

Une bibliothèque numérique encyclopédique,
Un atelier flexible d'analyse de corpus,
pour chercheurs et praticiens
des sciences expérimentales

WicriExplore

Introduction

De la musicologie au bois en passant par la santé

Bonjour, mon nom est Jacques Ducloy

- ▶ Octobre 67... de la règle à calcul aux cartes perforées
 - sur un ordinateur à 16 K mots de 18 bits...
- ▶ 73 : ingénieur système : un Iris 80 pour la fac
 - Et le TLF 1000 textes, un mois pour une analyse statistique
- ▶ 80 : création d'un GS valorisation de logiciels (Génie Logiciel -IA)
- ▶ 88 : directeur informatique de l'INIST
 - CNRS : 400 personnes pour 500.000 analyses d'articles par an
- ▶ 92 : Recherche appliquée (INIST, INRIA)
 - XML et l'analyse de corpus
- ▶ 2000 : Directeur Produits et Services à l'INIST
- ▶ 2005 : Lancement d'une revue numérique (AMETIST)
- ▶ 2007 : DRRT Lorraine, lancement WICRI
- ▶ 2010 : Projet LorExplor (ISTEX) Loria -> Paragraphe Paris 8

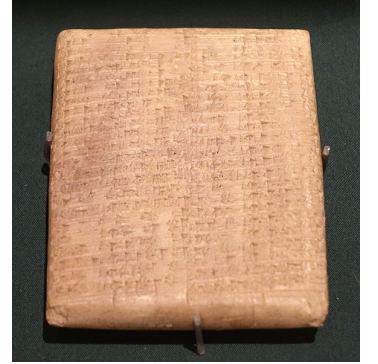
Dilib / Wicri / LorExplor / WicriExplore

- ▶ Dilib (INIST / Loria 1992 – 2002)
 - Une boîte à outils XML (*Document & Information Library*)
 - Un lego pour construire un système de recherche d'information
- ▶ Wicri (Univ. Lorraine 2008 – ...)
 - Un réseau de wikis sémantiques (
 - Wikis des Communautés de la Recherche et de l'Innovation)
- ▶ LorExplor (UL 2014 + ISTEEX – ...)
 - Exploration des besoins lorrains avec IITEX 22.000.000 d'articles)
- ▶ Et maintenant WicriExplore (Paragraphe 2020)
 - Applications de l'ensemble
 - avec un point fort dans les humanités (musicologie)
 - Et maintenant sur la Santé et les interactions arbres micro organismes



Changements de paradigmes pour les documents

PAGE



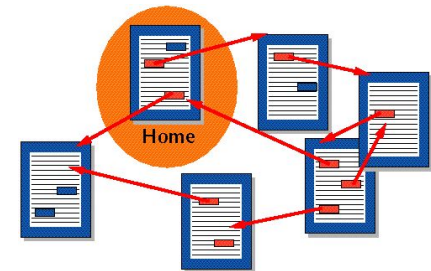
- ▶ 7000 av. J.-C. : jetons
- ▶ 3000 av. J.-C. : tablettes, papyrus, papyrus en rouleau

CODEX



- ▶ 300 av. J.C. : Codex manuscrit
- ▶ 1455 : Livre imprimé ; 1750 : Encyclopédie
- ▶ 1956 : Fieldata ; 1973 (Xerox Alto) ; 1984 : MacWrite ;
- ▶ 1992 : PDF

HTPERTEXTE



- ▶ 1963 : Hypertext ; 1987 : Hypercard ;
- ▶ 1989 : Html
- ▶ 2001 : Wikipédia (collectif)

Changements de paradigme dans les bibliothèques



Une bibliothèque => Une URL

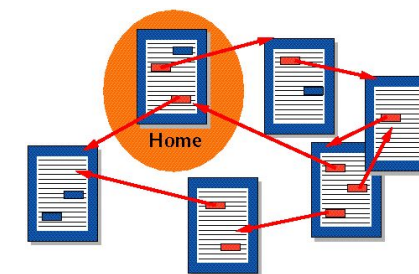
<http://gallica.bnf.fr>

<https://lorexplor.istex.fr/Wicri>

Wicri : un rayon => un wiki

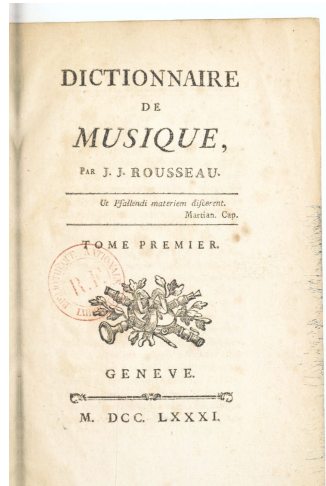


Des livres juxtaposés => hypertexte



Réédition numérique hypertexte Dictionnaire de musique (J.-J. Rousseau)

