



HAL
open science

Cabestan : à l'abordage des données linguistiques du Sudoc

Anna Svenbro

► **To cite this version:**

Anna Svenbro. Cabestan : à l'abordage des données linguistiques du Sudoc. Journées Abes 2024, May 2024, Montpellier, France. Zenodo, 2024, 10.5281/zenodo.11449450 . hal-04601803

HAL Id: hal-04601803

<https://hal.science/hal-04601803>

Submitted on 5 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cabestan



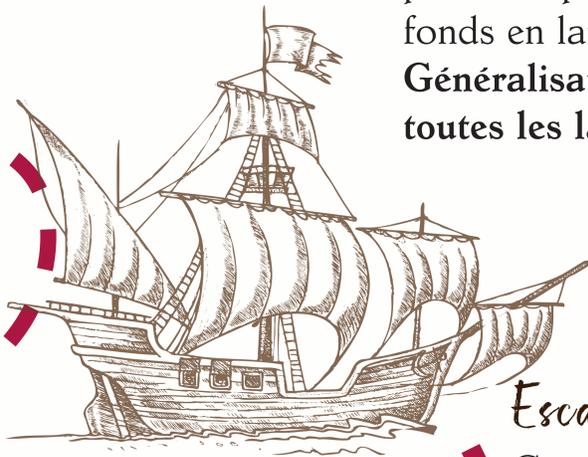
À l'abordage des données linguistiques du Sudoc !

L'aventure commence...

Mai 2023 Lancement de l'API SRU d'interrogation du Sudoc par l'ABES.

Objectif Initialement, utilisation de cette API combinée à celle d'IdRef pour comparer les collections de la Bibliothèque nordique avec d'autres fonds en langues scandinaves, finno-ougriennes et baltes.

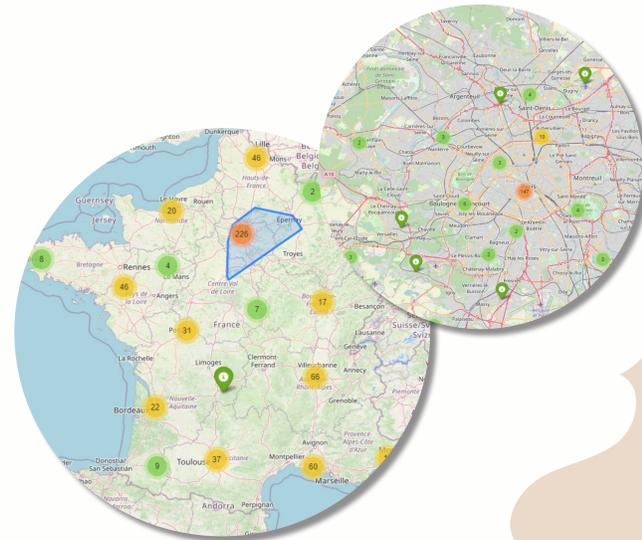
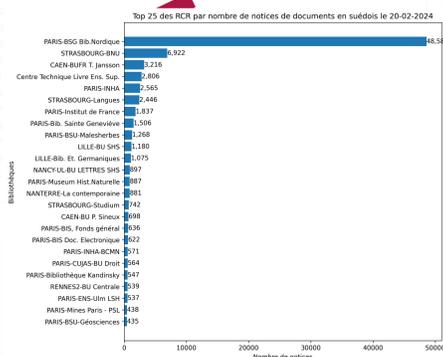
Généralisation du script pour établir une cartographie des fonds dans toutes les langues.



Escale : point méthodologie

CABESTAN utilise un script en **Python** pour localiser et quantifier les documents dans les bibliothèques du réseau Sudoc pour une langue donnée.

Il produit des données tabulées, des cartes localisant les bibliothèques, des diagrammes pour visualiser les résultats...



Premiers résultats !

Identifier, localiser et comparer les fonds les plus importants pour une langue donnée.

Utilité et publics visés

Les professionnels des bibliothèques

Pour piloter les politiques documentaires et valoriser les collections étrangères dans leurs établissements.

Les chercheurs

Pour localiser les fonds linguistiques et étudier la réception des langues étrangères en France.



Et maintenant... quelles perspectives ?

Développer CABESTAN en une plateforme web robuste.

Orienter son développement en recueillant et analysant les besoins des utilisateurs, notamment des bibliothécaires, des chercheurs et des étudiants.

Recourir à un prestataire extérieur pour le développement de l'infrastructure.

CABESTAN sera un outil précieux pour la cartographie et la quantification des fonds en langues étrangères du réseau Sudoc. Il offrira des perspectives prometteuses pour la recherche et la gestion des collections.