a contrario par une partie donnée des usagers d'un établissement



Des **croisements** à imaginer : avec les données de la scolarité, les grands indicateurs nationaux de la réussite étudiante, Libqual+...

MESURE, ezPAARSE, ezMESURE...

II. OUTILS COMMUNS, COLLABORATIONS

MESURE :

collecte COUNTER automatisée

- Les éditeurs certifiés **COUNTER** fournissent tous le protocole **SUSHI** (norme NISO, Standardized Usage Statistics Harvesting Initiative).
- Le **JUSP** (GB) moissonne depuis des années des rapports COUNTER (JR1) pour le compte des bibliothèques académiques anglaises
- **MESURE** = adaptation du portail JUSP opérée sous l'égide de Couperin.

<http://measure.couperin.org>

MESURE : collecte COUNTER automatisée

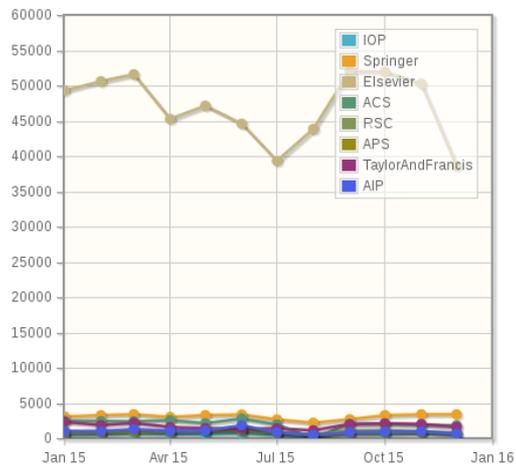


couperin.org
Consortium Universitaire de Périodiques Numériques

- [Résumé annuel](#)
- [Détail / éditeur](#)
- [Evolution annuelle / éditeur](#)
- [Moissonnage](#)
- [Feuille de route](#)
- [Contact](#)

Année :

Résumé annuel pour Université de Lorraine en 2015

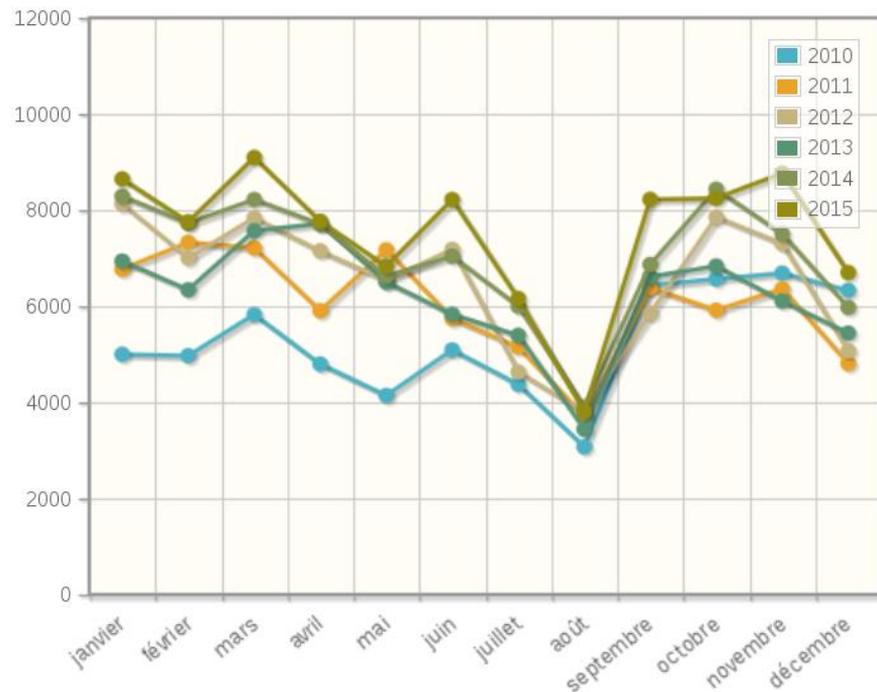
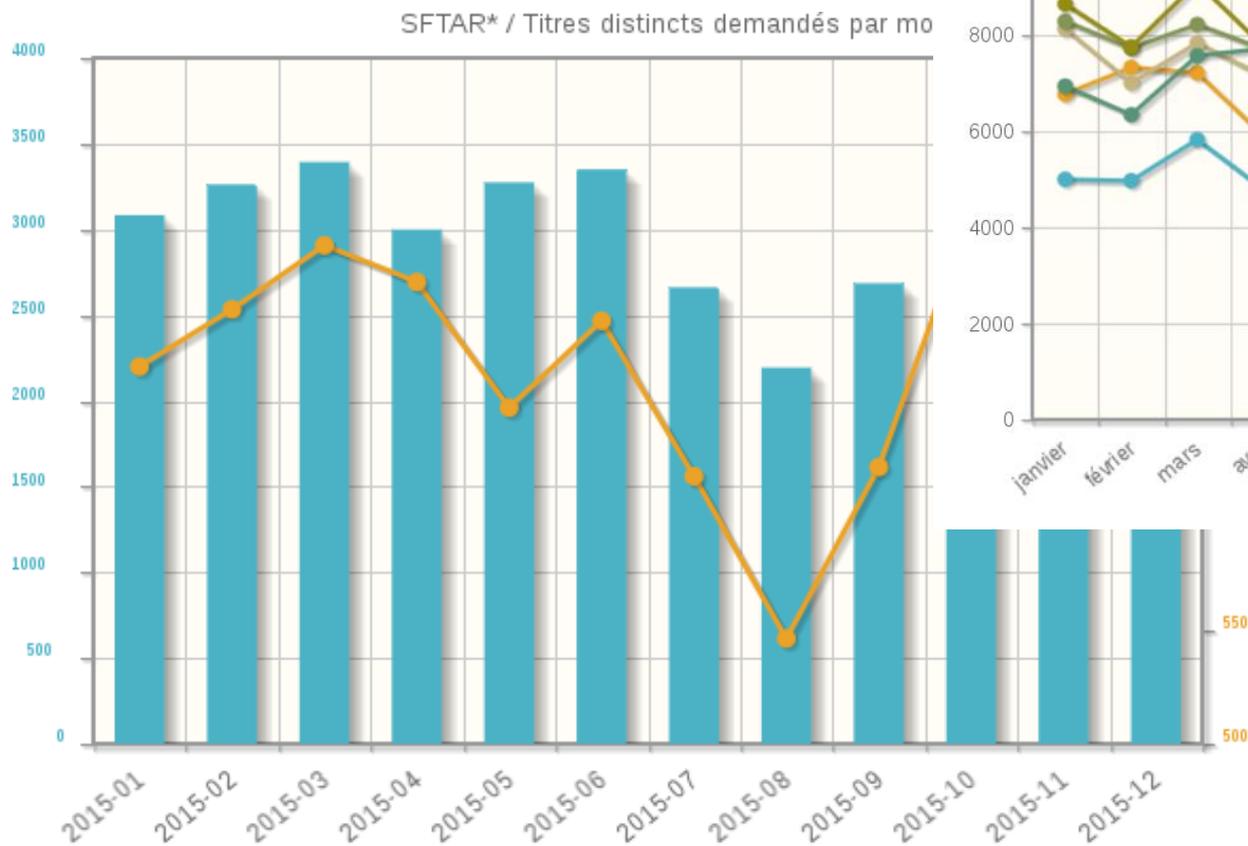


