



Les outils de contrôle des données Partie 2

Description

Formation pour la prise en main d'outils ou de commandes permettant la mise en place d'une activité de contrôle qualité sur les données du Sudoc



Public

Correspondant Catalogage



Intervenants

Raphaëlle Povéda
Laurent Piquemal

Le contrôle qualité...

- ... est une mission des correspondants Catalogage
- ... est une incitation de l'ABES
- De nombreux outils, commandes et procédures existent...
- ... mais ils n'ont **jamais** été réunis dans un même cours
- L'ABES propose aujourd'hui la 2ème partie

LÉGENDE

 The logo for WinIBW, featuring a stylized white book with a red cover and the text "WinIBW" to its right.	contrôle à effectuer avec WinIBW
 An icon representing a web service, showing the text "http://" inside a rounded rectangular box with a blue mouse cursor arrow pointing to the bottom right corner.	contrôle à effectuer avec un webservice

Les outils de contrôle des données

PARTIE 2

CONTRÔLE DE LA QUALITÉ DES DONNÉES

RAPPEL

avant d'analyser les données...
construire des requêtes avancées

RAPPEL

Les index de recherche utile *(pour un contrôle)*

Usage	Commande
catalogue des notices (par RCR)	CHE RBC <RCR>
notices <u>créées</u> (par RCR) <ul style="list-style-type: none">• par jour• par mois• par année	CHE CRN <RCR>JJMMAA CHE CRN <RCR>??MMAA CHE CRN <RCR>????AA
notices <u>exemplarisées</u> (pour l'ILN) <ul style="list-style-type: none">• aujourd'hui• sur une période	CHE SEL AUJ CHE SEL <JJ-MM-AAAA> JUS <JJ-MM-AAAA>

RAPPEL

- *Penser à combiner avec d'autres index*

Usage	Commande
Lister pour un RCR les notices localisées sur un corpus donné	CHE RBC <RCR> et SOU <code du corpus*>
Lister pour un ILN les notices exemplarisées automatiquement par l'ABES sur un corpus donné	CHE INL ABES et SOU <code du corpus>

**codes de corpus > voir dans le manuel Import*

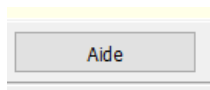


Voir la liste des **Index** dans le « manuel Interrogation » du Guide méthodologique

RAPPEL

- *Penser à utiliser les **limitations***

Usage	Commande
Lister les notices de ressources continues créées par un RCR	t do t; CHE CRN <RCR>JJMMAA
Lister les notices bibliographiques créées par un RCR	t no t; CHE CRN <RCR>JJMMAA
Lister les notices autorités créées par un RCR	t no e; CHE CRN <RCR>JJMMAA



Voir le liste des limitations dans l'**Aide** en ligne de WinIBW

Je veux vérifier dans WinIBW la qualité des données produites

(sur des zones spécifiques)





Vérifier les dates

Préalable : utiliser webstats

Statistiques de votre Etablissement > Catalogue > **nb notices localisées niveau RCR (sur critères) > Dates de publication** > visualiser

BOULOGNE-BU Psych. Henri Pieron

2017

Date de publication	Janv.	Fév.	Mars	Avril	Mai	Juin	Juil.	Août	Sept.
1500 et -	1	1	1	1	1	1	1	1	1
1501-1600	1	1	1	1	1	0	0	0	0
1701-1800	3	3	3	3	3	3	3	3	3
1801-1810	3	3	3	3	3	3	3	3	3
1811-1900	517	518	518	518	527	535	541	541	543

Exemple : RCR 920125201

Puis vérifier dans WinIBW

Nombre de résultats 3		Recherche effectuée : recherche rbc 920125201	
+Aax	Sèze, Victor de (1754-1...	Recherches physiologiques et philosophiques sur ...	chez Prault, imprimeur ... 1786
+Aax	Tardy de Montravel, A. A.	Journal du traitement magnétique de la Demoisell...	[s.n.] 1786
+Aax	Locke, John (1632-1704)	Abregé de l'Essay de Monsieur Locke, sur l'entend...	1741

