



Projet de numérisation des documents du Conseil du CERN

*- Sandrine Reyes -
Archives Historiques du CERN*



Le Conseil du CERN

- ❑ Le Conseil : Organe de tutelle du CERN
- ❑ 4 Réunions par an
- ❑ Délégations des 20 Etats Membres
- ❑ Président du Conseil + Directeur-Général
 - ❑ recherche en physique des particules
 - ❑ adopter les budgets
 - ❑ approuver ses programmes
 - ❑ fixer ses objectifs
 - ❑ nommer les membres de sa Direction
 - ❑ etc....
- ❑ Conformément à la Convention du CERN adoptée le 1^{er} juillet 1953 , le Conseil a également établi
 - ❑ le Comité des Directives Scientifiques (SPC) et
 - ❑ le Comité des Finances (FC)



Projet de numérisation

- En 2008, le Président du Conseil a demandé au groupe Information Scientifique de donner l'accès aux documents provenant du Conseil tout en respectant les règles d'accès
- Ce projet de numérisation de documents et gestion des métadonnées a commencé en janvier et doit se terminer en décembre 2009



Sommaire

- Analyse :
 - documents
 - métadonnées
 - outils
- Objectifs
- Procédure de travail :
 - Grille de saisie – Standardisation des champs
 - Uploader – Format MARC21
 - Codes-Barres - Numérisation
 - Titres français
 - Xenu – OCR - Chkall
- Résultats :
 - Visualisation sur CERN Documents Server (CDS) (format bref et détaillé)
 - Recherche sur CDS (dans les métadonnées et les fulltexts)



Analyse : Documents du Conseil

- ❑ Procès-verbaux édités généralement dans les 2 langues officielles du CERN : Anglais et français
- ❑ Un procès-verbal peut concerner 1-2 voire 3 comités
- ❑ Chaque document possède au minimum une référence : CERN/0250
- ❑ Respecter la confidentialité des documents



Analyse : Métadonnées

- ❑ Provient d'une base FileMakerPro tenue par le Secrétariat du Conseil
- ❑ 7 000 Notices accessibles sur le Serveur de Documents du CERN (CDS)
- ❑ Notices manquantes
- ❑ Métadonnées non conformes au format MARC21
- ❑ 2 Notices : Anglais – Français



Analyse : Outils

- ❏ Notices accessibles sur CDS-Invenio :
Extraction et Importation de données
- ❏ Possibilité de travailler sur BibEdit pour un travail à l'unité (Editeur développé au CERN pour remplacer le module de catalogage du système d'automatisation utilisé par la bibliothèque)
- ❏ Possibilité de travailler en série sous UNIX avec l'éditeur Emacs



Analyse : Outils

Exemples de notices sur BibEdit

BibEdit: Record #18641 [\(view history\)](#)

Record

18641 Rec ID

Search

Submit Cancel

View

Fields

Ready

Help

<input type="checkbox"/> 001		18641
<input type="checkbox"/> 003		SzGeCERN
<input type="checkbox"/> 035_	<input type="checkbox"/> \$9 ADMADM	
	<input type="checkbox"/> \$a 0001253	
<input type="checkbox"/> 041_	<input type="checkbox"/> \$a fre	
<input type="checkbox"/> 088_	<input type="checkbox"/> \$a CERN/FC/0500	
<input type="checkbox"/> 111_	<input type="checkbox"/> \$d 19 October 1961	
	<input type="checkbox"/> \$n FC 40	
<input type="checkbox"/> 245_	<input type="checkbox"/> \$a Transferts et dépassements de crédits	
<input type="checkbox"/> 260_	<input type="checkbox"/> \$c 1961	
<input type="checkbox"/> 269_	<input type="checkbox"/> \$c 30 septembre 1961	
<input type="checkbox"/> 595_	<input type="checkbox"/> \$a 2240/e	
<input type="checkbox"/> 65017	<input type="checkbox"/> \$2 SzGeCERN	
	<input type="checkbox"/> \$a BUDGET Transfers and Expenditure	
<input type="checkbox"/> 916_	<input type="checkbox"/> \$s n	
	<input type="checkbox"/> \$w 200319	
<input type="checkbox"/> 960_	<input type="checkbox"/> \$a 61	
<input type="checkbox"/> 961_	<input type="checkbox"/> \$c 20031203	
	<input type="checkbox"/> \$h 0854	
	<input type="checkbox"/> \$i MAND1	
	<input type="checkbox"/> \$x 19990126	
<input type="checkbox"/> 963_	<input type="checkbox"/> \$a PUBLIC	
<input type="checkbox"/> 970_	<input type="checkbox"/> \$a 000001253MAN	
<input type="checkbox"/> 980_	<input type="checkbox"/> \$a FINANCECOMM	