Les 10 titres aux SFTAR les plus élevés en 2015 (source : JR1)

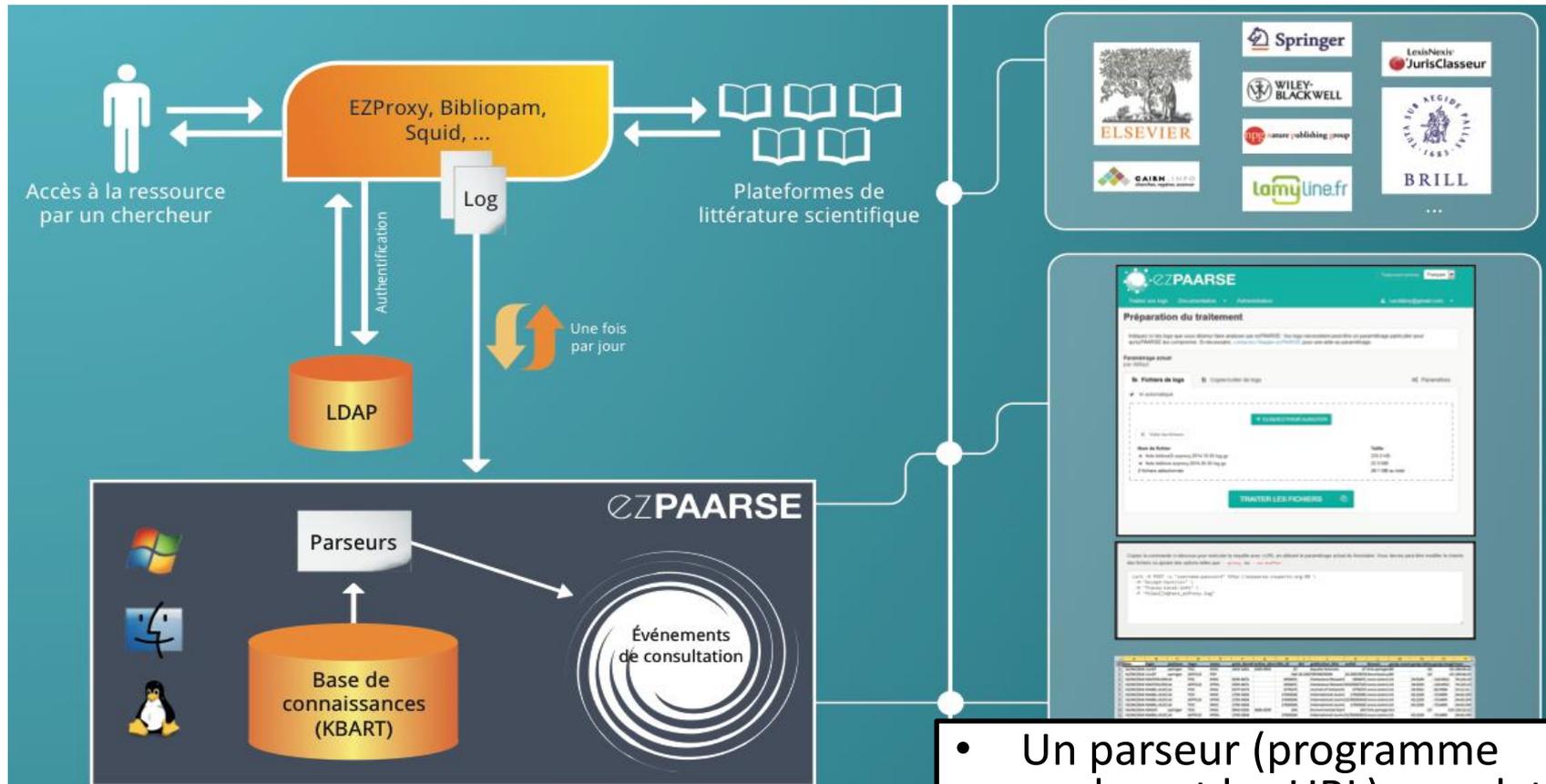
Titre	Fournisseur	Accès
Annales Médico-psychologiques, revue psychiatrique	Elsevier	5309
Materials Science and Engineering: A	Elsevier	5222
Acta Materialia	Elsevier	5067
Food Chemistry	Elsevier	4470
Neuropsychiatrie de l'Enfance et de l'Adolescence	Elsevier	4015
Physical Review B	APS	3923
Carbohydrate Polymers	Elsevier	3746
Archives de Pédiatrie	Elsevier	3566
Chemical Engineering Science	Elsevier	3449
Physical Review Letters	APS	3243