Commandes :

APU <ANNEE>; che rbc <RCR>

OU

APU <PERIODE>; che rbc <RCR>

Exemple : APU 1700-1799; che rbc 920125201



Vérifier les dates

AFF K:100:219

+Aax	100 0#\$a1975; 219 #1\$d1975;
+Aax	100 0#\$a1975\$f1964; 219 #0\$aParis\$cPayot\$dDL 1975;
+Aax	100 0#\$a1975\$f1965; 219 #0\$a[Mélin, Belgique]\$cTyl; 219 #2\$d1975;
+Aax	100 0#\$a1975\$d1975; 219 #0\$aParis\$cLe Hameau\$dDL 1977;

Exemple : les dates en 100\$a et 219\$d sont incohérentes. À vérifier

AFF K:100:210

Adx	100 0#\$a2017; 210 ##\$aParis\$cPUF\$d2017-;
Adx	100 0#\$a2017\$d2017-...; 210 ##\$aSaint-Denis\$cConnaissances et savoirs\$d[2017]-;
Adx	100 1#\$a2017; 210 ##\$aAmityville, N.Y.\$cBaywood Pub. Co.\$d2003;

Exemple : les dates en 100\$a et 210\$d sont incohérentes. À vérifier



Vérifier les langues

Préalable : utiliser webstats

Statistiques de votre Etablissement > Catalogue > **nb notices localisées niveau RCR (sur critères) > Langue de publication** > visualiser

BOULOGNE-BU Psych. Henri Pieron

2017

Langue de publication	Janv.	Fév.	Mars	Avril	Mai	Juin	Juil.	Août	Sept.
Albanais	1	1	1	1	1	1	1	1	1
Allemand	1 489	1 495	1 496	1 496	1 569	1 641	1 685	1 696	1 777
Amharique	1	1	1	1	1	1	1	1	1
Anglais	18 079	18 082	18 095	18 134	18 220	18 325	18 465	18 499	18 644
Arabe	5	5	5	5	5	5	5	5	5
Arménien	3	3	3	3	3	3	3	3	3

Exemple : RCR 920125201

Puis vérifier dans WinIBW

Commandes :

lan <code de langue> ; che rbc <RCR>

ou

che lai <code langue> et rbc <RCR>

Nombre de résultats 3 Recherche effectuée : recherche rbc 920125201 et lai arm

+Aax	Recueil de travaux scientifiques du Laboratoire de ...	Inst. pédagogique d'état	1965
+Aax	Recueil de travaux scientifiques	Inst. pédagogique d'état	1965
+Aax	Thouthoundjian, Hovsep La Formation et le rôle de l'attention dans l'éducati...	Impr. d'Etat	1959

Exemple : che rbc 920125201 et lai arm



Vérifier les langues de publication

AFF K:101:200

+Aax	101 0#\$fre; 200 1#\$aL'@évolution psychique de l'enfant\$int
+Aax	101 0#\$aund; 200 1#\$a@Pourquoi des relations humaines?\$
+Aax	101 0#\$fre; 200 1#\$aLa @vie affective de l'adolescent inadap

Exemple : en 101, le code langue doit être « fre »

Vérifier les pays de publication

AFF K:102:219

+Aax	102 ##\$aCA; 219 #0\$aQuébec, Québec\$cPresses de l'Université d
+Aax	102 ##\$aFR; 219 #0\$aMontréal\$cChenelière éducation\$dDL 2016;
+Aax	102 ##\$aFR; 219 #0\$aToulouse\$cOctarès éditions\$dDL 2016;

Exemple : en 102, le code pays doit être « CA »



Vérifier les types de contenu/médiation/supports

AFF K:181:182:183

```
181 ##$P01$cxt; 181 ##$P02$csti; 182 ##$P01$cn; 182 ##$P02$cn; 183 ##$P01$P02$anga;  
181 ##$P01$cxt; 182 ##$P01$cn; 183 ##$P01$anga;  
181 ##$P01$cxt; 182 ##$P01$cn;
```