BibEdit: Record #18640 [\(view history\)](#)

Record

18640 Rec ID

Search

Submit Cancel

View

Fields

Ready

Help

<input type="checkbox"/> 001		18640
<input type="checkbox"/> 003		SzGeCERN
<input type="checkbox"/> 035_	<input type="checkbox"/> \$9 ADMADM	
	<input type="checkbox"/> \$a 0001252	
<input type="checkbox"/> 041_	<input type="checkbox"/> \$a eng	
<input type="checkbox"/> 088_	<input type="checkbox"/> \$a CERN/FC/0500	
<input type="checkbox"/> 111_	<input type="checkbox"/> \$d 19 October 1961	
	<input type="checkbox"/> \$n FC 40	
<input type="checkbox"/> 245_	<input type="checkbox"/> \$a Transfers and Expenditure in Excess of Appropriations	
<input type="checkbox"/> 260_	<input type="checkbox"/> \$c 1961	
<input type="checkbox"/> 269_	<input type="checkbox"/> \$c 30 September 1961	
<input type="checkbox"/> 595_	<input type="checkbox"/> \$a 2240/e	
<input type="checkbox"/> 65017	<input type="checkbox"/> \$2 SzGeCERN	
	<input type="checkbox"/> \$a BUDGET Transfers and Expenditure	
<input type="checkbox"/> 916_	<input type="checkbox"/> \$s n	
	<input type="checkbox"/> \$w 200319	
<input type="checkbox"/> 960_	<input type="checkbox"/> \$a 61	
<input type="checkbox"/> 961_	<input type="checkbox"/> \$c 20031203	
	<input type="checkbox"/> \$h 0854	
	<input type="checkbox"/> \$i MAND1	
	<input type="checkbox"/> \$x 19990126	
<input type="checkbox"/> 963_	<input type="checkbox"/> \$a PUBLIC	
<input type="checkbox"/> 970_	<input type="checkbox"/> \$a 000001252MAN	
<input type="checkbox"/> 980_	<input type="checkbox"/> \$a FINANCECOMM	



Objectifs :

- ❑ Obtenir 1 notice pour les 2 versions
- ❑ Enregistrer les documents manquants
- ❑ Adapter les notices au format MARC21
- ❑ Accéder à 2 fichiers électroniques incluant l'OCR
- ❑ Respecter les règles d'accès
Confidentiel : 30 ans – Non confidentiel : 5 ans
- ❑ Donner une visualisation plus appropriée de la notice sur CDS (format bref et détaillé)



Procédure :

Grille de saisie - Standardisation

■ Établir une grille de catalogage standard en décrivant l'ensemble des champs qui peuvent être utilisés dans cette série

■ Standardiser les champs 111, 711, 269

269__c : 16 / 17 juin 1965 ⇒ 269__c : 16 - 17 Jun 1965

269__c : 6 février 1965 ⇒ 269__c : 06 Feb 1965



Procédure :

Uploader – Format MARC21

- ☐ Formater les métadonnées avec l’outil Uploader grâce à une configuration spécifique à cette source – Catalogage conforme au format MARC21