Editeur/Mois	01-2015	02-2015	03-2015	04-2015	05-2015	06-2015	07-2015	08-2015	09-2015	10-2015
IOP	453	514	618	551	522	642	427	303	624	544
Springer	3083	3262	3393	2999	3274	3350	2662	2194	2686	3260
Elsevier	49298	50598	51579	45252	47125	44582	39315	43835	52021	51900
ACS	2490	2420	2395	2535	2126	2776	1946	0	2038	2030
RSC	1048	1009	1134	1087	1105	1061	794	531	1095	1170
APS	694	786	898	875	1012	1236	772	467	732	808
TaylorAndFrancis	2333	1863	2176	1578	1417	1322	1406	1126	2002	2110

MESURE : collecte COUNTER automatisée



ezPAARSE : des données produites en local

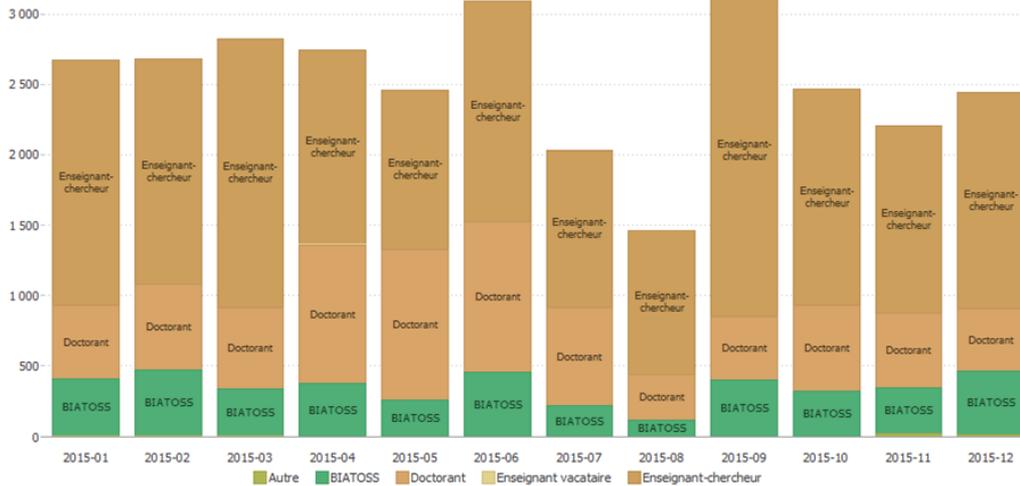


<http://ezparse.couperin.org> (instance globale)
<http://ang.couperin.org> (analyses collaboratives)

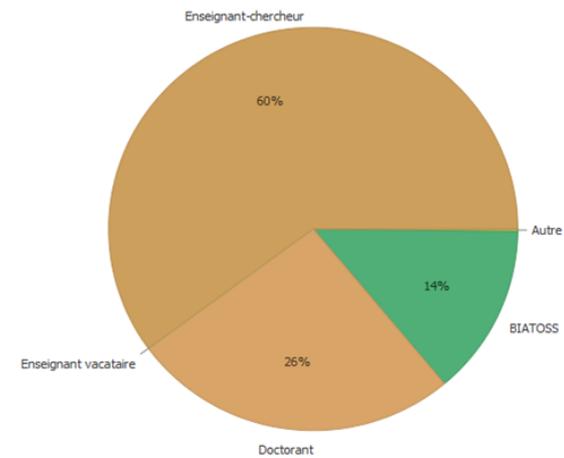
- Un parseur (programme analysant les URL) par plateforme
- Une mise à jour des parseurs et des analyses collaboratives

Exemple ezPAARSE : zoom sur un labo en sciences (UL)