Exemple : il manque une zone 183

AFF K:003:181:182:183

```
003 http://www.sudoc.fr/204049172; 181 ##$P01$cxt; 182 ##$P01$cn;  
003 http://www.sudoc.fr/204043433; 181 ##$P01$cxt; 182 ##$P01$cn; 183 ##$P01$anga;
```



Vérifier la présence des paires de zones « collections »

AFF K:225:410

+Aax	225 0#\$aL'@Harmattan Cameroun;
+Aax	225 2#\$a@Collection L'enfant, la psychiatrie et le psychanalyste\$fCentre Alfred Binet, ASM 13; 410 ##\$0080832938{#@Collection
+Aax	225 1#\$a@Pédagogies; 410 ##\$0002258455{#@Collection Pédagogies, ISSN 1158-4580};

Exemple : sur le 1^{er} exemple, il manque la zone 410 de lien à la collection

AFF K:225:461

+Aax	461 ##\$0040528715{#@Spirale (Ramonville Saint-Agne), ISSN 1278-4699}}\$v5;
+Aax	225 2#\$a@Spirale (Ramonville Saint-Agne) \$v9; 461 ##\$0040528715{#@Spirale (Ramonville Saint-Agne), ISSN 1278-4699}}\$v9;
+Aax	461 ##\$0040528715{#@Spirale (Ramonville Saint-Agne), ISSN 1278-4699}}\$v8;



Vérifier les code de fonction 7XX

AFF K:003:309

- +Aax 003 <http://www.sudoc.fr/024741094>; 309 ##\$aCORRIGER LE(S) CODE(S) DE FONCTION = 000 EN 7XX, L7XX et/ou E7XX.
- +Aax 003 <http://www.sudoc.fr/024714828>;
- +Aax 003 <http://www.sudoc.fr/024667803>;
- +Aax 003 <http://www.sudoc.fr/024356107>; 309 ##\$aCORRIGER LE(S) CODE(S) DE FONCTION = 000 EN 7XX, L7XX et/ou E7XX.
- +Aax 003 <http://www.sudoc.fr/024346063>; 309 ##\$aCORRIGER LE(S) CODE(S) DE FONCTION = 000 EN 7XX, L7XX et/ou E7XX.

Exemple : la présence de zones 309 indique le traitement à faire
(puis supprimer la 309)

AFF K:710:712

- +Abx 003 <http://www.sudoc.fr/039996700>; 712 02\$3031085989{{@Société d'andrologie de langue française}}\$4070;
- +Abx 003 <http://www.sudoc.fr/039932206>; 710 02\$3029732352{{@France. Conseil supérieur des bibliothèques}}\$4070;
- +Abx 003 <http://www.sudoc.fr/037398105>;
- +Abx 003 <http://www.sudoc.fr/048090239>;
- +Abx 003 <http://www.sudoc.fr/040327167>; 710 01\$3029708419{{@France. Direction de la programmation et du développement universitaire}}\$4070;
- +Abx 003 <http://www.sudoc.fr/039948862>; 712 02\$3026414058{{@Syndicat de la médecine générale (France)}}\$4340\$4650;

Exemple : en 712, le code fonction 070 est interdit



Usage

Cela peut service à :

Faire un contrôle régulier

Vérifier le travail en continu

Intervenir rapidement pour corriger

Pensez à vérifier sur les « créations » l'application des nouvelles règles RDA-FR :

181/182/183

210 uniquement pour les notices de ressources continues

219 conforme

etc.

