

Deux exemples d'implantation d'ezPAARSE :

INIST, Université de Lorraine

CÉCILIA FABRY (INIST)

THOMAS JOUNEAU (UL)



Plan

- Mise en place et suivi du projet
- Exploitation et résultats



Deux exemples d'implantation d'ezPAARSE : INIST, UL

MISE EN PLACE ET SUIVI DU PROJET

Le contexte INIST - CNRS

- L'INIST (en tant d'opérateur technique des portails du CNRS) dispose :
 - De logs de consultations grâce au reverse proxy
 - Des identifications des unités utilisatrices
- Création de **NUM_STAT** en 2009 (suite de portail_stat)
 - Base des logs traités par les parseurs écrits en PHP (petits programmes qui déchiffrent les logs bruts pour en sortir les informations importantes : plateformes; revues; issn; type de consultation; date; logins qui consultent)
 - Les indicateurs produits à partir de NUM_STAT suscitent un vif intérêt chez tous les gestionnaires de ressources électroniques lors des présentations externes mais cet outil est développé pour les données INIST et n'est pas “exportable” ailleurs
 - La maintenance de l'applicatif au niveau INIST-CNRS est très lourde : souvent dans le rouge (réécriture des parseurs, maintenance des analyses etc...)
 - Il n'existe pas de logiciel tiers qui réponde complètement à ce besoin spécifique



Le contexte UL (1)

- Fusion en 2012 : travail sur un portefeuille d'abonnements électroniques commun, évoluant en fonction de plusieurs critères :
- La politique documentaire, guidée par les besoins documentaires en recherche et formation
 - L'évolution des crédits (coupes budgétaires entre 2012 et 2013)
 - L'usage fait des ressources payées ou testées
- Manque de tableaux de bord précis sur ce dernier critère:
 - Statistiques éditeur incomplètes, sujettes à caution
 - Manque de mesure de l'usage par faculté ou composante, ennuyeux pour établissement de cette taille



Le contexte UL (2)

- ezProxy opérationnel dès octobre 2009
 - Trois puis quatre bouquets d'abonnements (INPL, UHP, N2 ; Metz dès la fin 2011) sur une même installation avec IP de sortie différente
 - Gestion par ezProxy des accès locaux et distants
- Recueil sommaire de statistiques de « sessions »
 - Nombre d'occurrences du domaine racine (déclaré dans EZP) dans les logs
 - Chiffres ne recoupant pas du tout ceux des éditeurs et au final peu fiables...



INIST et UL : Besoin de plus de précision et de réactivité...

- Qui consulte les ressources juridiques en-dehors des juristes?
 - Quelle est la part de consultation des laboratoires sur les portails de revues?
 - Quelle est la part d'utilisation des ressources pour la recherche, la formation?
 - Est-ce que la ressource très onéreuse et demandée par un petit nombre de chercheurs est réellement utilisée?
 - Comment analyser, améliorer la réactivité face aux pics de consultation?
 - Etc.
-
- Ces données sont extrêmement sensibles dans le cas de certaines ressources négociées.
 - La production en local apparaît comme le moyen le plus sûr de les générer.



Implantation pilote d'ezPAARSE à l'UL (1)

- Installation
 - L'utilisation en amont d'ezProxy pour l'ensemble des accès a évité toute adaptation préalable
 - Le suivi de l'installation, de la maintenance et du fonctionnement est assuré par l'administrateur ezProxy
- Participation active (l'administrateur ezProxy + 1 bibliothécaire) à l'équipe projet
 - Collecte des données, test des parseurs mis en place,
 - Travail en parallèle sur l'analyse de plates-formes préalable à l'écriture des parseurs, avec une concentration particulière sur l'analyse des plates-formes juridiques.