Consultations d'articles par type d'utilisateur et par mois



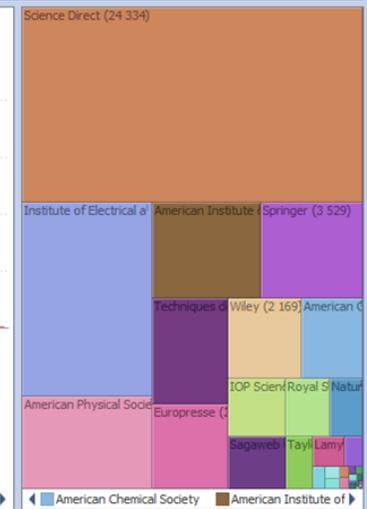
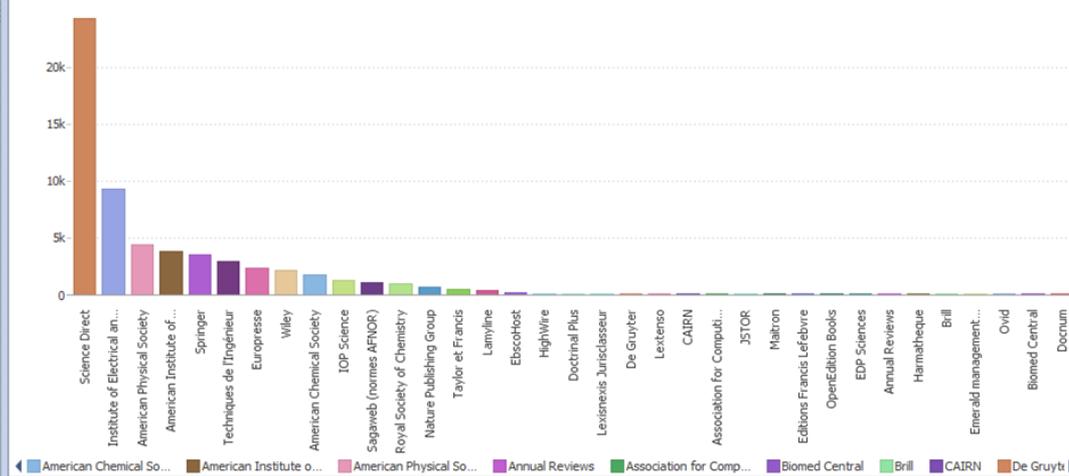
Consultations d'articles par type d'utilisateur



