



HAL
open science

Conditions de la mutualisation : les principes FAIR et HTR-United

Alix Chagué

► **To cite this version:**

Alix Chagué. Conditions de la mutualisation : les principes FAIR et HTR-United. Humanistica 2022, May 2022, Montréal, Canada. . hal-03685731

HAL Id: hal-03685731

<https://inria.hal.science/hal-03685731v1>

Submitted on 2 Jun 2022

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.



Distributed under a Creative Commons Attribution 4.0 International License

Alix Chagué : Conditions de la mutualisation : les principes FAIR et HTR-United. Humanistica 2022.

Les principes FAIR

- Favoriser la science ouverte
- Données, métadonnées, infrastructures



FINDABLE



ACCESSIBLE



INTEROPERABLE



REUSABLE



<https://www.go-fair.org/fair-principles/>

Résumé de l'épisode précédent

Alix Chagué, Thibault Clérice, Laurent Romary.
HTR-United : Mutualisons la vérité de terrain !.
DHNord2021. [⟨hal-03398740⟩](https://hal.archives-ouvertes.fr/hal-03398740)

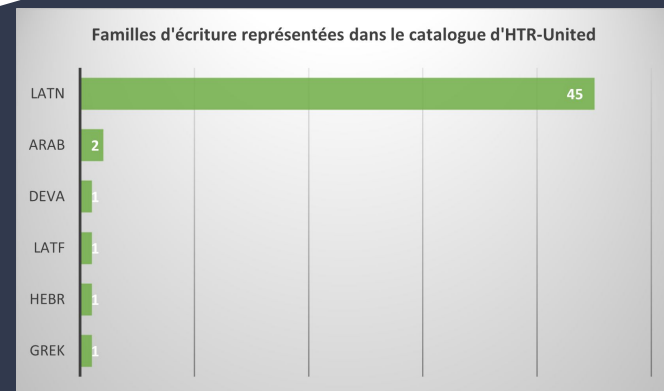
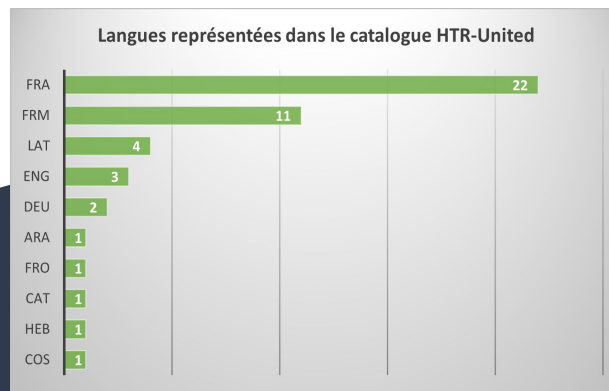
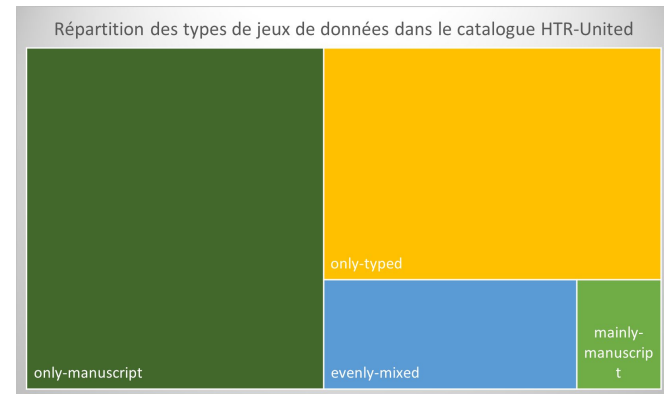
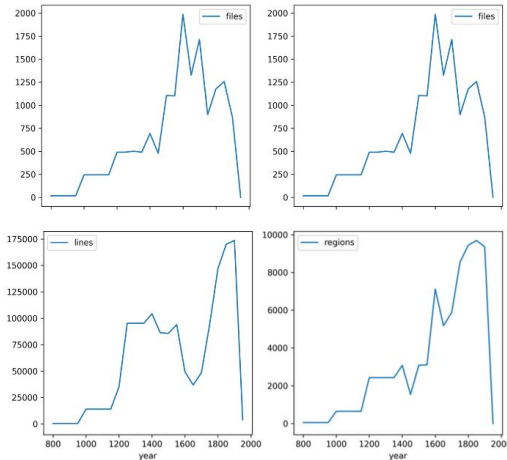
- HTR = Transcription automatique d'écritures manuscrites
- Enjeux important pour les GLAMs et la recherche
- Réduire le coût de production des modèles de transcription



<https://htr-extended.github.io>



- 47 jeux de données
- ~4 millions de caractères
- +220 000 lignes
- 19 projets contributeurs
- de l'an 800 à 1999



Un écosystème FAIR garanti par plusieurs composants

Standardisation des dépôts + template



Versionnage et transparence



Pérennisation par moissonnage



Catalogue facile à explorer et à parser

Charte de réutilisation des données + licences

Outils de contrôle qualité actionnables

Descriptions standardisées + schéma

