

La consolidation des données statistiques de Koha dans le réseau BRISE-ES

ENSSIB - 14 janvier 2016

Soledad Beudon

soledad.lida@enssib.fr



BRISE-ES

Préliminaires...

- ▶ BRISE-ES = Bibliothèques en Réseau Informatisé de Saint-Etienne - Enseignement Supérieur = 26 bibliothèques / 6 établissements
- ▶ Réinformatisation avec Koha en 2011
- ▶ Base de données MySQL
- ▶ Moteur d'indexation SOLR

À l'origine du projet

Le constat de plusieurs problèmes liés aux statistiques

Case départ : la table des transactions

datetime	branch	type	itemnumber	itemtype	borrowernumber
2015-06-01 17:44:54	BUDL	issue	213214	OUVRAGE	43796
2015-06-01 17:44:44	BUDL	issue	556699	OUVRAGE	43796
2015-06-01 17:44:15	BUDL	return	479423	EXCLU	43796
2015-06-01 17:40:17	BUDL	issue	297350	OUVRAGE	38278
2015-06-01 17:39:44	ENSASE	issue	72844	OUVRAGE	51425
2015-06-01 17:39:39	ENSASE	issue	4235	OUVRAGE	51425
2015-06-01 17:39:34	ENSASE	issue	197541	OUVRAGE	51425
2015-06-01 17:39:16	BUDL	return	479434	EXCLU	55440
2015-06-01 17:38:37	ENSASE	return	197541	OUVRAGE	51425
2015-06-01 17:38:32	ENSASE	return	72844	OUVRAGE	51425
2015-06-01 17:38:25	ENSASE	return	4235	OUVRAGE	51419
2015-06-01 17:34:26	BUDL	return	479391	EXCLU	34842
2015-06-01 17:34:00	BUDL	issue	100604	OUVRAGE	54702
2015-06-01 17:32:45	BUDL	return	479424	EXCLU	44000
2015-06-01 17:32:15	BUROANNE	renew	379874	OUVRAGE	61038
2015-06-01 17:32:14	BUROANNE	renew	512747	OUVRAGE	61038

Non-conformité CNIL

- ▶ La table des transactions stocke l'identifiant lecteur ET l'identifiant exemplaire pour chaque transaction :

datetime	branch	type	itemnumber	itemtype	borrowernumber
2015-06-01 17:44:54	BUDL	issue	213214	OUVRAGE	43796
2015-06-01 17:44:44	BUDL	issue	556699	OUVRAGE	43796
2015-06-01 17:44:15	BUDL	return	479423	EXCLU	43796

- ▶ Or la réglementation CNIL limite à 4 mois la conservation de telles données une fois la transaction terminée (<http://www.cnil.fr/vos-obligations/declarer-a-la-cnil/declaration/mon-secteur-dactivite/mon-theme/je-dois-declarer/declaration-selectionnee/dec-mode/DISPLAYSINGLEFICHEDECL/dec-uid/17/>)

Écart entre l'enregistrement et l'extraction (1/2)

- ▶ La table des transactions est relativement pauvre en matière de champs enregistrés

datetime	branch	type	itemnumber	itemtype	borrowernumber
date et heure	bib de transaction	type de transaction	id exemplaire	catégorie de prêt	id emprunteur

- ▶ Pour croiser ces infos avec d'autres tables, il faut une jointure

Table des transactions

datetime	branch	type	itemnumber
2015-06-01 17:44:54	BUDL	issue	213214

Donnée du moment de l'enregistrement



Table des exemplaires

itemnumber	itemcallnumber
213214	306 CER

Donnée du moment de l'extraction

Écart entre l'enregistrement et l'extraction (2/2)

- ▶ Plus le moment de l'extraction est éloigné du moment de l'enregistrement de la transaction, plus le risque d'obtenir des résultats faux augmente
- ▶ Au niveau des exemplaires
 - ▶ Changements de cote
 - ▶ Changements libre accès > magasin, déménagement...
 - ▶ Désherbage, suppression d'exemplaires
- ▶ Au niveau des dossiers lecteurs
 - ▶ Changement d'année d'études
 - ▶ Changement de catégorie d'emprunteur
 - ▶ Changement de cursus
 - ▶ Suppression du dossier

Au niveau local, deux besoins supplémentaires

- ▶ Avoir un rétrospectif sur plusieurs années
 - ▶ Mise en place d'une politique documentaire formalisée
 - ▶ Volonté de récupérer la table statistique de notre SIGB précédent (Loris)
- ▶ Pouvoir interroger une base à jour sans peser sur la base de production