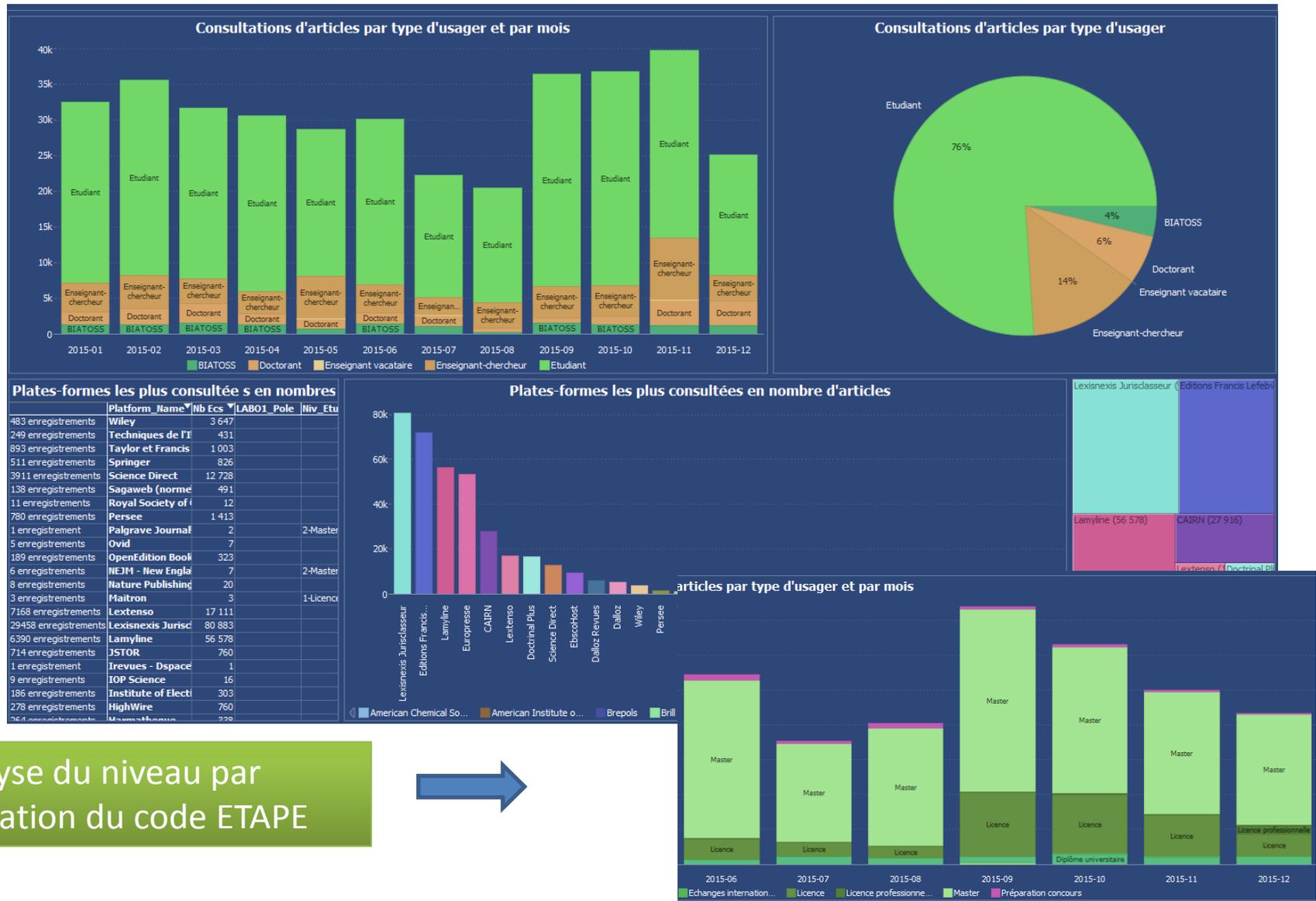
Plates-formes les plus consultées

enregistrements	Platform Name	Nb Ecs	Co
13345	Science Direct	24 334	
4660	Institute of Electrical an...	9 309	
1582	American Physical Society	4 374	
1185	American Institute of Physics	3 816	
2385	Springer	3 529	
1601	Techniques de l'Ingénieur	2 914	
477	Europresse	2 324	
1107	Wiley	2 169	
1198	American Chemical Society	1 785	
714	IOP Science	1 240	
388	Sagaweb (normes AFNOR)	1 090	
637	Royal Society of Chemistry	954	
191	Nature Publishing Group	703	
438	Taylor et Francis	505	
55	Lamyline	354	
141	EbscoHost	196	
69	HighWire	95	
43	Doctrinal Plus	70	
15	Lexisnexis Jurisclass	51	
26	De Gruyter	37	
11	Lextenso	35	
12	CAIRN	22	
8	Association for Computing Machinery	19	
18	JSTOR	19	
8	Maitron	13	
7	Editions Francis Lefebvre	9	

Plates-formes les plus consultées en nombre d'articles



Exemple ezPAARSE : zoom sur les UFR DEG (UL)



Analyse du niveau par utilisation du code ETAPE



ezPAARSE > ezMESURE

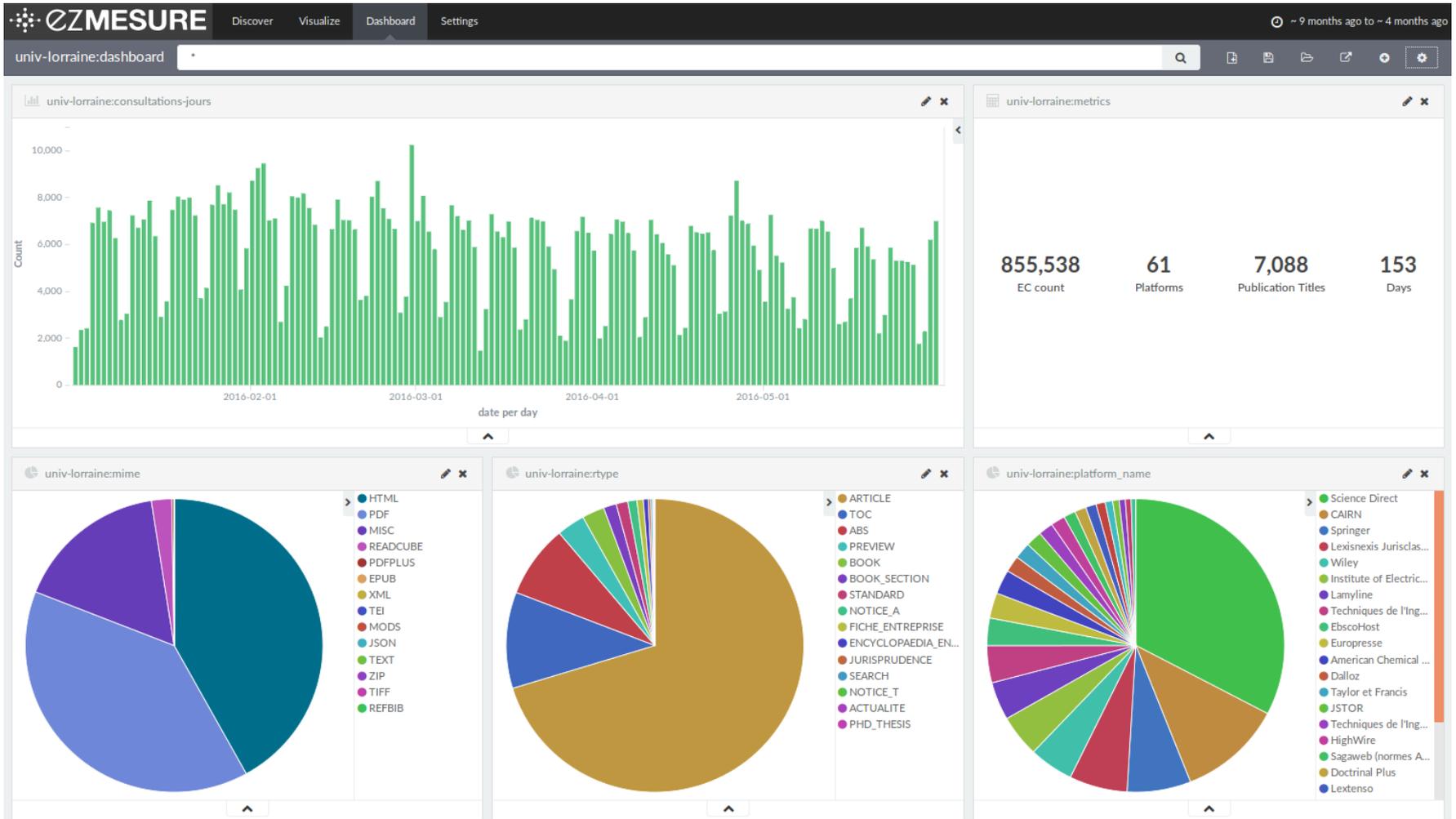


Le but : construire un entrepôt national des données d'usage produites par les différentes instances d'ezPAARSE installées et gérées par un nombre croissants de membres de Couperin