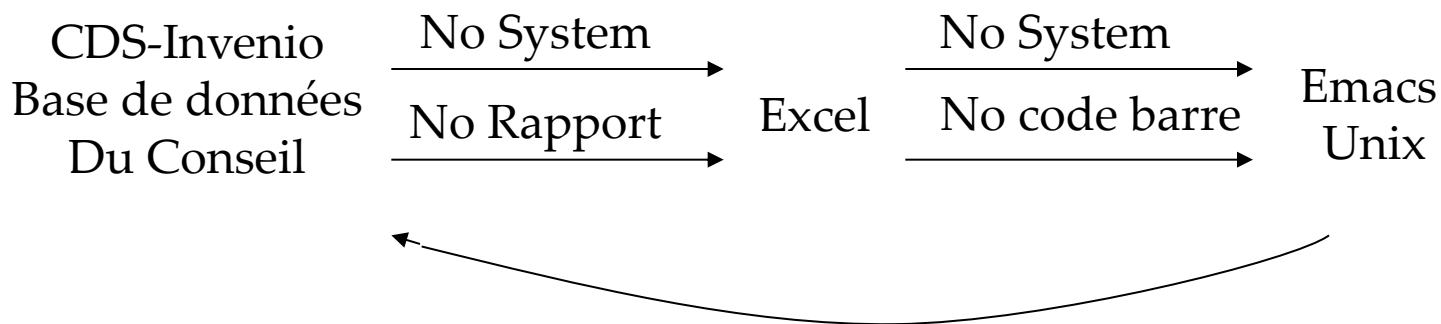
<u>AVANT</u>	<u>APRES</u>
041 __ a eng	041 __ a eng 041 __ a fre
111 __ d 16 - 17 Jun 1965 __ n C 30	111 __ 9 19650616 __ a 30th Session of Council __ n 30 __ c CERN, Geneva, Switzerland __ d 16 – 17 Jun 1965 __ f 1965 __ g cern19650616 __ w CH __ z 19650617
711 __ d 20 – 21 jun 1965 __ n FC 25	711 __ 9 19650620 __ a 25th Meeting of Finance Committee __ n 25 __ c CERN, Geneva, Switzerland __ d 20 – 21 Jun 1965 __ f 1965 __ g cern19650620 __ w CH __ z 19650621



Procédure :

Codes-barres - Numérisation

- ❑ Equiper les documents d'un code-barre pour la numérisation
- ❑ Ajouter ces codes-barres dans le champs 088__9 des notices anglaises : Extraction et importation de données



- ❑ Envoyer les documents au service du scanning



Procédure :

Exemple tableau Excel – fichier Unix

	A	B	C	D	E
1	19268	088	\$\$\$	CM-P00076660	-e CERN/0500
2	19309	088	\$\$\$	CM-P00076661	-e CERN/0503
3	19305	088	\$\$\$	CM-P00076662	-e CERN/0505
4	19319	088	\$\$\$	CM-P00076663	-e CERN/0506
5	19315	088	\$\$\$	CM-P00076664	-e CERN/0506/Rev.
6	19380	088	\$\$\$	CM-P00076665	-e CERN/0507/Draft
7	19373	088	\$\$\$	CM-P00076666	-e CERN/0510
8	19383	088	\$\$\$	CM-P00076667	-e CERN/0510/Rev.
9	19351	088	\$\$\$	CM-P00076668	-e CERN/0511
10	19393	088	\$\$\$	CM-P00076669	-e CERN/0512
11	19391	088	\$\$\$	CM-P00076670	-e CERN/0513
12	19456	088	\$\$\$	CM-P00076671	-e CERN/0514/Add.
13	19458	088	\$\$\$	CM-P00076672	-e CERN/0514/Draft
14	19402	088	\$\$\$	CM-P00076673	-e CERN/0516
15	19406	088	\$\$\$	CM-P00076674	-e CERN/0516/Appendices
16	19465	088	\$\$\$	CM-P00076675	-e CERN/0517
17	19525	088	\$\$\$	CM-P00076676	-e CERN/0517/Corr.

```
[lxlplus225] ~/public > more 088_9-5
19268 088_  $$$CM-P00076660-e
19309 088_  $$$CM-P00076661-e
19305 088_  $$$CM-P00076662-e
19319 088_  $$$CM-P00076663-e
19315 088_  $$$CM-P00076664-e
19380 088_  $$$CM-P00076665-e
19373 088_  $$$CM-P00076666-e
19383 088_  $$$CM-P00076667-e
19351 088_  $$$CM-P00076668-e
19393 088_  $$$CM-P00076669-e
19391 088_  $$$CM-P00076670-e
19456 088_  $$$CM-P00076671-e
19458 088_  $$$CM-P00076672-e
19402 088_  $$$CM-P00076673-e
19406 088_  $$$CM-P00076674-e
19465 088_  $$$CM-P00076675-e
19525 088_  $$$CM-P00076676-e
19430 088_  $$$CM-P00076677-e
19477 088_  $$$CM-P00076678-e
19444 088_  $$$CM-P00076679-e
19463 088_  $$$CM-P00076680-e
19410 088_  $$$CM-P00076681-e
```

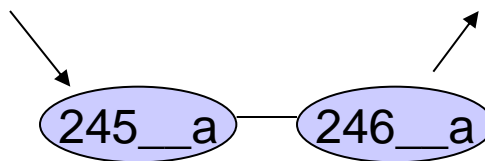


Procédure :

Titres français

- 📄 Ajouter les titres français en 246__a dans les notices anglaises : Extraction et importation de données

