

ATELIERS DATAVIZZ

Liste des projets

| Etablissement porteur du projet | Descriptif rapide | Données disponibles (typologie / format) | Résultat attendu | Objectif final | Historique |
|---------------------------------|--|---|--|---|---|
| DBU Sorbonne Nouvelle - Paris 3 | Utiliser la dataviz pour mieux exploiter les données d'usages de notre portail documentaire : http://www.dbu.univ-paris3.fr | Données issues de Google Analytics Données issues du serveur web (logs) | Produit web avec mise à jour régulière (quotidienne/hebdomadaire) | Indicateur de pilotage et aide à la décision lors des refontes (nouvelles fonctionnalités, toilettage graphique) Données d'utilisation essentielles pour le prochain projet web (2019) | Aucune tentative précédente Peu d'utilisation des données pour le moment |
| ENS Paris-Saclay | Exploitation des données lecteurs / prêts issues de notre SIGB (Koha) | BD SQL (Koha) | Alimentation automatique (périodicité à définir) au fur et à mesure des inscriptions et interface web de visualisation avec possibilité d'export tableur | Reporting et présentation aux instances avec objectif final d'automatiser tout le rapport d'activité | Rapports SQL dans Koha mais pas automatisés et qui n'alimentent rien |
| Université Le Havre-Normandie | Exploiter les données sur les collections afin d'avoir un outil visuel et partagé de pilotage de la politique documentaire de l'établissement. | Données issues du SIGB Horizon Base SQL adossée (BiblioGQL) Fichiers excel récapitulatifs sur les collections de livres papier (budget, acquisitions, désherbage etc) et leurs usages (prêts, âge, rotation, disponibilité, etc.) | Application web interne partagée entre les acquéreurs Mise à jour mensuelle | Objectif de pilotage pour la gestion des acquisitions et des collections par les acquéreurs et par le responsable de la politique documentaire | Aucune tentative précédente |

| | | | | | |
|--|--|--|--|---|--|
| Université Jean Moulin Lyon 3 | Exploiter les données prêt / niveau d'étude du SIGB Koha pour valoriser les documents qui sortent (incitation pour les autres étudiants de même niveau et discipline) et pour favoriser les interactions avec les acquéreurs | Fichiers CSV/excel (données extraites du SIGB Koha) | Produit visant à enrichir : - le site web (pages collections disciplinaires) - le catalogue Koha (et/ou l'outil de découverte) - l'intranet des étudiants (informations poussées en fonction des disciplines) | Objectif principal : favoriser la réussite des étudiants de licence 1) Renforcer le travail des acquéreurs en favorisant les interactions avec les étudiants 2) Valoriser les fonds pour les étudiants de licence en améliorant leur visibilité et leur adéquation avec les besoins des usagers | Aucune tentative précédente |
| Université Jean Monnet - Saint-Etienne | Mise en valeur des données d'activité du SCD (dans un cadre autre ou pour compléter l'annuel rapport d'activité) | Toutes les données relatives au SIGB sont dans une base dédiée que l'on requête via MySQL (phpMyAdmin) Données autres : sous fichiers Excel | Application web externe Mise à jour annuelle Visualisation de l'activité du SCD sur plusieurs années | Utiliser la datavisualisation à des fins de communication institutionnelle, pour "glamouriser", rendre plus attractives et compréhensives les données d'activité du SCD, auprès des politiques et des lecteurs | Aucune tentative précédente pour ce qui est de la datavisualisation, mais une nouvelle approche de la représentation des données via la rédaction d'un "Digest" original du rapport d'activité |
| Bibliothèque SciencesPo Paris | Mise en valeur des données d'interrogation de l'outil de découverte (environ 600 000 mots annuels) | Extrait des requêtes sur l'outil de découverte | Présentation chronologique et /ou thématique des sujets et des ouvrages et des auteurs recherchés | Mieux comprendre les intérêts des étudiants ainsi que la façon dont ils interrogent le catalogue | Aucune exploitation de ces données dans le passé |
| Université de Bourgogne | Mettre en valeur les données d'activité du SCD et les principaux indicateurs en matière de politique documentaire | Données extraites du SIGB AbsysNet interrogé via BO Données rassemblées sous forme de fichiers Excel | Application web avec mises à jour annuelles | Mieux communiquer sur l'activité du SCD en direction des publics. Valoriser l'action du service auprès de la gouvernance de l'université | Pas d'expérience antérieure |