Une solution stéphanoise :

Mise en place de tables statistiques spécifiques sur
une base Koha à j-1

Mise en place de l'import quotidien de la base de production

- ▶ Sauvegarde des permissions, des rapports et des préférences système de la base statistique à $j-1$
- ▶ Import de la base de données et des index SOLR de la base de production dans la base Koha statistique à $j-1$
- ▶ Réintroduction dans la base Koha statistique à $j-1$ des permissions, des rapports et des préférences système sauvegardés avant import

Constitution de tables spécifiques (1/2)

Les 5 tables statistiques brisées_ :

- ▶ Emprunts
- ▶ Réservations
- ▶ Acquisitions
- ▶ Fascicules
- ▶ (Lecteurs actifs)

Constitution de tables spécifiques (2/2)

- ▶ Recueil des besoins du réseau
- ▶ Sélection des champs et tableau de correspondance Loris / Koha / tables brisees_
- ▶ Création de la table dans la base de données Koha statistique
- ▶ Insertion des données de la table statistique de notre ancien SIGB Loris (2003-2011)
- ▶ Insertion des données du rétrospectif Koha
- ▶ Mise en place de l'alimentation quotidienne dans la foulée de l'import de la base de production

Exemple : l'alimentation de la table brisees_emprunts

▶ Insertion des données transactions

▶ Mise à jour des données :

▶ exemplaires

▶ lecteurs

▶ attributs lecteurs

▶ biblio 1

▶ biblio 2

▶ Anonymisation des champs temporaires contenant des identifiants



Table emprunts



Requetes emprunts

Les évolutions ultérieures

- ▶ Cryptage de l'identifiant lecteur dans `brisees_emprunts` et `brisees_réservations`
- ▶ Suppression de la table `brisees_lecteurs_actifs`
- ▶ Mise à jour rétrospective des domaines des exemplaires dans `brisees_emprunts` et `brisees_réservations`
- ▶ À chaque début d'année, suppression des données de l'année la plus ancienne (10 ans de statistiques courantes)
- ▶ Changement de machine, augmentation de l'espace disque et montée de version du système d'exploitation Debian

Un bilan très positif

- ▶ Gain de fiabilité
- ▶ Résultats statistiques enrichis sur les emprunts, les réservations, les acquisitions et les fascicules



Transactions 2015



Emprunts
domaines

- ▶ Plus de problèmes liés à des suppression d'exemplaires, de dossiers lecteurs ou d'abonnements
- ▶ ESGBU
- ▶ Mise en place de l'anonymisation des transactions
- ▶ Tables enrichies = moins de jointures, requêtes allégées
- ▶ Des noms de champs explicites et francophones

Exploitation du module « Bilans et statistiques »

Modifications et adaptations locales

L'idée initiale...

- ▶ Ne pas dépendre des « Assistants statistiques » de Koha
- ▶ Des rapports à variables, limités en nombre mais souples en usages
- ▶ Une interface d'interrogation aisée pour un groupe d'utilisateurs non experts en MySQL

Amélioration du tableau des rapports sauvegardés

► Tri des colonnes

ID	Nom du rapport	Type	Notes
----	----------------	------	-------

► Choix du nombre de rapports affichés

Rapports sauvegardés

Choisir le rapport à exécuter à partir de la liste

Montrer entrées par page

► Filtre sur le contenu de tous les champs du tableau

filtrer (sur toutes les colonnes et toutes les lignes) :

Les rapports sauvegardés (1/2)

- ▶ Rapports préfixés en fonction du module : ACQ -, CAT-, CIRC-, PERIO -

ID	Nom du rapport ▲	Type ▼	Notes ▼
41	ACQ - Décompte des dépenses à un instant t avec comparaison calendrier	1	Décompte des dépenses à un instant t par poste budgétaire avec comparaison avec l'estimé du calendrier et écart estime-dépense Variable : poste budgétaire Tables Koha acq
2	ACQ - Décompte des exe reçus par poste budgétaire avec détails	1	Décompte des exe reçus et sommes des montants dépensés, répartis par poste budgétaire, fournisseur, domaine, provenance, code de collection Variable : poste budgétaire Table brisees_reception_acq

- ▶ Deux types de variables
 - ▶ Menu déroulant avec liste de valeurs autorisées
 - ▶ Champ libre

Les rapports sauvegardés (2/2)

- ▶ La requête SQL réellement écrite contient...

items.homebranch=<<Bibliothèque de l'exemplaire | branches>>

items.itemcallnumber like << Cote commence par (mettre % derrière début de cote ou % seul pour toutes les cotes) >>

- ▶ ...mais l'utilisateur voit...

