



AlgoLiens : un nouveau service de "détection d'absence de liens 5XX, 60X 7XX dans les notices" pour les catalogueurs du Sudoc

Description

Formation pour la prise en main d'AlgoLiens, générateur de rapports d'absence de liens dans les zones de lien du Sudoc

Public

Correspondant Autorité
Correspondant catalogage
Coordinateur Sudoc



Intervenants

François Mistral, responsable IdRef-Autorités
modérateur : Raphaëlle Poveda, service formation et documentation

PLAN

- Le contexte
- L'algorithme
- Générer un rapport dynamique
- Engager un chantier de correction

RAPPEL DU CONTEXTE

Qualité du catalogue

- Approcher la notion de qualité du catalogue par les autorités et les liens.
- Compléter la boîte à outils des catalogueurs.

Les anomalies

- Pathologies de liens :
 - Déficit de liens
 - Déficit de notices d'autorité
 - Mauvais liens

Les traitements

➤ Les thérapeutiques :

- Rendre visible les anomalies
- Déterminer des corpus pertinents
- Engager des chantiers
- La force du collectif et de la coordination
- L'alliance « Humain + Machine »



Les enjeux

1. Exhaustivité du signalement catalographique
2. Valorisation scientifique de l'IST
3. Valorisation patrimoniale des BU
4. Contribution au Web de données liées



L'ALGORITHME D'ALGOLIENS

AlgoLiens ?

- C'est :

un Web service de génération de rapport
d'anomalies sur les zones de liens des notices Sudoc



- Ce n'est pas :

une espèce humanoïde originaire du Quadrant
Alpha ou Beta, rencontrée dans Star Trek

Fonctionnement d'AlgoLiens

- Un programme scanne toutes les notices bibliographiques et d'autorité du Sudoc modifiées dans la journée et enchaîne une série de tests sur les zones de liens :
 - A 5XX
 - B 500, 60X et 7XX
- 4 exceptions :
 - Notices de RCR
 - Indexation avec \$2mesh ou \$2lc
 - Niveau bibliographique 0
 - Zones 4XX

Le test / les tests

- Le TEST = absence de liens dans une zone de liens
- Les TESTS = de nombreux filtres
 - Sur les ILN
 - Sur les RCR
 - Sur les types de documents
 - Sur chacune des zones de liens
 - Sur les blocs A5XX, B60X, B70X, B71X et B72X
 - Sur les unicas
 - Sur la date de création de la notice

**GÉNÉRER UN RAPPORT
DYNAMIQUE**

Principes

- La manipulation : une url dans un navigateur
 - Au préalable, passer des paramètres dans l'url pour personnaliser la requête
- Le résultat : un fichier .csv
 - À enregistrer en local sur son poste
 - À modifier comme on veut (tri..)
 - Pour servir de feuille de route pour des corrections à faire dans WinIBW / IdRef

Exemple d'un rapport dynamique (1/3)

1	Rapport genere le 2016-11-29T13:28:47Z										
2	Url utilisee : http://www.idref.fr/AlgoLiens?iln=12										
3	PPN	ILN	RCR	Texte	Date	Code	TypDoc	LibelleTypDoc			
4	196708451	12	250562103	une 702 sans lien	2016-11-23 00:00:00.0	B702	Ba	Document audiovisuel			
5	059390654	12	250562102	une 550 autorité sans lien	2016-10-25 00:00:00.0	A550	null	null			
6	059358939	12	250562102	une 550 autorité sans lien	2016-10-25 00:00:00.0	A550	null	null			
7	078009499	12	250562101	une 540 autorité sans lien	2015-10-03 00:00:00.0	A540	null	null			
8	098789090	12	900102102	une 500 autorité sans lien	2013-05-30 00:00:00.0	A500	null	null			
9	196740746	12	250562101	une 702 sans lien	2016-11-24 00:00:00.0	B702	Aa	Monographie imprimée			
10	196789966	12	253882101	une 700 sans lien	2016-11-28 00:00:00.0	B700	Aa	Monographie imprimée			
11	059451793	12	250562102	une 550 autorité sans lien	2016-10-25 00:00:00.0	A550	null	null			
12	059645423	12	250562102	une 550 autorité sans lien	2016-10-25 00:00:00.0	A550	null	null			
13	059637080	12	250562102	une 550 autorité sans lien	2016-10-25 00:00:00.0	A550	null	null			