| | | | | | |
|--|--|---|---|---|--|
| <p>SCD Aix-Marseille Université</p> | <p>Exploitation des données lecteurs : inscriptions, prêts A titre indicatif : - Inscrits 2016 : 40 580 - Prêts 2016 : 396 492</p> | <p>Données issues SIGB Koha</p> | <p>Affichage web pour la visualisation des données avec : > mise à jour automatique (périodicité à définir) > conservation de l'historique des affichages antérieurs</p> | <p>Pilotage et communication externe : suivre et rendre visible l'activité de prêts/inscriptions dans son évolution</p> | <p>Présentation des rapports d'activité SCD 2014 et 2015 en utilisant Piktochart pour la datavisualisation (optique communication). Mais aucune tentative pour l'automatisation des données.</p> |
| <p>SCD de l'Université Paris Descartes (Paris V)</p> | <p>Avoir un outil moderne qui exploite pleinement les indexations des notices bibliographiques</p> | <p>Notre corpus de notices bibliographiques</p> | <p>Une navigation (web) graphique dans les thésaurus, rameau par exemple (mais ça pourrait être fmesh) en fonction des données de notre catalogue : nœuds liés par de nombreuses occurrences communes, zoom avant/arrière pour atteindre les nœuds parents/enfants, affichage des notices correspondant à la descendance du nœud sélectionné (et non pas que le nœud comme sur gallica), auteurs phares pour cette descendance etc...</p> | <p>Ce serait pour faire une interface publique pour naviguer plus intuitivement dans le catalogue</p> | <p>Pas d'historique</p> |

SCD de l'Université
Rennes 2

Avec la mise en œuvre des projets EZparse/EZmesure et AGIMUS des progrès importants ont été consentis pour tendre vers une meilleure analyse et compréhension des usages des ressources et des services numériques.

Cependant, nous continuons (et c'est sûrement le cas d'autres établissements) à rencontrer des difficultés pour professionnaliser la collecte d'indicateurs plus traditionnels de l'activité de nos bibliothèques.

Notre volonté serait donc de capitaliser sur l'expérience des projets sus-cités (et éventuellement de se rapprocher d'autres comme prevu.fr) en mettant cette fois-ci, le focus sur la question de l'usage des ressources imprimées en bibliothèque universitaire.

Partir de l'exploitation des données d'usage issues des SIGB locaux, en s'appuyant en particulier sur les données de circulation (enrichies de données bibliographiques ou encore de données caractérisant les utilisateurs) pour construire une solution permettant une remontée automatique des indicateurs de l'utilisation de la documentation imprimée.

Données provenant du SIGB open source KOHA. Plus précisément, il s'agira d'extractions SQL issues de plusieurs tables (statistics, borrowers, biblio, biblioitems, items).

Il est possible de fournir les données dans plusieurs formats si nécessaire : sql, csv, json

Ce corpus de données permettra d'obtenir :

- des données caractérisant les documents/exemplaires empruntés : identifiant notice, titre, auteur, éditeur, date de publication, collection, identifiants (ISBN, ISSN), PPN, type de document, identifiant exemplaire, code-barres, cote, type de prêt, site, localisation, statut, date de création, niveau cible
- des données caractérisant les emprunteurs : catégorie de lecteur, année inscription, statut, établissement, niveau, code étape, diplôme, discipline, affectation, campus
- des données caractérisant les transactions réalisées (données d'exemplaire) : site, date et heure, type de transaction

Volumétrie : 2 400 000 lignes
Rétrospectif : 2013 à aujourd'hui

La solution pourrait être constituée de plusieurs parties :

- un module de traitement des données, servant à la collecte, à la transformation et à l'enrichissement des données de circulation. Dans l'idéal, la collecte pourrait se faire quotidiennement/hebdomadairement de manière (semi)-automatique ;
- un moteur d'indexation (elasticsearch, influxDB, autre ?) servant au stockage des données enrichies ;
- une application de rendu visuel (kibana, etc.) qui permettra un affichage dynamique des indicateurs par requête temps réel ou encore la création de tableaux de bord ;
- des services web qui permettront de réutiliser, publier facilement les visualisations/extractions générées dans différents contextes, sur différentes plateformes, dans différents formats.