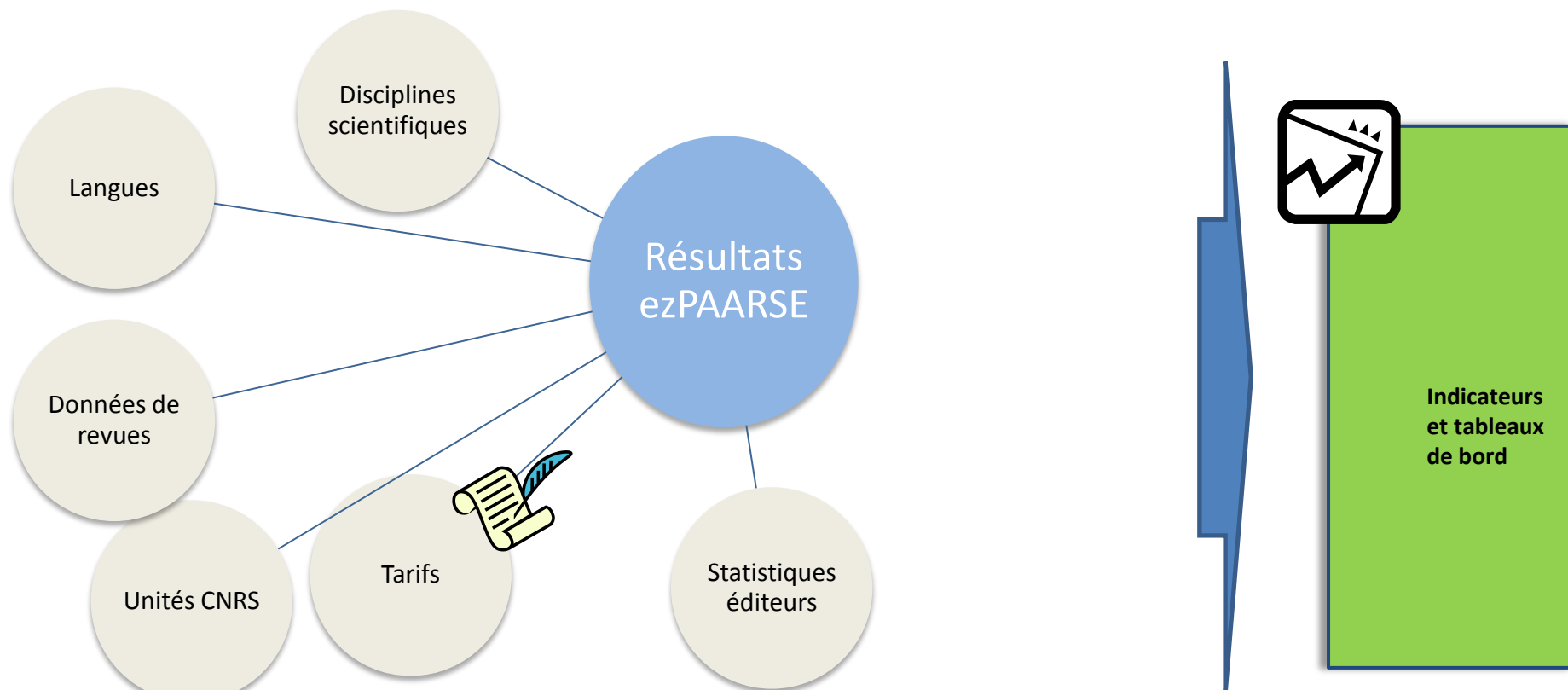


Implantation pilote d'ezPAARSE à l'UL (2)

- Travail transversal avec la direction du numérique pour l'utilisation des référentiels utilisateurs
- Traitement final des fichiers assuré au moyen du logiciel Omnisciope, de la société Visokio.
 - D'autres logiciels (Qlikview...) peuvent être utilisés : manque pour le moment d'expertise dans leur utilisation
 - D'autres solutions (développement de programmes, scripts ou procédures en interne) devraient idéalement pouvoir être considérées et envisagées



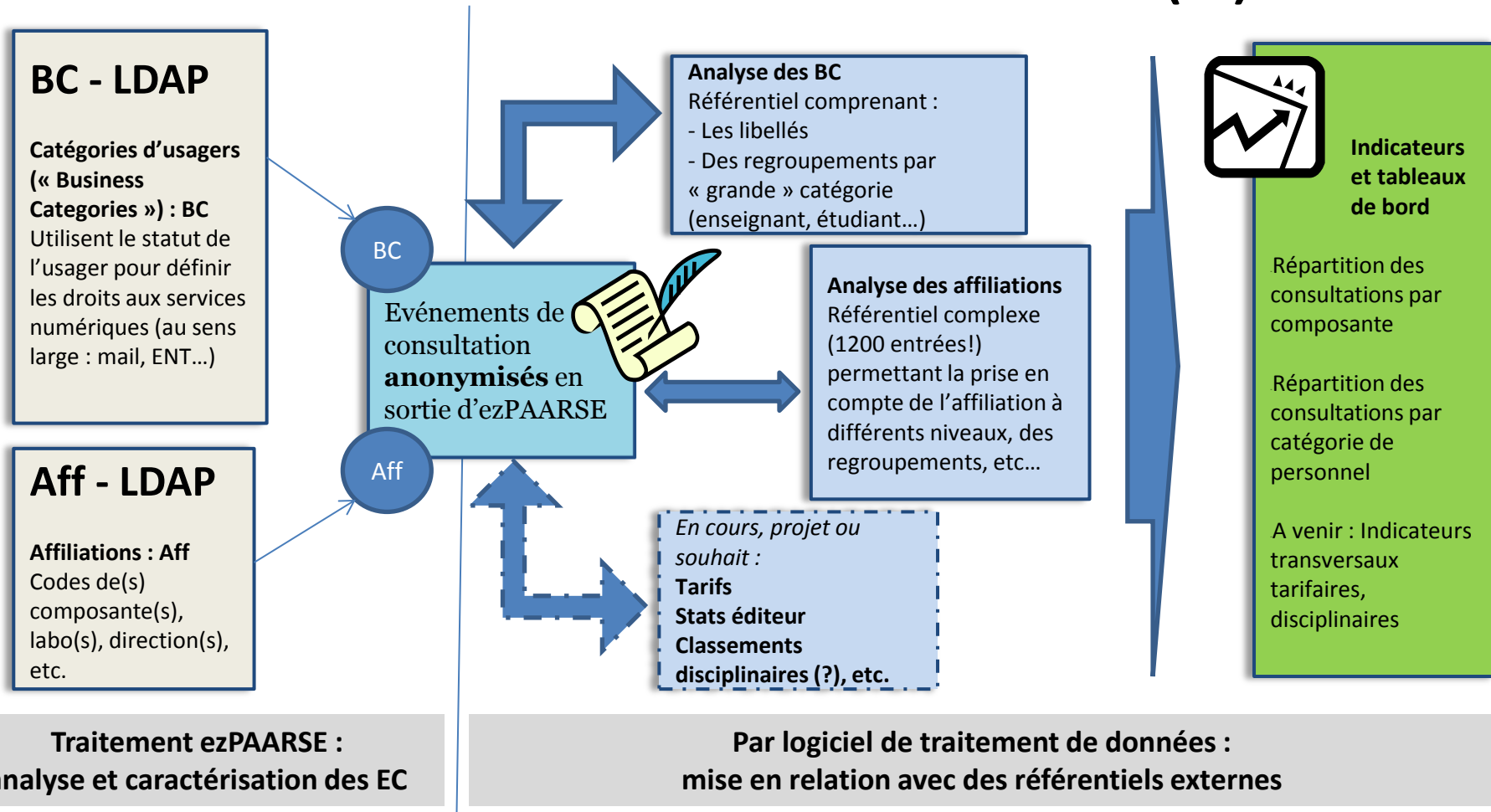
Enrichissement et traitement des données ezPAARSE : à l'INIST