	A	B	C	D
1	FRE		ENG	
2	19296	CERN/CC/0502	19295	CERN/CC/0502
3	19308	CERN/CC/0504	19307	CERN/CC/0504
4	19350	CERN/CC/0508/Draft	19349	CERN/CC/0508/Draft
5	19354	CERN/CC/0509	19353	CERN/CC/0509
6	19486	CERN/CC/0515/Draft	19485	CERN/CC/0515/Draft
7	19443	CERN/CC/0520	19442	CERN/CC/0520
8	19501	CERN/CC/0528	19500	CERN/CC/0528
9	19599	CERN/CC/0532/Draft	19598	CERN/CC/0532/Draft
10	19574	CERN/CC/0534	19573	CERN/CC/0534
11	19583	CERN/CC/0535	19582	CERN/CC/0535
12	19669	CERN/CC/0545	19668	CERN/CC/0545
13	19712	CERN/CC/0545/Rev.	19711	CERN/CC/0545/Rev.
14	19691	CERN/CC/0548/Draft	19690	CERN/CC/0548/Draft
15	19737	CERN/CC/0555	19736	CERN/CC/0555
16	19778	CERN/CC/0555/Rev.	19777	CERN/CC/0555/Rev.
17	19752	CERN/CC/0556/Draft	19751	CERN/CC/0556/Draft
18	19822	CERN/CC/0559/Draft	19821	CERN/CC/0559/Draft





Procédure :

Xenu – OCR - ChKall

- ❏ Supprimer les notices françaises de la base
- ❏ Programme Xenu pour détecter les adresses incorrectes des documents revenus du Scanning
- ❏ Envoyer ces adresses au Service Informatique qui utilise le programme OCRopus développé par Google pour la reconnaissance de texte
- ❏ Programme ChKall – Outil configuré pour contrôler le bon formatage et le bon contenu des métadonnées



Résultats : Visualisation sur CDS

En format Bref

Council and Committee of Council

1 notices trouvées


La recherche a duré 0.01 secondes.


1. **The Ford Foundation and the CERN Visiting Scientist Programme**

La Fondation Ford et le programme des savants visiteurs de CERN

CERN/0500

24th Session of Council ; 1963

French -  [PDF](#)

English -  [PDF](#)



[Notice détaillée](#) - [Notices similaires](#)

AJOUTER AU PANIER

En format Détaillé



Council

Report number	CERN/0500
Title	The Ford Foundation and the CERN Visiting Scientist Programme
Titre français	La Fondation Ford et le programme des savants visiteurs de CERN
Meeting title	24th Session of Council 20 Jun 1963 - CERN, Geneva, Switzerland
Date of document	1963
Access to documents	French -  PDF English -  PDF



Résultats :

Recherche sur CDS dans les métadonnées

Council and Committee of Council

Recherche:

weisskopf

tous les champs

Recherche

Liste

[Conseils de recherche](#) :: [Recherche avancée](#)

Recherche collections:

Council and Committee of Council

*** ajouter une nouvelle collection ***

Trier par:

les plus récents en premier

décroissant

- ou ordonner par -

Afficher:

10 résultats

regrouper par collection

Format de sortie:

HTML brief

Council and Committee of Council

1 notices trouvées

La recherche a duré 0.00 secondes.

- 1. **Proposed Composed of the Scientific Policy Committee (by the Chairman of the Committee) (Weisskopf, Lehmann, Salvini)**

Composition proposée du Comité des Directives scientifiques (par le Président du Comité)

CERN/1191

55th Session of Council ; 1975

French - [PDF](#)

English - [PDF](#)

[Notice détaillée](#) - [Notices similaires](#)



Résultats :

Recherche sur CDS dans les fulltexts

Council and Committee of Council

Recherche:

fulltext:weisskopf

tous les champs

Recherche

Liste

[Conseils de recherche](#) :: [Recherche avancée](#)

Recherche collections:

Council and Committee of Council

*** ajouter une nouvelle collection ***

Trier par:

Afficher:

Format de sortie:

les plus récents en premier

décroissant

- ou ordonner par -

10 résultats

regrouper par collection

HTML brief

Council and Committee of Council 125 notices trouvées 1 - 10 aller vers la notice: 1 La re

- 1. **The Fellowship Programme**
Programme pour les boursiers
CERN/0273
10th Session of the Council ; 1958
English - [PDF](#)
French - [PDF](#)
[Notice détaillée](#) - [Notices similaires](#)
- 2. **Progress Reports of the Director-General and Divisional Directors**



Questions ?