Exemple d'un rapport dynamique (2/3)

Date et heure de génération du rapport dynamique

url de génération du rapport dynamique

```
1 Rapport genere le 2016-11-29T13:28:47Z
2 Url utilisee : http://www.idref.fr/AlgoLiens?iln=12
3 PPN      ILN      RCR      Texte  Date  Code  TypDoc  LibelleTypDoc
```

Intitulés des colonnes



Un rapport a une durée de vie de 24h car le programme d'analyse des données du Sudoc tourne chaque nuit.

Exemple d'un rapport dynamique (3/3)

Rapport genere le 2016-11-30T13:43:03Z							
Url utilisee : http://www.idref.fr/AlgoLiens?&iln=4&rownum=100&sort=rcr							
PPN	ILN	RCR	Texte	Date	Code	TypDoc	LibelleTypDoc
04864997X	4	130012101	une 700 sans	2012-08-07 0	B700	Aa	Monographie imprimée
052262480	4	130012101	une 700 sans	2013-05-09 0	B700	Aa	Monographie imprimée
052262480	4	130012101	une 702 sans	2013-05-09 0	B702	Aa	Monographie imprimée
056263112	4	130012101	une 700 sans	2006-07-11 0	B700	Aa	Monographie imprimée
056284306	4	130012101	une 710 sans	2016-07-26 0	B710	Aa	Monographie imprimée
056527853	4	130012101	une 606 sans	2006-05-04 0	B606	Aa	Monographie imprimée
056533918	4	130012101	une 600 sans	2002-03-18 0	B600	Aa	Monographie imprimée
058573720	4	130012101	une 710 sans	2016-02-09 0	B710	Ad	Collection imprimée
058611819	4	130012101	une 500 sans	2010-06-14 0	B500	Aa	Monographie imprimée
05863584X	4	130012101	une 500 sans	2011-03-16 0	B500	Aa	Monographie imprimée
058641793	4	130012101	une 500 sans	2012-06-11 0	B500	Aa	Monographie imprimée
058550275	4	130012101	une 500 sans	2012-01-13 0	B500	Aa	Monographie imprimée
058550402	4	130012101	une 606 sans	2003-04-04 0	B606	Aa	Monographie imprimée
05856344X	4	130012101	une 712 sans	2014-04-09 0	B712	Aa	Monographie imprimée
058861734	4	130012101	une 600 sans	2008-10-31 0	B600	Aa	Monographie imprimée
058869980	4	130012101	une 702 sans	2008-04-18 0	B702	Aa	Monographie imprimée

Le rapport dynamique peut être retravaillé dans un tableur :

- ajout / suppression de colonnes
- tri
- mise en forme...

Comprendre le fonctionnement du webservice AlgoLiens

- Appeler le WS par url

<http://www.idref.fr/AlgoLiens?>

- À compléter par des paramètres

- rechercher par zones de liens **&code=**

- rechercher à partir d'un jour **&date=**

- rechercher par type de document **&typdoc=**

- rechercher les notices d'un établissement **&iln=** ou **&rcr=**

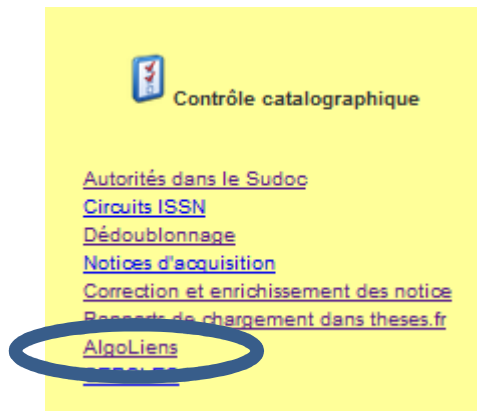
- rechercher les unicas **&unica=rcr** ou **&unica=iln**