**Par logiciel de traitement de données :
mise en relation avec des référentiels externes**



Enrichissement et traitement des données ezPAARSE : à l'UL (1)



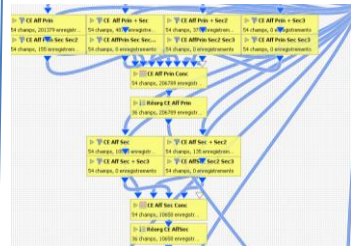
Enrichissement et traitement des données ezPAARSE : à l'UL (2)

Format du fichier en sortie d'ezPAARSE, enrichi d'informations relatives au login (23 champs)

Domain
Rtype
Mime
Doi
Pissn
Eissn
Title
platform
Unitid
datetime
Login
Host
session
Date
url
Status
Size
User
BCPrin
ABC
AffPrin
AAff
cas

Traitement ezPAARSE :
analyse et caractérisation des EC

Format du fichier de données après mise en relation avec les référentiels externes (54 champs)



Par logiciel de traitement de données :
mise en relation avec des référentiels externes

Domain	AffPrinc Code Mere
Rtype	AffPrinc Libelle Court
Mime	AffPrinc Libelle Long
Doi	AffPrinc Libelle Type Structure
Pissn	AffPrinc Type Structure
Eissn	AffSec 2 Code Mere
Title	AffSec 2 Libelle Court
Platform	AffSec 2 Libelle Long
Unitid	AffSec 2 Libelle Type Structure
Datetime	AffSec 2 Type Structure
Login	AffSec 3 Code Mere
Host	AffSec 3 Libelle Court
Session	AffSec 3 Libelle Long
Date	AffSec 3 Libelle Type Structure
Url	AffSec 3 Type Structure
Status	AffSec Code Mere
Size	AffSec Libelle Court
User	AffSec Libelle Long
BCPrinc	AffSec Libelle Type Structure
BCSec	AffSec Type Structure
Sec 2	BCPrin Nomcourt
Princ	BCPrin Regroupement
Sec	BCSec Nomcourt
Sec 2	BCSec Regroupement
Sec 3	BCSec 2 Nomcourt
s particuliers	BCSec 2 Regroupement
is	Nom Plate-forme



Deux exemples d'implantation d'ezPAARSE : INIST, UL

EXPLOITATION ET RÉSULTATS

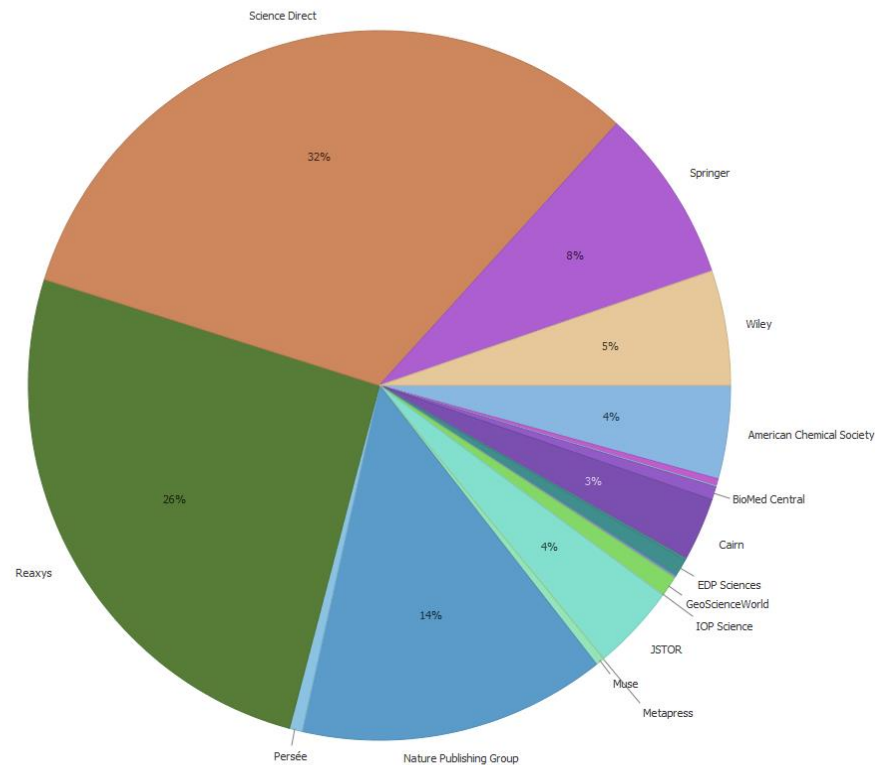
Enrichissement et traitement des données : il importe de noter que...