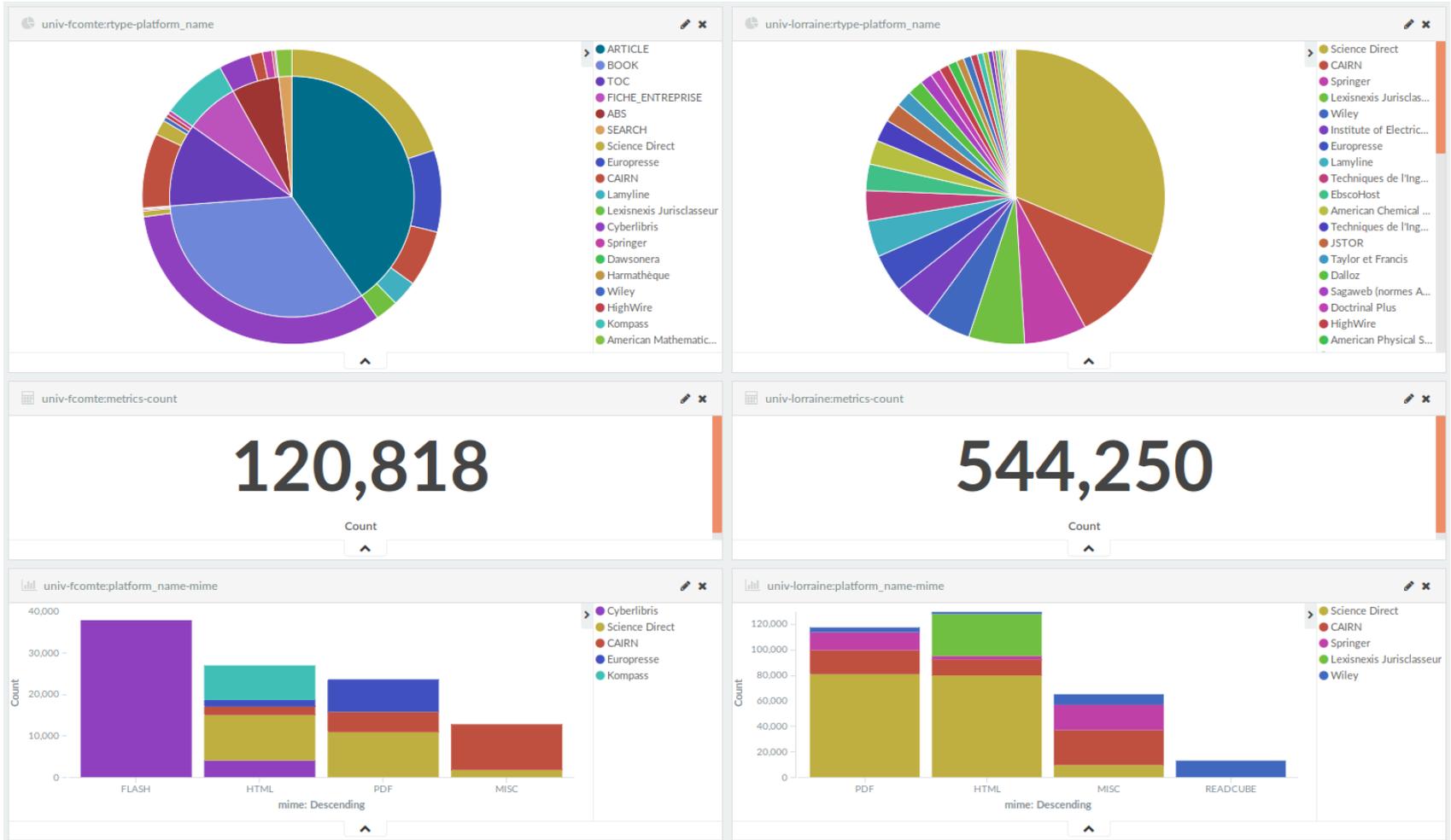
- des données fiables et indépendantes de la fourniture par les éditeurs
- des visualisations graphiques partagées et dynamiques
- des indicateurs implémentés à différents niveaux (beaucoup de macro)



ezMESURE : une institution...



... deux institutions...



... trois institutions (et ce n'est pas fini!)