- Exemple

<http://www.idref.fr/AlgoLiens?&code=xxx&date=xxxxxxx&typdoc=Xx&iln=x&unica=rcr>

Le manuel

- Se reporter à la documentation :
http://documentation.abes.fr/sudoc/manuels/controle_bibliographique/algoliens/index.html
- Accessible depuis la page Manuel du GM :

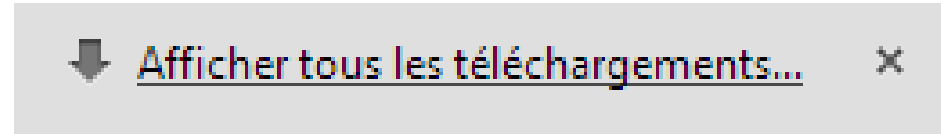
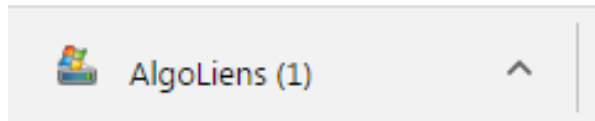


Télécharger le fichier .csv

- Le téléchargement du fichier peut se faire différemment selon la configuration de l'ordinateur et du navigateur que vous utilisez.
- Attention au « 0..... » liminaire des PPN !

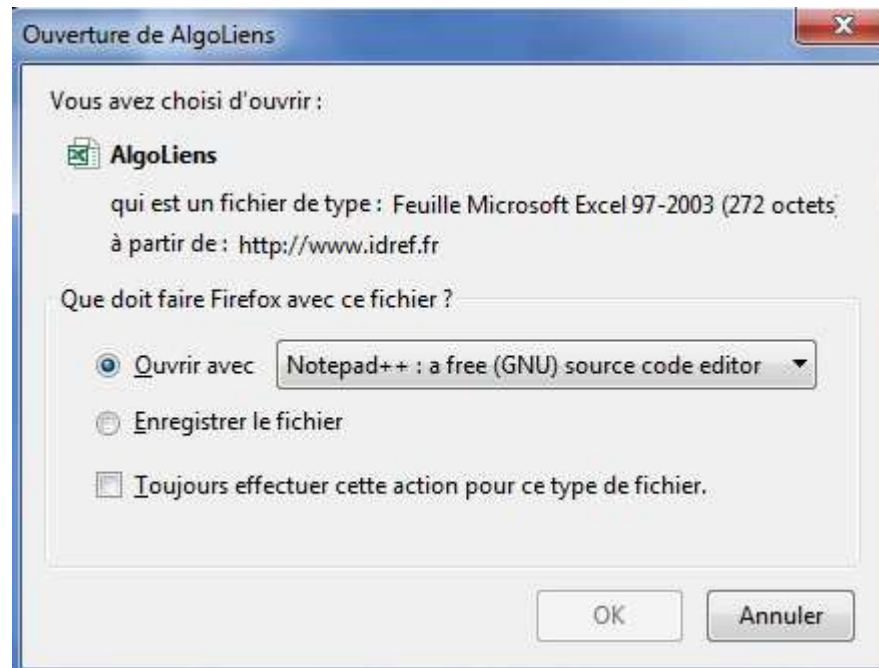
Exemple : Chrome

- En bas du navigateur, à gauche et à droite



Exemple : Firefox

- Un pop-up qui propose le programme grâce auquel on peut ouvrir le fichier .csv



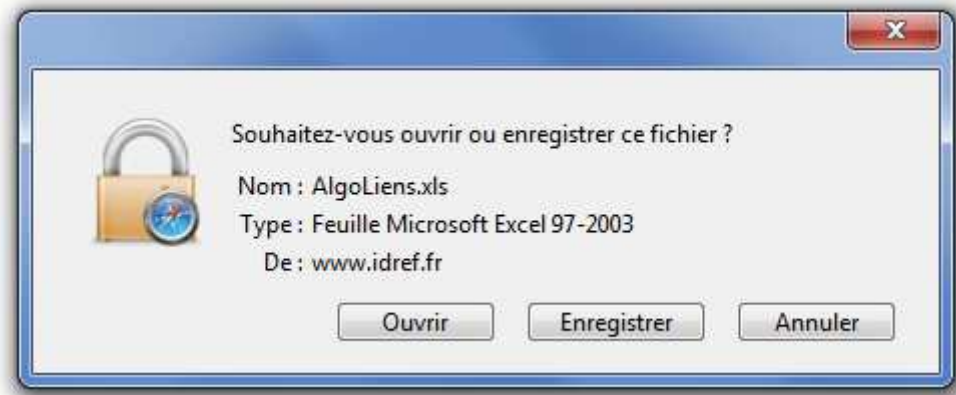
Exemple : Internet Explorer

- Un pop-up qui propose le programme grâce auquel on peut ouvrir le fichier .csv



Exemple : Safari

- Dans une fenêtre téléchargements, un pop-up



DEMO

**ENGAGER UN CHANTIER DE
CORRECTION**

Quelques conseils

- Commencez par traiter les anomalies à votre échelle
 - votre établissement,