Je veux vérifier **la qualité des notices produites**

(sur un ensemble de zones dans une notice)





CheckSudoc

http://domybiblio.net/check_sudoc/

vérifie la conformité des notices sur :

- Cohérence des dates des zones Unimarc 100 et 210
 - Présence des zones 225 et 410 ou 461
 - Présence des zones 181/182/183 et du sous-champ 200\$b
 - Présence des liens dans les zone 6XX et 7XX
 - Conformité du code de fonction \$4 des zones 7XX
 - Contrôle sur la zone 991 (signalement rétrospectif)
-
- **à partir d'un liste de PPN**



Astuce

Télécharger une liste de PPN

1. CHE CRN <RCR>JJMMAA
2. puis TEL S<n° du lot> 1-20 PPN
3. Ouvrir : C > oclcpica > WinIBW30 > Profiles > NOM UTILISATEUR > download.txt
4. Copier les PPN
5. Coller dans CheckSudoc



CheckSudoc

Présentation en « vignettes »

Retour



[PPN 197531555]

● Genres scolaires

PPN : 197531555

Années : 100 = 2016 210_d=2016

Champs 181/182 :

Champ 200 : 200_a Aucun double espace dans le titre. - 200_b 0

Champ 210 : 210_a : Pas de Mention latine S.I. - 210_c : Pas de Mention latine S.n. - 210_d : Pas de Mention latine S.d.

Champ 225 : **Il manque une zone 225 pour le PPN 197531555 alors qu'il y a une zone 461 (monographie en plusieurs volumes)!**

Champ 305 : 1

Champ 410 : 0

Champ 461 : 1

Champ 606 : Aucune anomalie constatée

Champ 712 : Aucune anomalie relative au \$3 Aucune anomalie relative au \$4 **Le contenu du code fonction en \$4 ne correspond pas à ce qui est attendu en 712**



[PPN 195521986]

● Les manuels dans la classe de français

PPN : 195521986

Années : 100 = 2016 210_d=2016

Champs 181/182 :

Champ 200 : 200_a Aucun double espace dans le titre. - 200_b 0

Champ 210 : 210_a : Pas de Mention latine S.I. - 210_c : Pas de Mention latine S.n. - 210_d : Pas de Mention latine S.d.

Champ 225 : **Il manque une zone 225 pour le PPN 195521986 alors qu'il y a une zone 461 (monographie en plusieurs volumes)!**

Champ 305 : 1

Champ 410 : 0

Champ 461 : 1

Champ 606 : Aucune anomalie constatée

Champ 702 : Aucune anomalie relative au \$3 Aucune anomalie relative au \$4 Aucune anomalie relative au code de fonction



CheckSudoc

Présentation en tableau => rapport téléchargeable (.csv)

Résultats

PPN	Année zone 100	Année zone 210	181	182	200_a	200_b	210_a	210_c	210_d	225	305	410	461	461 \$0	461_t
197531555	2016	2016	2	2	Aucun double espace dans le titre.	0	Pas de Mention latine S.I.	Pas de Mention latine S.n.	Pas de Mention latine S.d.	0	1	0	1	1	1
195521986	2016	2016	2	2	Aucun double espace dans le titre.	0	Pas de Mention latine S.I.	Pas de Mention latine S.n.	Pas de Mention latine S.d.	0	1	0	1	1	1
195245970	2016	2016	2	2	Aucun double espace dans le titre.	0	Pas de Mention latine S.I.	Pas de Mention latine S.n.	Pas de Mention latine S.d.	1	0	1	0	0	0
190382910	2015	2015	2	2	Aucun double espace dans le titre.	0	Pas de Mention latine S.I.	Pas de Mention latine S.n.	Pas de Mention latine S.d.	0	0	0	0	0	0
188974156	2015	2015	2	2	Aucun double espace dans le titre.	0	Mention latine S.I.	Pas de Mention latine S.n.	Pas de Mention latine S.d.	0	0	0	0	0	0

[Retour](#)



Fichier CSV

Usage

Cela peut servir à :

Faire un contrôle rapide sur la production de son ILN

Détecter des erreurs de catalogage récurrentes

Cela ne peut PAS servir à :

Faire d'autres contrôles comme la conformité des données codées, la validité des ISBN et ISSN, etc.