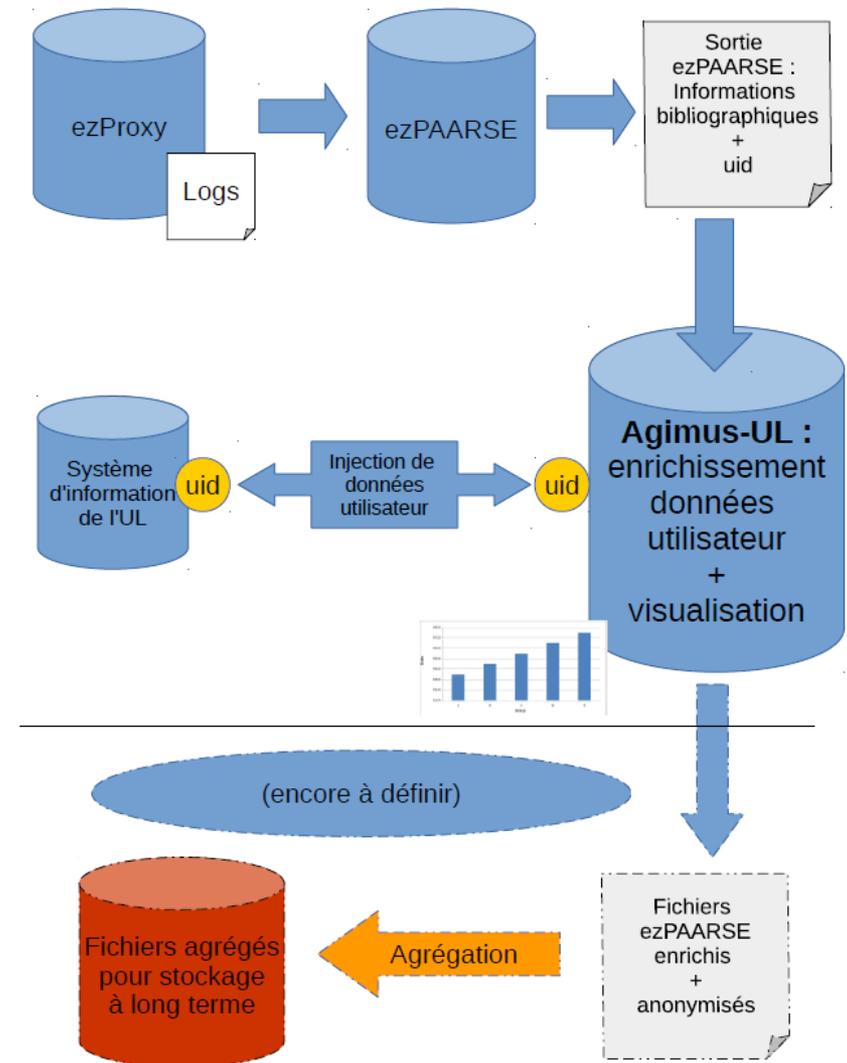
ezAGIMUS = AGIMUS-NG + ezPAARSE

AGIMUS-NG est une plateforme qui permet d'utiliser les fichiers de traces et le système d'information d'un établissement (annuaire LDAP, base de données, ...) afin de produire un entrepôt de données pour produire des indicateurs statistiques d'usage des services numériques.

Pour ce faire, la plateforme ingère des logs, les enrichit avec des informations du SI, les stocke dans un entrepôt et permet de créer des tableaux de bord de rendu visuel (Kibana).