 - vos unicas,
 - un type de document,
 - etc.
- Quand vous modifiez une notice, ne vous contentez pas de corriger strictement selon le rapport
 - relecture globale de la notice
 - complétude et exactitude des données
- Vous pouvez informer l'ABES de l'avancée de vos corrections et nous solliciter pour :
 - de l'accompagnement technique
 - de l'accompagnement de service

Ventilation des unicas par ILN

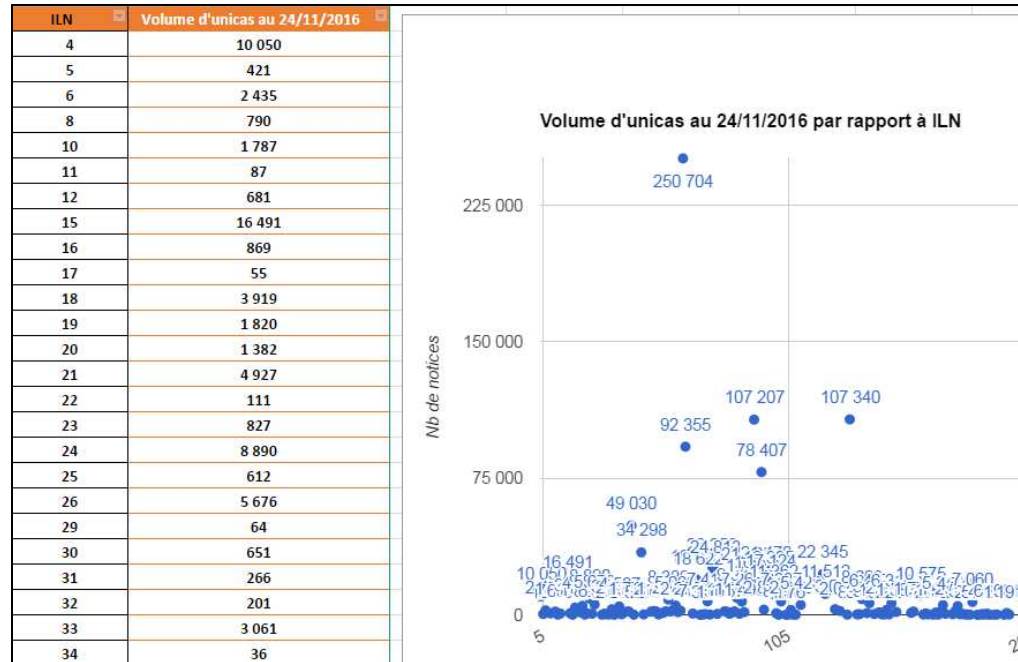


Tableau de bord et visualisation

<https://docs.google.com/spreadsheets/d/1Bf1yWmG97Qd3XYm5FVBInxnDBKLD36pdGgySUspHPsk/edit?usp=sharing>

En résumé

- Toutes les zones de liens des notices (biblio & autorité) du Sudoc ont vocation à être liées.
- Le niveau de qualité du catalogue est une fonction croissante du nombre et de la fiabilité de ces liens.
- L'ABES met à disposition des établissements dans un rapport dynamique les listes de PPN concernés et suggère les interventions à faire.
- Les rapports dynamiques sont générés à partir d'un webservice AlgoLiens, dont l'url peut être paramétrée.
- L'ABES vous laisse libre d'organiser le travail de correction comme vous l'entendez.