Romilly

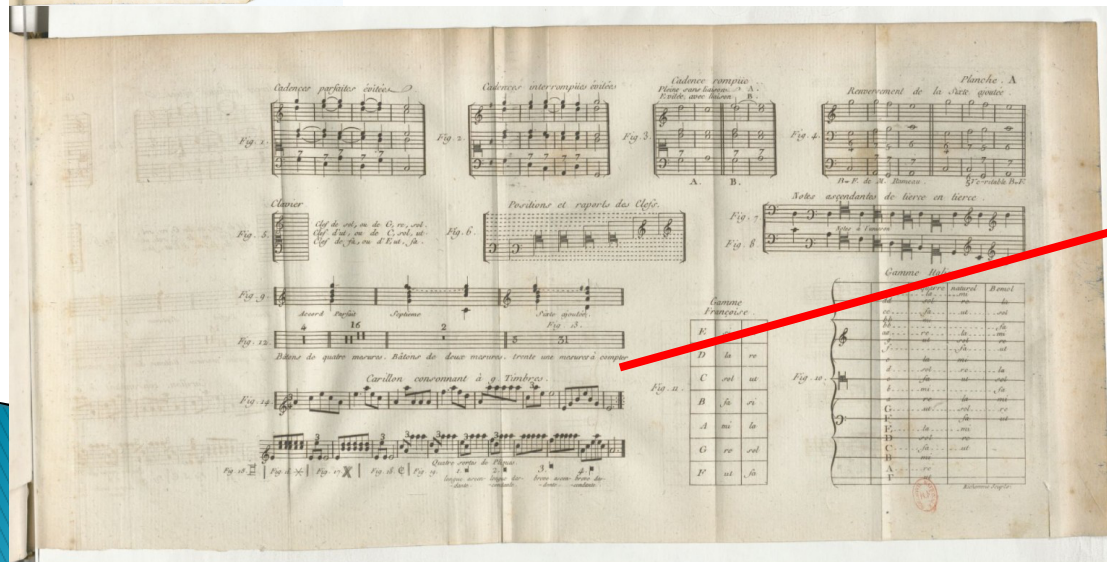


Dans le tome 1

pays des vaines spéculations.
CARRILLON. Sorte d'Air fait pour être exécuté par plusieurs Cloches accordées à différens Tons. Comme on fait plutôt le Carrillon pour les Cloches que les Cloches pour le Carrillon, l'on n'y fait entrer qu'autant de Sons divers qu'il y a de Cloches. Il faut observer de plus, que tous leurs Sons ayant quelque permanence, chacun de ceux qu'on frappe doit faire Harmonie avec celui