- Ces deux exemples ne sont que des illustrations de ce qu'il est possible de faire
- Les informations utilisables pour la caractérisation des EC vont étroitement dépendre des outils et ressources disponibles au niveau de chaque établissement
- Le rapatriement de ces informations peut se faire à plusieurs niveaux :
 - UL : injection des codes seuls dans la sortie ezPAARSE, puis enrichissement à partir de ces codes dans le logiciel de traitement de données
 - INIST : dans le logiciel de traitement de données à partir (not.) du login de l'unité



INIST : Répartition des consultations par plate-forme tous portails confondus

Subdiviser: Aucun

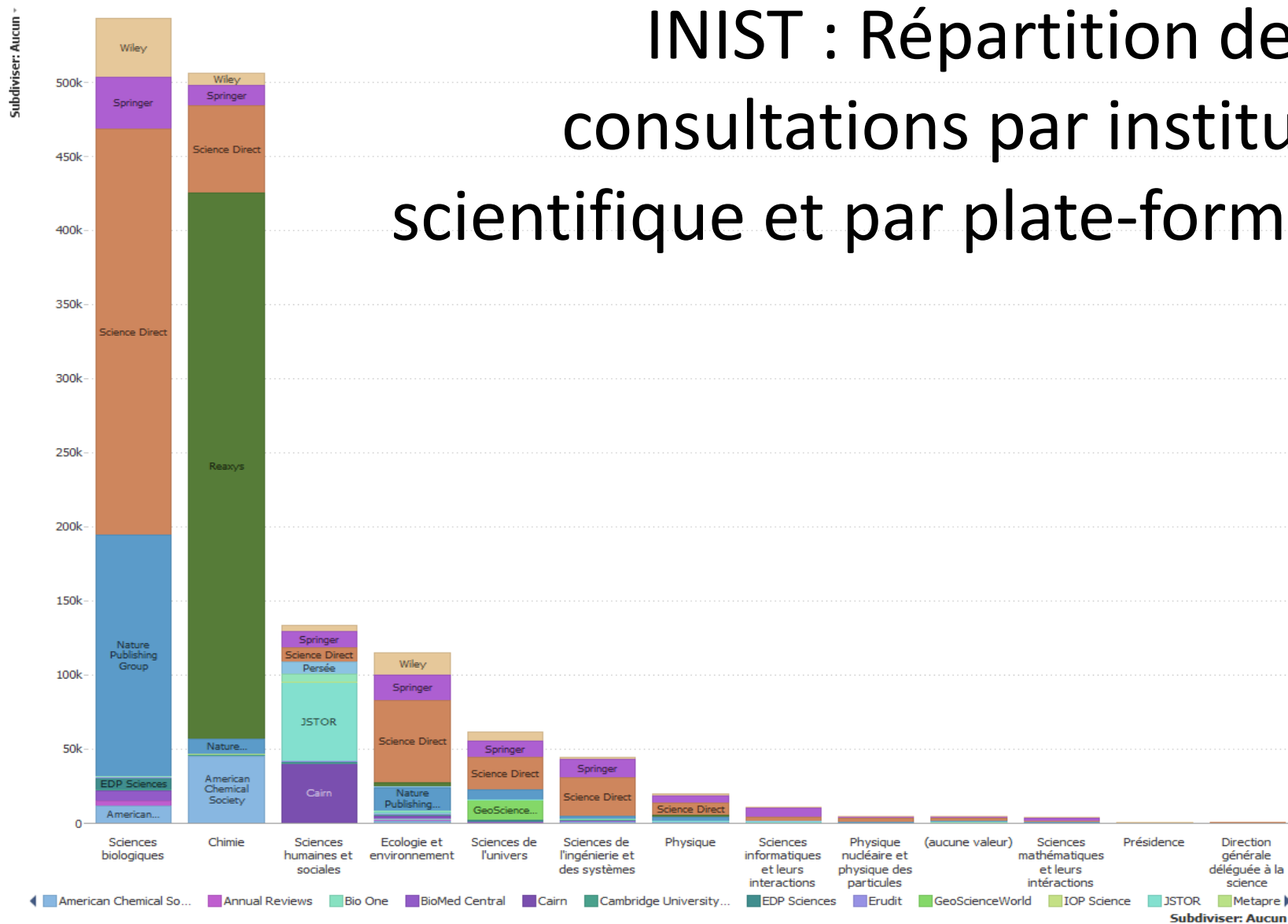


Subdiviser: Aucun

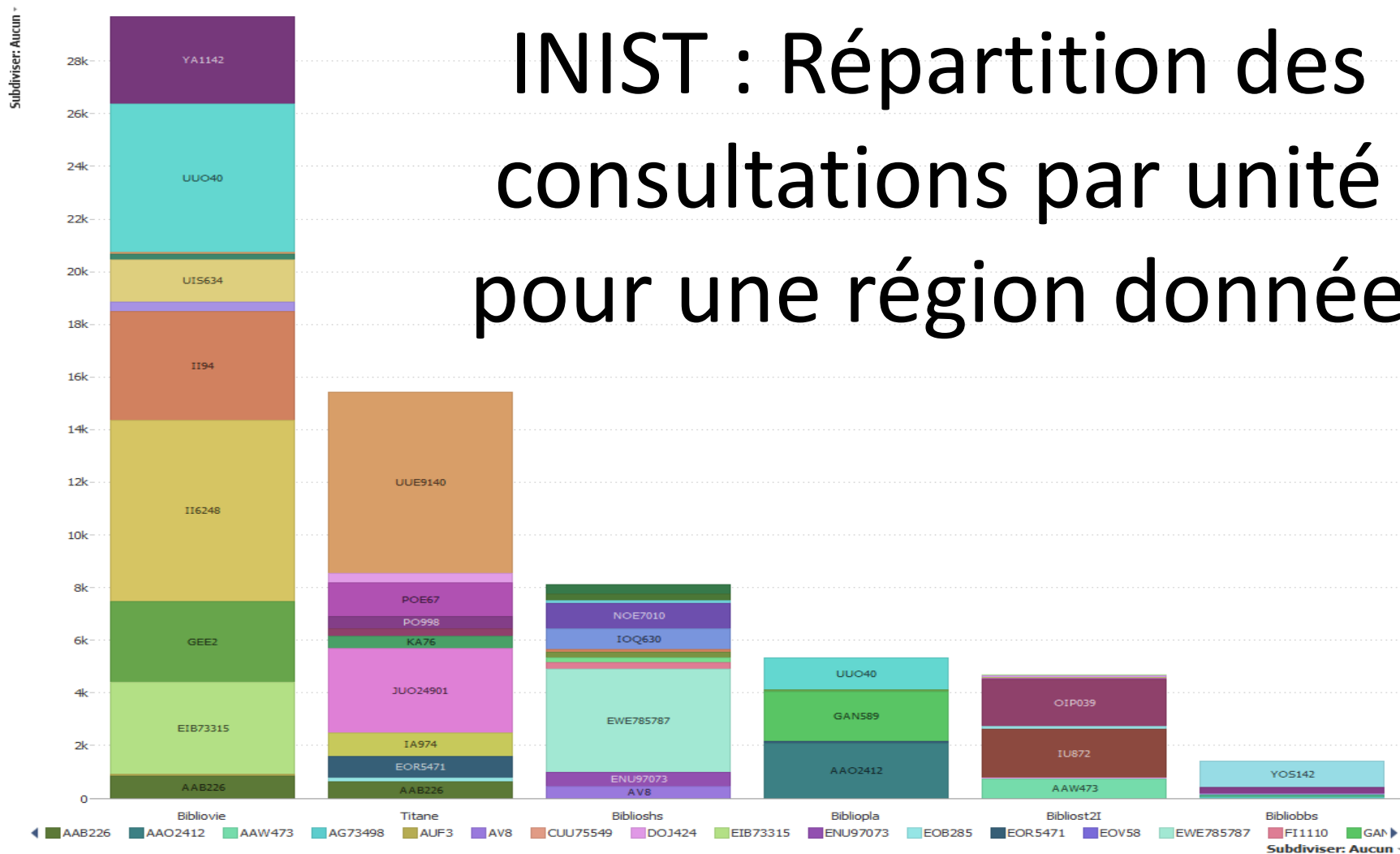


couperin.org

INIST : Répartition des consultations par institut scientifique et par plate-forme