Bibliothèque de l'exemplaire :

Cote commence par (mettre % derrière début de cote ou % seul pour toutes les cotes) :

- ▶ ...et a la description du rapport sous les yeux au moment de le lancer

Liste d'exe par bibliothèque, localisation, cote, date de création de l'exemplaire et code de collection biblio

Variables: bibliothèque, localisation, cote, dates de création de l'exemplaire et code de collection biblio

Champs exportés : Code de collection notice, Date de création Notice, Auteur, Titre, Complément de titre, PPN, ISBN, ISSN, Editeur, Date de publication, Titre collection, Volume collection, Bibliothèque, Localisation

Code barres, Cote, Numéro inventaire, Date de création exemplaire, imputation, Notes exemplaire, Numéro de périodique, Nombre de prêts, Date de retour prévue

Tables Koha biblio et exe

L'écran de lancement

Saisir les paramètres pour le rapport **CAT - Liste** d'exemplaires par bib, loc, cote, date de création de l'exemplaire et code de collection biblio:

Liste d'exemplaires par bibliothèque, localisation, cote, date de création de l'exemplaire et code de collection biblio

Variables: bibliothèque, localisation, cote, dates de création de l'exemplaire et code de collection biblio

Champs exportés : Code de collection notice, Date de création Notice, Auteur, Titre, Complément de titre, PPN, ISBN, ISSN, Editeur, Date de publication, Titre collection, Volume collection, Bibliothèque, Localisation

Code barres, Cote, Numéro inventaire, Date de création exemplaire, imputation, Notes exemplaire, Numéro de périodique, Nombre de prêts, Date de retour prévue

Tables Koha biblio et exe

Code de collection de la notice :

ET Bibliothèque de l'exemplaire :

ET localisation de l'exemplaire contient (mettre la loc sous forme de code ou % seul pour toutes les loc) :

ET Cote commence par (mettre % derrière début de cote ou % seul pour toutes les cotes) :

ET Date de création de l'exemplaire inférieure ou égale à (mettre sous la forme AAAA-MM-JJ) :

ET Date de création de l'exemplaire supérieure ou égale à (mettre sous la forme AAAA-MM-JJ) :

L'écran de résultats

CAT - Liste d'exé par bib, loc et note de l'exé

Liste d'exé par bibliothèque, localisation, note et date de création de l'exé

Variables : bibliothèque, localisation, note et dates de création de l'exé

Champs exportés : Code de collection notice, Date de création Notice, Auteur, Titre, Complément de titre, PPN, ISBN, ISSN, Editeur, Date de publication, Titre collection, Volume collection, Bibliothèque, Localisation, Catégorie de prêt, support physique, Domaine, Etat, Code barres, Cote, Numéro inventaire, Date de création exemplaire, imputation, Notes exemplaire, Numero de périodique, Nombre de prêts, Date de retour prévue

Tables Koha biblio et exe

Nombre total d'enregistrement correspondant à la requête : 21.

```
SELECT ifnull(biblioitems.itemtype,'NULL') as 'code de collection notice', ifnull(biblio.datecreated,'NULL') as 'Date de creation Notice', ifnull(biblio.author,'NULL') as 'Auteur', ifnull(biblio.title,'NULL') as 'Titre', ifnull(biblioitems.volume,'NULL') as 'Complement de titre', ifnull(biblioitems.number,'NULL') as 'PPN', ifnull(biblioitems.isbn,'NULL') as 'ISBN', ifnull(biblioitems.issn,'NULL') as 'ISSN', ifnull(biblioitems.publishercode,'NULL') as 'Editeur', ifnull(biblioitems.publicationyear,'NULL') as 'Date de publication', ifnull(biblioitems.collectiontitle,'NULL') as 'Titre collection', ifnull(biblioitems.collectionvolume,'NULL') as 'Volume collection', ifnull(items.homebranch,'NULL') as 'Bibliothèque exe', ifnull(items.location,'NULL') as 'Localisation exe', ifnull(items.type,'NULL') as 'Catégorie de prêt exe', ifnull(items.statisticvalue,'NULL') as 'support physique exe', ifnull(items.copynumber,'NULL') as 'domaine_exe', ifnull(items.notforloan,'NULL') as 'Etat exe', ifnull(items.barcode,'NULL') as 'Code barres exe', ifnull(items.itemcallnumber,'NULL') as 'Cote exe', ifnull(items.stocknumber,'NULL') as 'Numero inventaire exe', ifnull(items.dateaccessioned,'NULL') as 'Date de creation exe', ifnull(items.materials,'NULL') as 'imputation exe', ifnull(items.itemnotes,'NULL') as 'Notes exe', ifnull(items.enumchron,'NULL') as 'Numero de periodique exe', ifnull(items.issues,'NULL') as 'Nombre de prêts exe', ifnull(items.onloan,'NULL') as 'Date de retour prévue' FROM items LEFT JOIN biblioitems on (items.biblioitemnumber=biblioitems.biblioitemnumber) LEFT JOIN biblio on (biblioitems.biblioitemnumber=biblio.biblioitemnumber) WHERE items.homebranch='GATELSE' and items.location like '%&' and items.itemnotes like '%chaire%' and items.dateaccessioned <='2013-12-31' and items.dateaccessioned >='0000-00-00' << < 1 2 >>
```

