

MODELE DE FICHE CONSORTIUM CORLI

<https://corli.huma-num.fr/>

« Exploration de corpus : outils et pratiques » :

<http://explorationdecorpus.corpusecrits.huma-num.fr/>

Recensement d'outils et établissement d'une fiche par outil

Fiche descriptive, et fiches « d'usage », possibilité de « noter » l'outil avec des étoiles par exemple, fiche-page de TXM : <http://explorationdecorpus.corpusecrits.huma-num.fr/txm/>

FICHE DESCRIPTIVE DE L'OUTIL

Dernière version	TXM 0.7.6 (juillet 2014)
Responsable du projet	Serge Heiden
Accessibilité	Téléchargement
Site Web	http://textometrie.ens-lyon.fr/
Interface – Ergonomie	<p>► Interface conviviale : interface graphique utilisateur classique avec fenêtres multiples, widgets, barre d'outils et windows manager intégré (Eclipse RCP pour la version bureau ou GWT pour la partie client dans le navigateur web de la version portail) ;</p> <p>► Ergonomie : assistants de construction de requêtes, de sous-corpus ou de partitions, interface traduite en français, anglais et russe ;</p> <p>► Interface experte : version bureau pilotable également par scripts Groovy ou R, version portail interrogeable par API Restful (prototype).</p>
Systèmes	<p>► Windows (32 et 64 bits) ;</p> <p>► Mac OS X ;</p> <p>► Linux (32 et 64 bits) ;</p> <p>► Version portail Web ;</p>
Licence	GPLv3
Format des corpus en entrée	Une quinzaine de formats d'import déjà définis + import personnalisé possible par script. Les formats déjà définis vont du texte brut (presse-papier, fichiers .txt) au texte structuré (XML dont une dizaine d'usages TEI différents), en passant par des formats de production de corpus (Factiva, Europresse), de transcription de l'oral (Transcriber), de corpus annoté (Cordial, Tiger Search), de corpus parallèle (TMX) ou d'analyse de corpus (Hyperbase, Alceste). Travaille en UTF-8 mais peut importer et exporter d'autres encodages.
Format des corpus en sortie	<p>► Exports des tableaux par copier/coller ou au format tableur générique csv ;</p> <p>► Export des graphiques en svg, pdf, bmp, jpeg, png, gif ;</p> <p>► Production d'une version XML-TEI du corpus</p>
Fonctionnalités	<p>► Concordances ;</p> <p>► Cooccurences ;</p> <p>► Lexique et Index ;</p> <p>► Spécificités sur une partition</p>

	► Recherche de motifs...; voir plus de fonctionnalités
Annotations gérées par l'outil	Toute annotation XML, lexicale (étiquettes) ou/et textuelle (structuration) ; annotation Cordial ; possibilité d'ajout d'un étiquetage TreeTagger à la volée.
Documentation	Manuel d'utilisation
Type d'outil	Analyse

PRATIQUES ET USAGES DE TXM

- ***Fiches renseignées par :***

[Sergen Heiden](#) (ICAR, Université de Lyon) – [Céline Poudat](#) (BCL, Université de Nice Sophia Antipolis) – [Béatrice Turpin](#) (Université de Cergy Pontoise) – [Denis Vigier](#) (ICAR, Université Lumière Lyon II)

- ***Fiche 1 :***

Objectifs de recherche	Conception (S.Heiden) / Etude du genre de l'article de recherche (C.Poudat) / Etude quantitative et qualitative de la combinatoire de mots lexicaux ou grammaticaux (D. Vigier) / Analyse de corpus journalistiques (B.Turpin)
Pratiques	Concordance
Scénarios	Concordance sur corpus annotés, gestion du multiniveau avec CQP, utilisation d'expressions régulières dans tous les éléments de la requête, nombreuses possibilités de tris, paramétrage du contexte, retour au texte, interconnexion des fonctionnalités (cooccurrences, ventilations, tables lexicales...)

- ***Fiche 2 :***

Objectifs de recherche	Identifier et caractériser des usages linguistiques propres à une typologie textuelle
Pratiques	Mettre au jour les spécificités d'une partition
Scénarios	Spécificités, loi hypergéométrique

[Signaler des erreurs et/ou proposer des modifications](#)

Donner votre avis sur l'outil :

Interface - Ergonomie

Fonctionnalités

Facilité d'utilisation

Moyenne