

Corpus du consortium COSME² et leur correspondance aux principes FAIR

Corpus	Format(s)	Taille	Écosystème FAIR	Findable	Accessible	Interoperable	Reusable
Europange	pdf, jpg, sql	12 Go	infrastructure personnelle, Isidore	Oui	Oui	Partiellement	Partiellement
Transscript	sql, jpg	quelques Mo	Non FAIRisé*	Partiellement	Partiellement	Partiellement	Partiellement
Aposcripta	sql	quelques Mo	à venir	Partiellement	Partiellement	Partiellement	Partiellement
Armma	sql, jpg	texte: 21 Mo; images: 17,2 Go	à venir	Partiellement	Partiellement	Partiellement	Non
Bullarium Franciscanum	sql (TELMA), json (Zenodo)	quelques Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Cartulaires italiens	en cours de création sur TELMA	quelques Mo	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation
Cartulaires walllons	sql (TELMA), json (Zenodo)	quelques Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
CartulIR	sql (TELMA), json (Zenodo)	quelques Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Castellanie	xml, jpg	≈ 1Go	à venir	Partiellement	Partiellement	Partiellement	Partiellement
CBMA (Corpus Burgundiae Medii Aevi)	fmp12, csv, tab, xlsx, xml, jpg, txt, pdf, doc/docx	≈ 30-50 Mo	dépôt de données (Nakala)	Oui	Oui	Oui	Oui
CEMA (Cartae Europae Medii Aevi)	xml, txt, csv, ods, sql, pdf	textes: 700 Mo; bibliographie: < 1 Mo; édition numérisé (3200 pdf): 198 Go	en cours de création du jeu de données	Partiellement	Partiellement	Partiellement	Non

Corpus	Format(s)	Taille	Écosystème FAIR	Findable	Accessible	Interoperable	Reusable
CIFM-CBMA (Corpus des Inscriptions de la France Médiévale) (Corpus Burgundiae Medii Aevi)	csv (export de CBMA) > txt; xml (textes lemmatisés pour TXM)	≈ 1 Mo	dépôt de données (Nakala)	Oui	Oui	Oui	Oui
Col&Mon (Collégiales et monastères de la réforme carolingienne au Concile de Trente)	sql, jpg, png	500 Ko (base Col&Mon); 15 Mo (base Collégiales) + 100 Mo images; 10 Mo (base Monastères) + 1 Go images	à venir	Partiellement	Partiellement	Partiellement	Partiellement
Comptabilités Principautés S/E	xml, jpg	> 600 Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Devise	sql, jpg	texte: 5,2 Mo; images: 116 Mo	à venir	Partiellement	Partiellement	Partiellement	Non
Distinguo	xml	≈ 15 Mo	à venir	Partiellement	Partiellement	Partiellement	Partiellement
e-Cartae	xml, jpg	2,74 Mo (266 fichiers)	à venir	Partiellement	Partiellement	Partiellement	Partiellement
Golubovic	transformation en cours vers xml	quelques Mo	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation	Les données ne sont pas prêtes pour la FAIRisation
Lettres Jean XXII	sql (TELMA), json (Zenodo)	< 3 Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Originaux 1&2	sql	quelques Mo	Non FAIRisé*	Partiellement	Partiellement	Partiellement	Partiellement

Corpus	Format(s)	Taille	Écosystème FAIR	Findable	Accessible	Interoperable	Reusable
PALM (Plateforme d'analyse linguistique médiévale)	xml	22 Mo	à venir	Partiellement	Partiellement	Partiellement	Partiellement
Répertoire des anciens inventaires d'archives ecclésiastiques ...	xml (site AN, Zenodo); csv (migration vers TELMA)	123 Ko	infrastructure personnelle (salle des inventaires virtuelle des AN); dépôt de données (Zenodo)	Oui	Oui	Oui	Partiellement
Ressources médiévistes	sql (Heurist), csv (Zenodo)	< 1 Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Scripta	xml	65.4 Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Scriptorium	sql	20,7 Mo	à venir	Partiellement	Partiellement	Partiellement	Non
Sigilla	sql, jpg (Drupal), csv (Zenodo)	quelques Mo	dépôt de données (Zenodo)	Oui	Oui	Oui	Oui
Studium parisiense	documents json (MongoDB)	impossible à définir	à venir	Partiellement	Partiellement	Partiellement	Non
Titulus	xml, jpg	impossible à définir	à venir	Partiellement	Partiellement	Partiellement	Partiellement
Prosopange	sql, pdf, json, geoJson	62 Mo	à venir	Partiellement	Partiellement	Partiellement	Non

* - En l'absence d'un accord explicite de(s) auteur(s) du corpus, les données n'ont pas pu être FAIRisées.