En interne (pour le SCD) :

- aider à l'évaluation et au pilotage des acquisitions
- connaître et comprendre l'usage qui est fait de la documentation imprimée
- tendre vers une meilleure adéquation de l'offre et des besoins
- prévisualiser l'activité de la bibliothèque et rendre compte des rythmes universitaires
- connaître les pratiques informationnelles, comprendre les stratégies documentaires de recherche des publics, les typologies et les communautés de lecteurs
- disposer d'indicateurs d'usage pour la pédagogie : impact des formations documentaires, etc.
- identifier les mutations et les tendances en œuvre dans les usages

Pour la communauté universitaire :

- faire connaître l'activité de la bibliothèque par la valorisation de ses collections imprimées = meilleure (re)connaissance du SCD par le public (usager ou non) ainsi que par la tutelle
- imaginer des actions de valorisation/sensibilisation physiques (matérialisées dans l'espace) en utilisant la data-visualisation : visualisation des données sous la forme d'expositions par exemple (en lien avec certains enseignements réalisés à l'université (partenariats à construire avec le Master Arts - Spécialité Création et Management Multimédia ou encore le Master ergonomie)
- constituer d'un socle favorable à la conception/construction de nouveaux services à valeur ajoutée :
 - * valoriser, recommander les collections par l'usage qui en est fait
 - * dispositif de médiation numérique et de recherche dans le catalogue et/ou le site web de la bibliothèque
 - * favoriser de nouvelles pratiques de recherche documentaire et de nouveaux usages scientifiques par le biais de visualisations des données

Si le projet est porté au niveau de plusieurs établissements :

L'ambition d'un tel projet pourrait être de ne pas forcément se limiter à des expériences locales ou parfois isolées. En effet, agréger l'ensemble de ces données dans un entrepôt unifié facilement exploitable (comme le propose Ezmesure pour la documentation électronique) permettrait plusieurs choses :

- disposer d'un outil libre générique qui collecte à la fois localement et nationalement les données d'usage et les analyse sans dépendre de la technique ou des

Fonctionnement actuel basé sur le module "Rapports" de Koha. Plusieurs requêtes SQL pré-paramétrées. Quelques tableaux de bord réalisés sous des logiciels de type tableau (excel, libreoffice calc).

| | | | | | |
|---|--|---|--|--|--|
| | | | | <p>modules/outils fournis par les prestataires desdits SIGB (= sortir les SCD de l'artisanat dans la collecte des indicateurs)</p> <ul style="list-style-type: none"> - favoriser la mutualisation et le partage d'expérience entre établissements : réflexion sur la définition des données d'usage attendues, des indicateurs, des scénarios de visualisation, sur les questions d'informatique et libertés - proposer une interface en ligne de visualisation dynamique, consolidée et comparative des données entre les établissements ce qui pourrait offrir une alternative potentielle à ASIBU (par exemple). | |
| DDE de l'Université de Lorraine | J'aimerais établir une cartographie des documents empruntés par tranches de cotes ou grands domaines et par filières ou types de lecteurs | <p>2 fichiers Excel :</p> <p>1 avec la liste des documents empruntés et leurs cotes avec les 5 dernières dates d'emprunt, les 5 dernières dates de retour et les 5 codes barres des lecteurs ayant empruntés</p> <p>1 avec la liste des codes barres lecteurs et leurs filières ou leurs catégories de lecteurs</p> | <p>2 tableaux annuels sur web interne avec 1 tableau des filières de lecteurs et les tranches de cotes empruntés avec leurs périodes</p> <p>et 1 ou plusieurs tableaux par bibliothèques avec les tranches de cotes et les filières emprunteuses avec leurs périodes</p> | <p>Cartographie de ce qui est emprunté pour les acquéreurs et les responsables de pôles et plus largement pour le reste du personnel</p> | <p>Pour l'instant à l'état de réflexion (Macro sur Excel ?)</p> |
| Université Paris-Est Marne-la-Vallée (UPEM) | Représenter les co-publications présentes dans HAL entre les structures de la recherche du (nouveau) I-SITE Future — rassemblant l'EIVP, l'ENSAVT, l'ENSG, l'ESIEE Paris, l'Ifsttar et enfin l'UPEM. | <p>Données bibliographiques/affiliations des publications scientifiques (articles, actes, ouvrages...) référencées dans HAL, sur les 5 dernières années. Format au choix JSON, XML, CSV</p> | <p>Un rendu interactif web externe serait souhaitable, et idéalement un affichage en temps réel (via l'API de HAL).</p> | <p>Rendre compte des connexions déjà présentes entre les nombreuses structures de la recherche de cette nouvelle entité qu'est l'I-SITE Future.</p> <p>Placer HAL, et par la même l'Open Access, comme outil central pour référencer et diffuser la production scientifique écrite.</p> | <p>Deux rapports bibliométriques réalisés (avec HAL, Scopus et le Wos) à l'échelle de l'UPEM</p> |