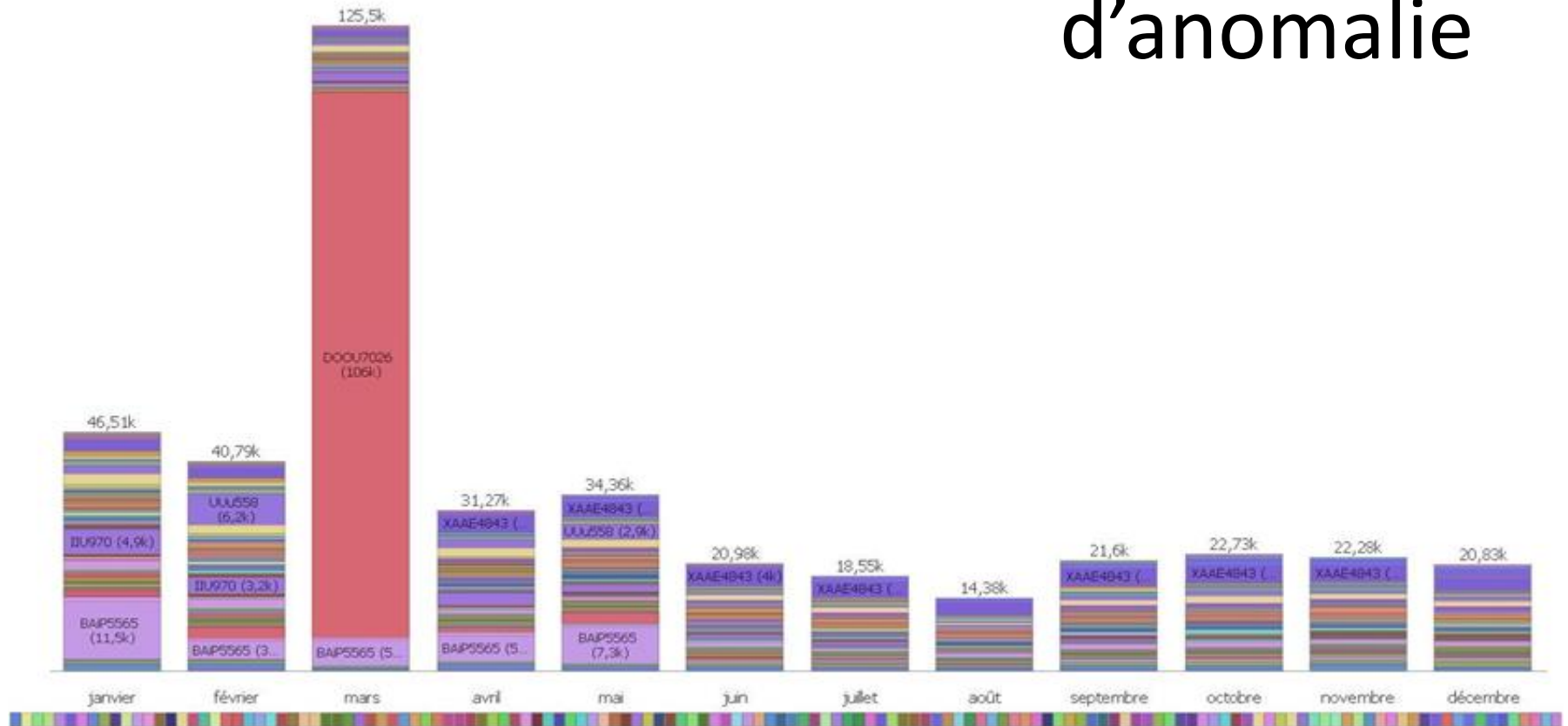


INIST : Répartition des consultations par unité pour une région donnée