Je veux vérifier la **qualité des notices de thèses** produites dans mon établissement





AlgoSudoc

<http://www.theses.fr/algosudoc?>

détecte **des améliorations** à apporter aux notices **ou des erreurs bloquantes** empêchant le chargement des données dans theses.fr sur :

- Unicité du numéro national de thèse (029)
- Date et langue (100\$a et 101\$a)
- Mention de thèse originelle (105\$bm)
- Note de thèse structurée (328)
- Absence/erreur de zones de lien 4XX
- Code domaine TEF 686
- Présence d'un lien auteur 700

✓ URL à paramétrer : par ILN ou RCR, par Type de document...

✓ Export des résultats (fichier .csv)



Guide Méthodologique > Manuels > Contrôle catalographique > AlgoSudoc, rapports de chargement des données dans theses.fr



AlgoSudoc

<http://www.theses.fr/algosudoc?>

Exemple pour une erreur bloquante

<http://www.theses.fr/algosudoc?type=erreur&rownum=10&sort=nnt&codeEtab=AMIE>

Rapport généré le 2017-10-11T15:03:32Z								
Url utilisée : http://www.theses.fr/AlgoSudoc?type=erreur&rownum=10&sort=nnt&codeEtab=AMIE								
CodeEtab	RCR	Année	NNT	PPN	Etat	Gravité	Libellé	Consigne
AMIE	800212102	1992	1992AMIEM069	41078802	56	Bloquante	Une notice de thèse originelle doit contenir la mention d'un directeur de thèse en 702	Ajoutez le directeur de thèse dans une zone 702 et ajoutez le code de fonction \$4727. Créez sa notice d'autorité et faites le lien. Reportez cet ajout aux éventuelles notices décrivant la thèse non originelle.



AlgoSudoc

<http://www.theses.fr/algosudoc?>

Exemple pour une amélioration à apporter

[http://www.theses.fr/algosudoc?type=amelioration
&rownum=10&sort=nnt&rcr=800212102](http://www.theses.fr/algosudoc?type=amelioration&rownum=10&sort=nnt&rcr=800212102)

Rapport généré le 2017-10-11T15:24:48Z						
Url utilisée : http://www.theses.fr/AlgoSudoc?type=amelioration&rownum=10&sort=nnt&rcr=800212102						
CodeEtab	RCR	Année	PPN	Gravité	Libellé	Consigne
null	800212102	2010	145918270	Non bloquante : cette no	Déduction de lien par le lien	Corrigez les notices pour faire les liens réciproques.
				Bloquante : cette notice n'est pas versé dans		
null	800212102	2010	145918270	theses.fr	Déduction de lien par l'auteur	Corrigez les notices pour faire les liens réciproques.

Je veux vérifier les **liens d'indexation**
« Rameau » dans une notice





Script « ContrôleSyntaxe60X »

détecte les erreurs de liage sur une notice :