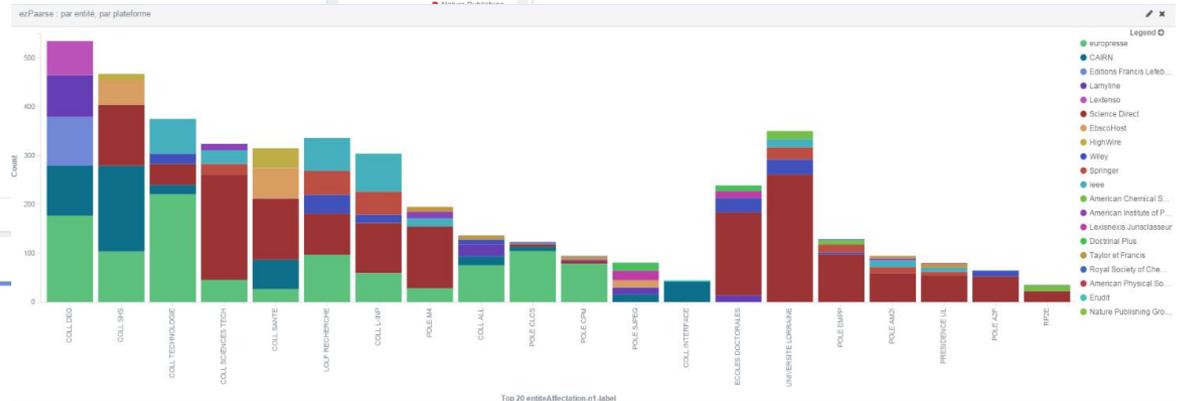
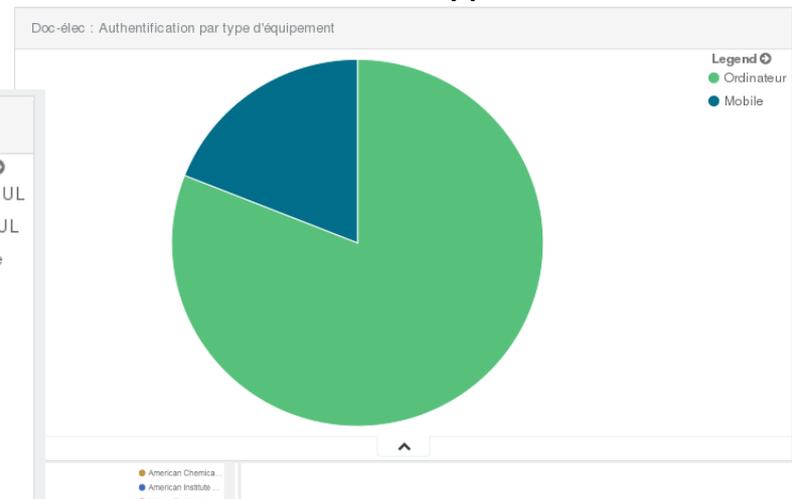
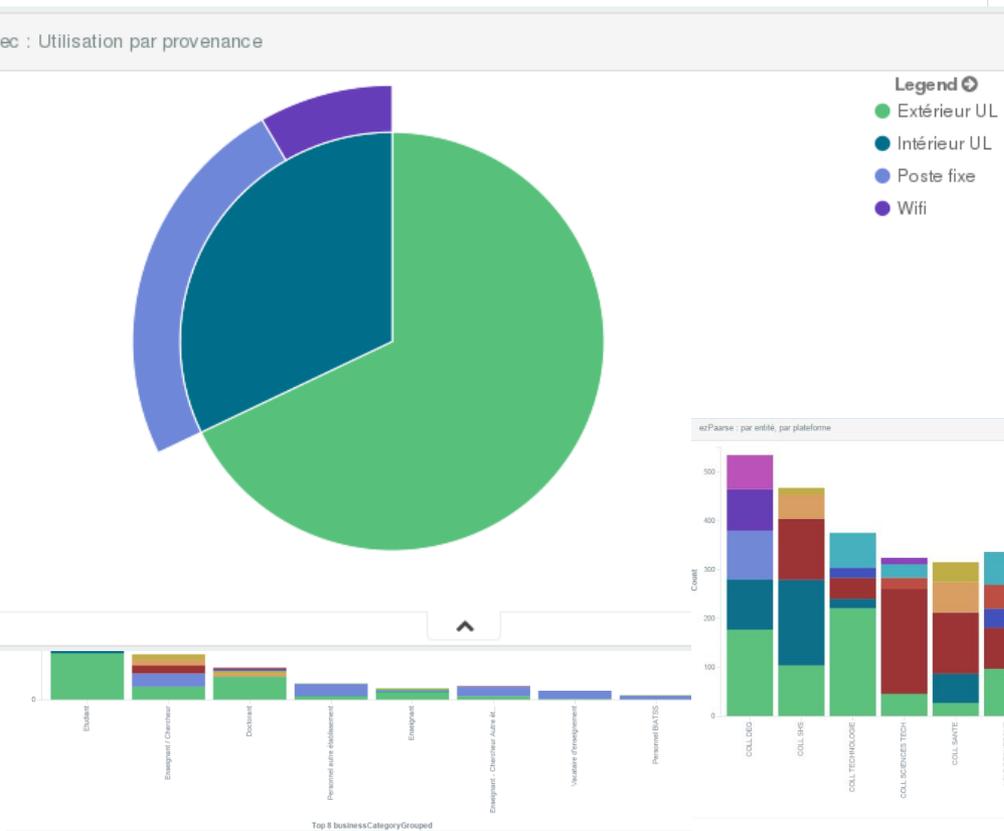
Voir <https://www.esup-portail.org/wiki/display/AGIMUSNG/Documentations>

ezAGIMUS est le nom (officiel) d'un essai en cours de branchement des sorties ezPAARSE sur AGIMUS-NG.



ezAGIMUS = AGIMUS-NG + ezPAARSE

- Il est possible de disposer d'un outil simple et performant des données ezPAARSE
- Il est possible de croiser les données ezPAARSE avec d'autres données comme par exemple la répartition entre accès interne / extérieur ; le type de terminal (PC ou mobile).



Quelques groupes de travail nationaux...

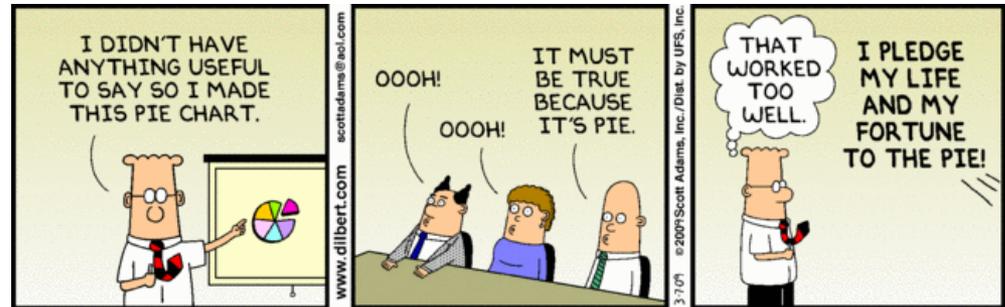
- Couperin : groupe de travail Indicateurs (2014-...)
- ADBU : Commission Evaluation
- EPRIST : Groupe de travail Usages ressources numériques

Quelques travaux du GTI

- Promouvoir l'usage de COUNTER auprès des éditeurs français
 - Participation au comité exécutif de COUNTER
 - Traduction de plusieurs documents COUNTER (mis à disposition sur le site du Projet, <http://www.projectcounter.org>)
 - Mise en place d'une solution de certification en France (société Dekra)
 - Participation à venir à l'évaluation collaborative de Counter v5
- Travail sur les normes : propositions d'évolution
- « Mettre en musique » les sources de données existantes : projet de plateforme statistique nationale...?
- Une journée d'étude à venir (juin 2017)

A garder à l'esprit...

Produire un camembert est facile...



... mais attention à la composition!