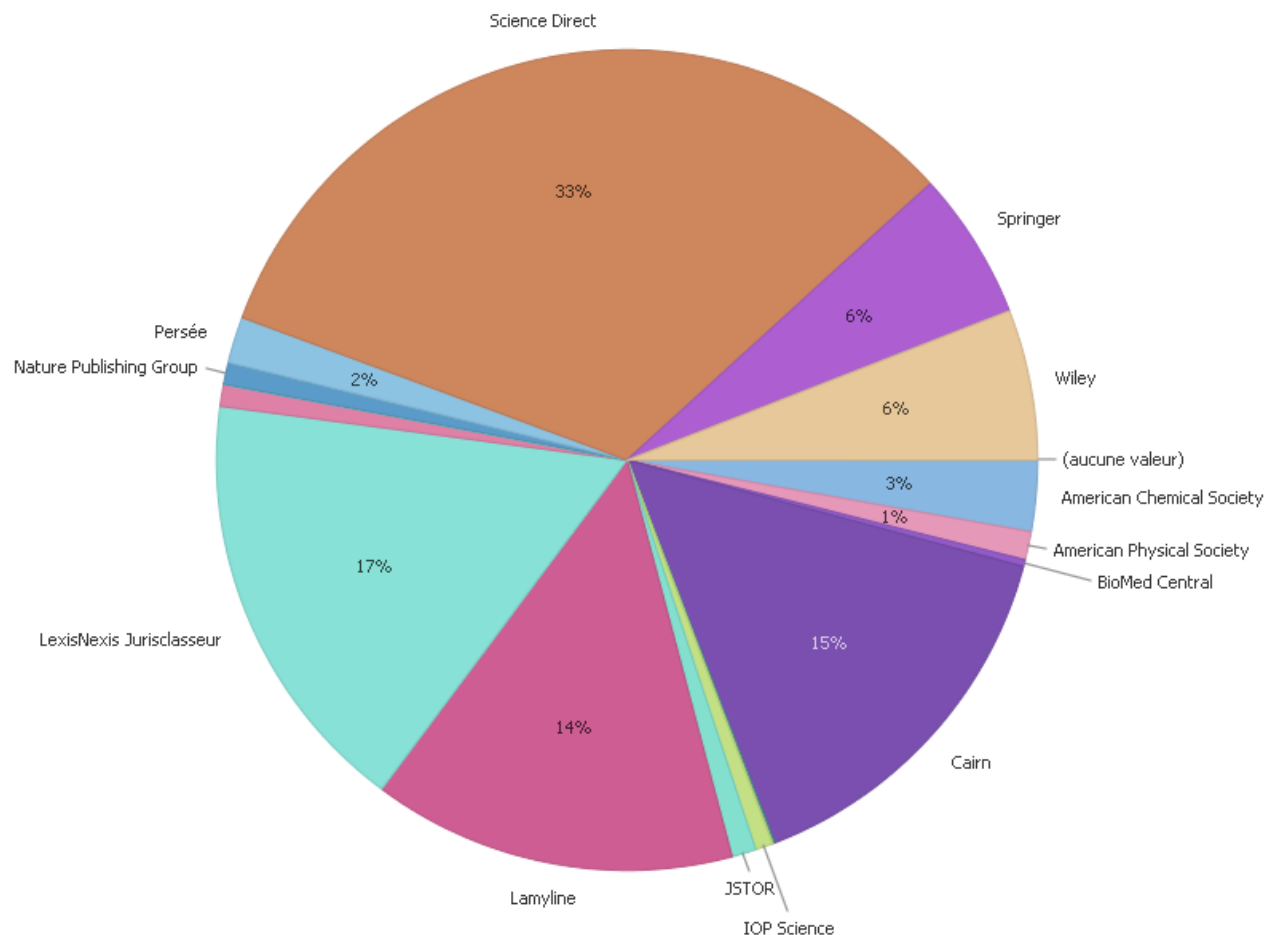


INIST : exemple de détection d'anomalie

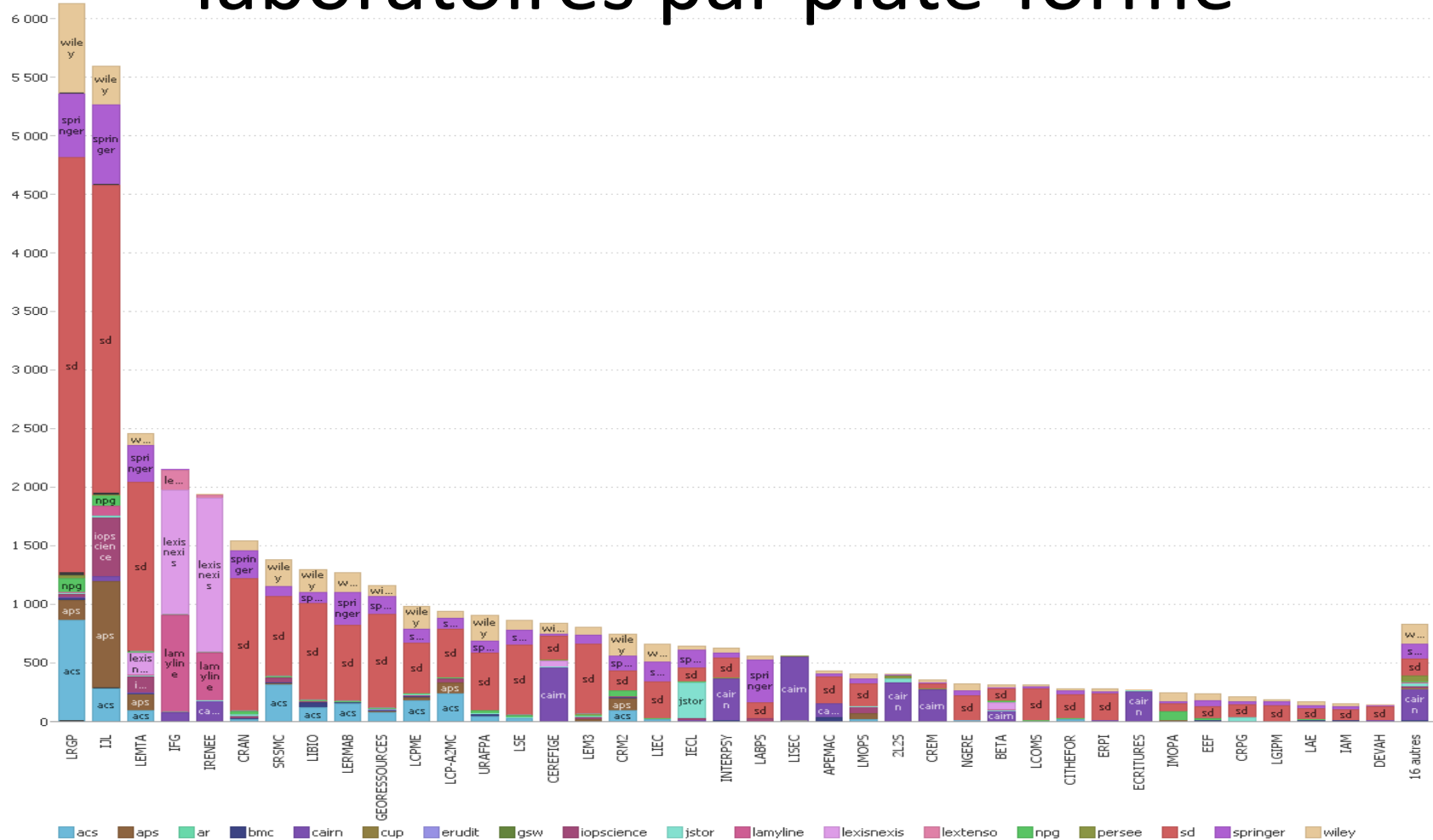
Logs TitaneSciences 2010