code de collection notice	Date de creation Notice	Auteur	Titre	Complement de titre	PPN	ISBN	ISSN	Editeur	Date de publication	Titre collection	Volume collection	Bibliothèque exe	Localisation exe	Catégorie de prêt exe	support physique exe	domaine_exe	Etat exe	Code barres exe	Cote exe	Numero inventaire exe	D c
LIVRE	2011-11-30	Steven G. Medema	The hesitant hand	taming self-interest in the history of economic ideas	136887589	9780691122960 9780691		Princeton University Press	2009			GATELSE	BIB	OUVRAGE	LIVRE	GATE-ECONOMIE	0	34200012505703	B0 MED	GAT 7964	20
LIVRE	2011-12-01	John B. Davis	Individuals and identity in economics		153774096	9781107001923 11070019		Cambridge University Press	2011			GATELSE	BIB	OUVRAGE	LIVRE	GATE-ECONOMIE	0	34200012505745	A12 DAV	GAT 7965	20
LIVRE	2011-12-01	Daniel M. Hausman, Michael S. McPherson	Economic analysis, moral philosophy and public policy		106995553	0521846293 052160866X		Cambridge university press	2006			GATELSE	BIB	OUVRAGE	LIVRE	GATE-ECONOMIE	0	34200012505737	A12 HAU	GAT 7966	20
LIVRE	2011-12-01	William H. Riker	Liberalism against populism	a confrontation between the theory of democracy and the theory of social choice	156763036	0881333670 97808813336		Waveland Press	1982			GATELSE	BIB	OUVRAGE	LIVRE	GATE-ECONOMIE	0	34200012505729	D71 RIK	GAT 7967	20

Faciliter la lecture des résultats

- ▶ Distinguer les champs non renseignés des champs NULL

GATELSE	BIB	OUVRAGE	LIVRE	GATE-ECONOMIE	0	34200012506685	A12 REI	GAT 8079	2013-11-21	NULL	
---------	-----	---------	-------	---------------	---	----------------	------------	----------	------------	------	--

- ▶ Afficher des totaux, sous-totaux

ACQ - Décompte des exe reçus par poste budgétaire global

Décompte des exe reçus et sommes des montants dépensés, répartis par budget père, poste budgétaire

Variable : poste budgétaire

Table brisees_reception_acq

Nombre total d'enregistrement correspondant à la requête : 6.

```
SELECT ifnull(budget_pere, 'Total') as budget_pere, case WHEN poste_budgetaire is not null THEN poste_budgetaire WHEN budget_pere IS NOT NULL THEN 'Sous-total' ELSE sum( nbr_exe_recus ) AS 'quantite recue' FROM brisees_reception_acq WHERE 'poste_budgetaire' LIKE '%BUDL%2014%' GROUP BY budget_pere, 'poste_budgetaire' with
```

budget_pere	poste_budgetaire	somme prix fois quantité	quantite recue
BUDL CQ 2014	BUDL CQ 2014	27975,87	1316
BUDL CQ 2014	Sous-total	27975,87	1316
BUDL DGF 2014	BUDL DGF PEDA 2014	31258,64	1244
BUDL DGF 2014	BUDL DGF RECH 2014	6692,58	219
BUDL DGF 2014	Sous-total	37951,22	1463
Total	Total	65927,09	2779

Télécharger le rapport

Un groupe d'utilisateurs relais

- ▶ Formation d'un groupe à l'utilisation des rapports
- ▶ À chaque site son référent
- ▶ La liste des rapports à disposition des collègues

Merci de votre attention