Lien

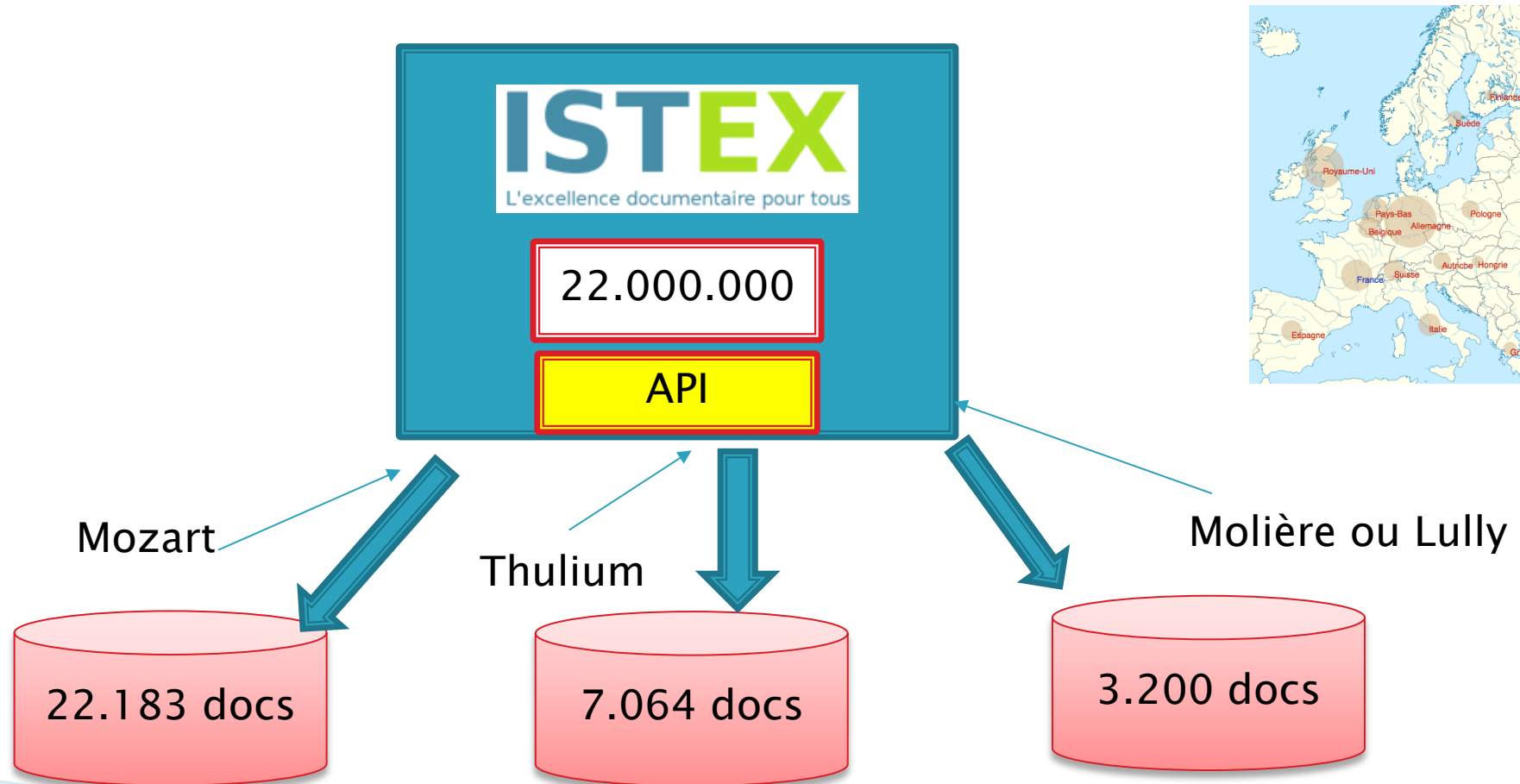
Dans le tome 2



Le lecteur « non solfégiste » peut écouter la musique

ISTEX

des millions de documents à explorer



En santé : volumétrie débordante

▶ Covid

- Google : 6 270 000 000 pages
- Google Scholar : 1 830 000 pages ou articles scientifiques (1 580 000)
- PubMed : 75 800 / 30.000.000 (24 913)
- PubMed Central : 76.000 (25 552)
- HAL : 1 600 (562)
- Wikipédia (fr) : 8 175 (4 877)

▶ Sur WicriExplore , thématique des épidémies grippales

- 24 serveurs d'exploration
- 70.000 documents

Le Covid avec WicriExplore

| | ◆ HAL ◆ | ISTEX ◆ | PubMed ◆ | PMC ◆ | Pascal ◆ | Francis ◆ | Total ◆ | Total corrigé ◆ |
|---|---------|---------|----------|-------|----------|-----------|---------|-----------------|
| Serveur d'exploration Chloroquine | 262 | 4 500 | 701 | 1 125 | 104 | 0 | 6 692 | 6 435 |
| Serveur d'exploration Covid | 13 | 1 052 | 767 | 814 | 71 | 1 | 2 718 | 2 451 |
| Serveur d'exploration Covid (26 mars) | 13 | 1 052 | 1 421 | 1 144 | 71 | 1 | 3 702 | 3 214 |
| Serveur d'exploration H2N2 | 340 | 2 508 | 665 | 1 454 | 140 | 0 | 5 107 | 4 660 |
| Serveur d'exploration MERS | 25 | 4 000 | 4 326 | 2 049 | 118 | 1 | 10 519 | 8 450 |
| Serveur d'exploration SRAS | 203 | 3 707 | 5 359 | 2 563 | 989 | 6 | 12 827 | 10 380 |
| Serveur d'exploration Stress et Covid | 224 | 3 573 | 1 145 | 1 000 | 294 | 88 | 6 324 | 5 969 |
| Serveur d'exploration Tocilizumab | 86 | 2 264 | 3 019 | 1 825 | 263 | 1 | 7 459 | 6 869 |
| Serveur d'exploration sur la grippe au Canada | | | 1 386 | | | | 1 386 | 1 386 |
| Serveur d'exploration sur la grippe en Allemagne | | | 634 | | | | 634 | 634 |
| Serveur d'exploration sur la grippe en Belgique | | | 130 | | | | 130 | 130 |
| Serveur d'exploration sur les pandémies grippales | 837 | 3 967 | 3 311 | 1 574 | 3 046 | 113 | 12 849 | 10 482 |

Plan

1. Ecrire science et culture dans un wiki
2. Développements par les chercheurs avec les modèles
3. Maitriser les ontologies avec les wikis sémantiques
4. Coopérer, mutualiser avec les réseaux de wikis
5. Gérer la volumétrie avec l'ingénierie XML
6. Conclusion