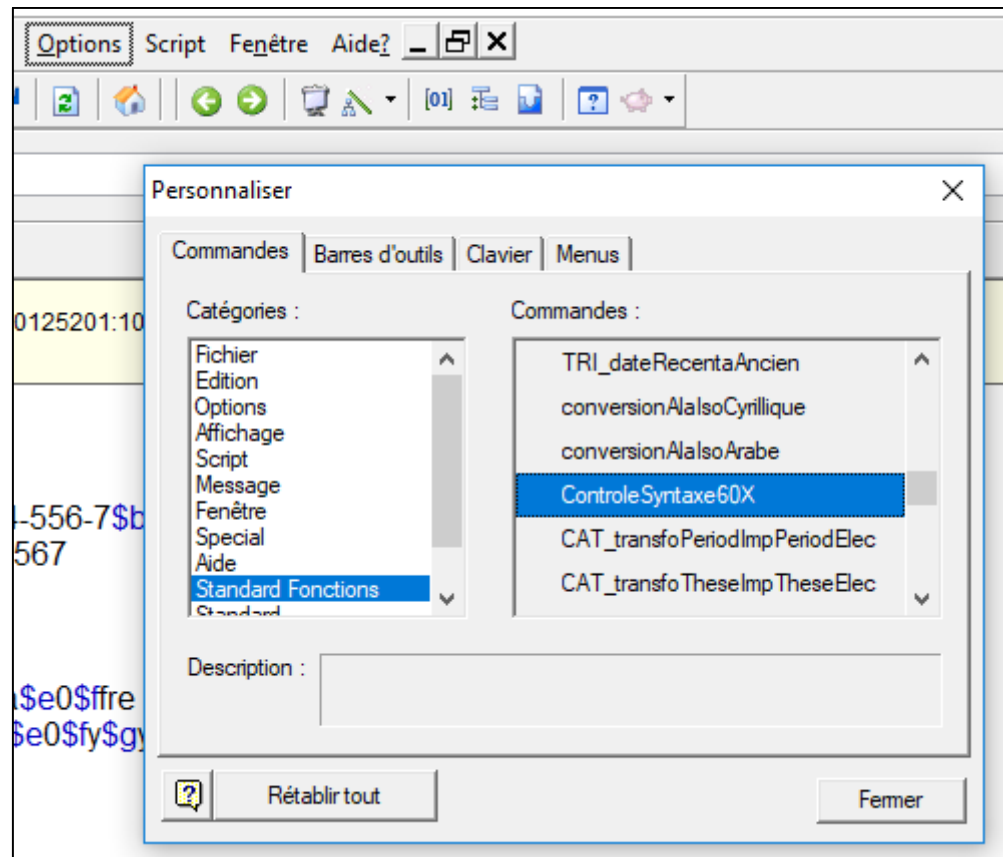
- **liens non faits**
- **liens faux syntaxiquement** quand le type de la notice d'autorité liée n'est pas autorisé dans la zone 6XX où le ppn se trouve





Script « ContrôleSyntaxe60X »

Installer le script sur vos barre d'outils depuis le Menu **Options > Personnaliser > Commandes > Standard Fonctions**



Script « ContrôleSyntaxe60X »

Exemple



1. Activer le script

2. ANALYSER : cliquer sur « contrôle de la syntaxe des zones 60X »

3. RESULTAT : Détection d'un lien incompatible

4. CORRIGER : modifier 600 en 607

The screenshot shows the 'ControleSyntaxe60X' application window. The main area displays a list of items under 'Lot 24'. A dialog box titled 'Controle Syntaxe zone 60X' is open, showing a message: 'Le script va parcourir toute la notice quand vous cliquerez !' and a button labeled 'Contrôle de la syntaxe des zones 60X'. Below this, another dialog box titled 'Controle de syntaxe zones 60X' displays an error message: 'zone 600 occurrence 1 : ppn 030200105 Tête de vedette Tg incompatible avec l'étiquette de la zone'. The list of items includes various codes and descriptions, with the entry '600 #1\$3030200105@Afrique orientale -- Civilisation\$2rameau' highlighted in red.

ControleSyntaxe60X

CBS

Controle Syntaxe zone 60X

Le script va parcourir toute la notice quand vous cliquerez !

Contrôle de la syntaxe des zones 60X

Annuler

Controle de syntaxe zones 60X

zone 600 occurrence 1 : ppn 030200105 Tête de vedette Tg incompatible avec l'étiquette de la zone

OK

Lot 24 | Nombre de

Création: 999999997:1

002 \$a2000070321\$2

003 http://www.sudoc.fr

008 \$aAar2

010 ##\$a2-221-07868-3\$bBr\$d179 FRF

035 ##\$aelec2000070321

100 0#\$a2000

101 0#\$afr

102 ##\$aFR

104 ##\$b|\$c|\$dba\$e0\$ffre

105 ##\$c0\$d|\$e|\$fy\$g|

106 ##\$ar

200 1#\$a@Voyages de découvertes

210 ##\$aParis\$cLaffont\$d2000

215 ##\$a1110 p\$d20 x 13 cm

225 2#\$a@Bouquins\$x0244-5913

410 ##\$t@Bouquins\$x0244-5913

600 #1\$3030200105@Afrique orientale -- Civilisation\$2rameau

606 ##\$3027242285@Récits de voyages\$2rameau

606 ##\$3027225550@Explorateurs\$2rameau

676 ##\$a916.2 600\$v99

676 ##\$a914\$v99a

702 #1\$aRicard\$bAlain \$4340

801 #3\$aFR\$bElectre\$c20000504\$gAFNOR



Usage

Cela peut servir à :