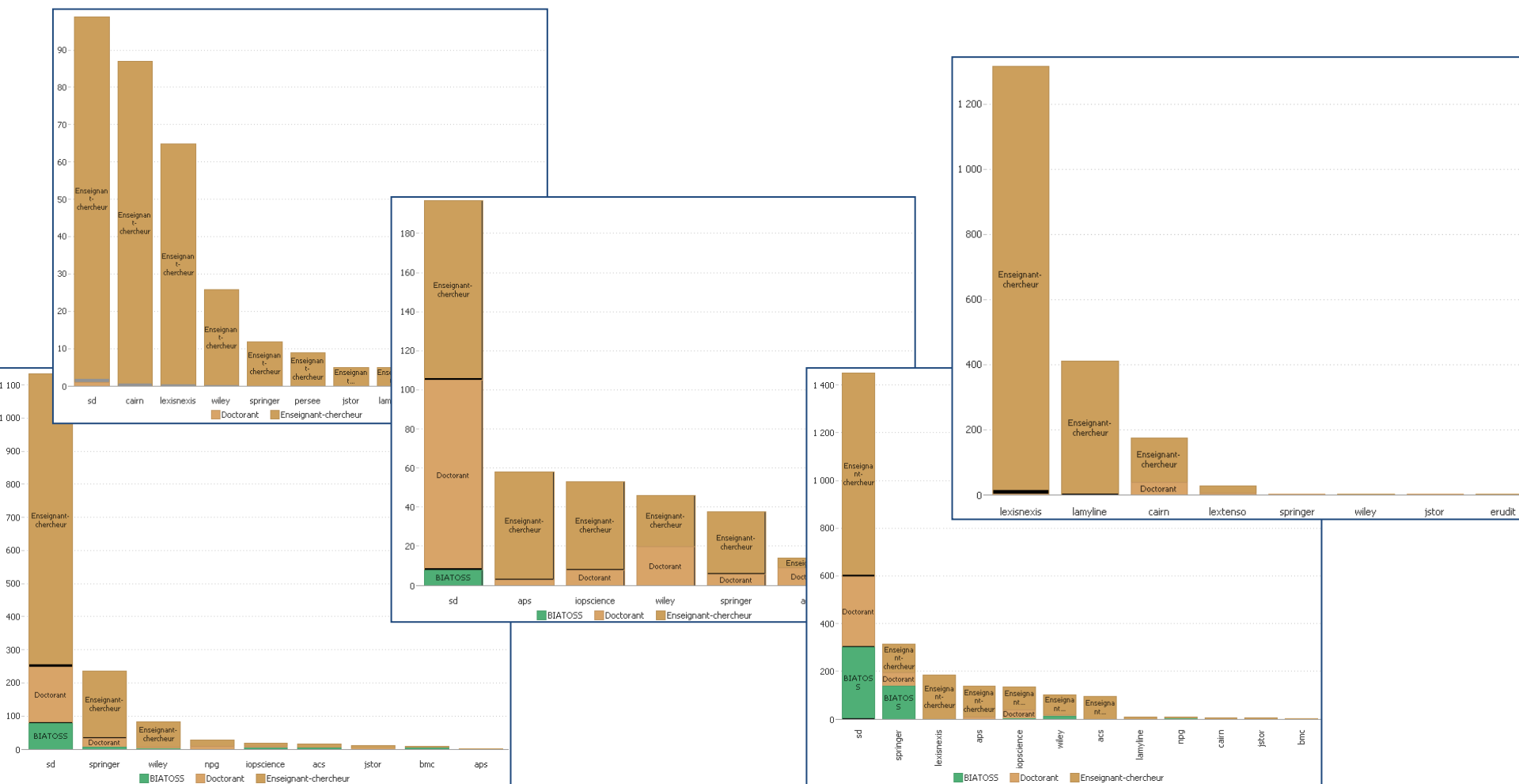
UL : répartition des consultations par plate-forme toutes composantes confondues