Vérifier le travail en continu

Intervenir rapidement pour corriger

Je veux vérifier les **liens aux « autorités »** des notices produites





AlgoLiens

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

évalue le liage de chaque notice et détecte les **liens manquants** en **7XX, 5XX, 60X**

- URL à paramétrer : par ILN ou RCR, par Type de document...
- Export des résultats (fichier .csv)



Guide Méthodologique > Manuels > Contrôle catalographique > AlgoLiens



Plateforme d'autoformation > Utilisateurs du réseau Sudoc > Contrôle Qualité



AlgoLiens

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

Exemple pour l'ILN 004 (4 résultats d'erreurs demandés)

<http://www.idref.fr/AlgoLiens?iln=4&rownum=4>

Url utilisée : http://www.idref.fr/AlgoLiens?iln=4&rownum=4							
PPN	ILN	RCR	Texte	Date	Code	TypDoc	LibelleTypDoc
195614224	4	130552101	une 607 sans lien	2016-	B607	Aa	Monographie imprimée
195614224	4	130552101	une 700 sans lien	2016-	B700	Aa	Monographie imprimée
195614224	4	130552101	une 702 sans lien	2016-	B702	Aa	Monographie imprimée
195614224	4	130552101	une 710 sans lien	2016-	B710	Aa	Monographie imprimée

Je veux vérifier le **liage complet** d'une revue
(avec historique)





Metarevues

<http://www.sudoc.fr/services/metaperiodical/PPN>

génère l'historique complet d'une revue - imprimée et électronique

- ✓ URL à paramétrer avec un n°PPN
- ✓ Export des résultats (fichier .csv)





Metarevues

<http://www.sudoc.fr/services/metaperiodical/PPN>

Exemple pour la revue « le nouvel observateur »

<http://www.sudoc.fr/services/metaperiodical/039225666>

ISSN	PPN	TITLE	PUBLISHER	DATE	MEDIA	STARTDATE	ENDDATE
1273-8980	038887258	L'Observateur politique, économique et littéraire	Paris : [éditeur inconnu] , 1950-1953	(1950)-(1953)	a	1950	1953
1273-8999	038887266	L'Observateur d'aujourd'hui (Paris)	Paris : [éditeur inconnu] , 1953-1954	(1953)-(1954)	a	1953	1954
1273-9006	037456407	France-observateur (Paris)	Paris : [éditeur inconnu] , 1954-1964	(1954)-(1964)	a	1954	1964
0029-4713	039225666	Le Nouvel observateur (Paris)	Paris : Nouvel observateur , 1964-2014	(1964)-(2014)	a	1964	2014
2416-8793	181387700	L'Obs	Paris : Nouvel observateur , 2014-	(2014)-...	a	2014	
2490-9947	191088757	L'Obs (En ligne)	Paris : le Nouvel Observateur du monde , 2014-	(2014)-...	l	2014	

EN RÉSUMÉ

- Consulter la documentation

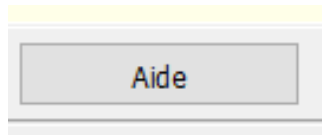


Guide Méthodologique > Manuels > Interrogation > **Liste des index**

Guide Méthodologique > Manuels > Contrôle catalographique > **AlgoSudoc**

Guide Méthodologique > Manuels > Contrôle catalographique > **AlgoLien**

Guide Méthodologique > Manuels > Logiciel WinIBW > **Script**



WinIBW > Bouton Aide

- Utiliser les index de recherche utiles (RBC, CRN, SEL)
- Construire des requêtes avancées
- Créer des scripts : *faciliter vos procédures répétitives en les automatisant*
- Utiliser les webservices (lien, thèses) de façon régulière