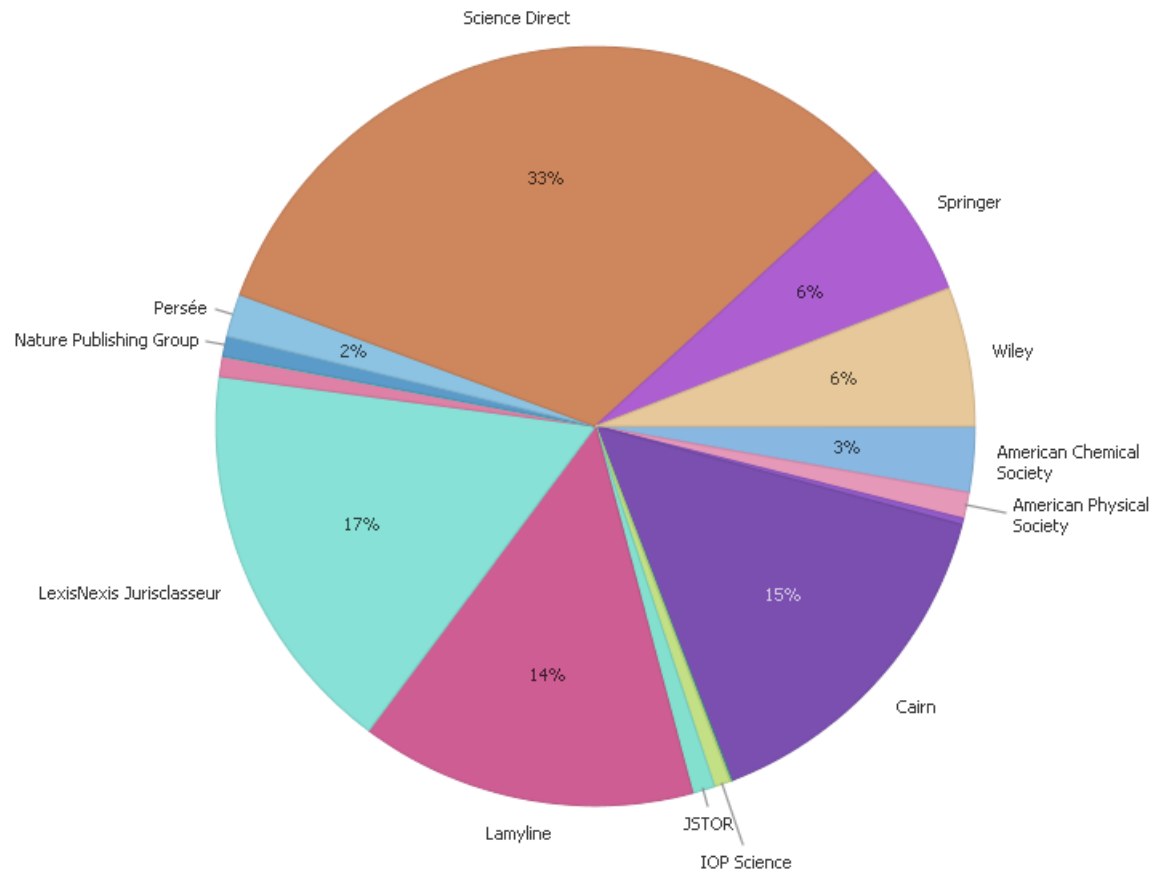
UL : consultation d'articles des laboratoires par plate-forme



... et le détail en « focus » par laboratoire



UL : consultation d'articles par les étudiants, par plate-forme



UL : par composante d'enseignement et par plate-forme

