POLITIQUE D'INDEXATION ET D’INTÉGRATION DE L'INSTRUMENT DE RECHERCHE DANS LE PORTAIL D’ARCHIVES DE L’AISNE

Table des matières

[1 Objet 1](#_Toc194140904)

[2 Choix des normes de métadonnées 1](#_Toc194140905)

[3 Indexation : 1](#_Toc194140906)

[4 Alignements sur les référentiels 1](#_Toc194140907)

[**5** **Typologie documentaire** 2](#_Toc194140908)

[6 Validation du document 2](#_Toc194140909)

[7 LIMITES – LIEN AVEC D’AUTRES FORMATS 2](#_Toc194140910)

# Objet

L’instrument de recherche doit correspondre aux attentes des Archives de l’Aisne mais aussi, pour répondre aux besoins d’interopérabilité des données, aux spécifications des plateformes agrégatives comme France Archives au niveau français ou par exemple Europeana au niveau européen.

# Choix des normes de métadonnées

Le format XML de l’instrument doit bien évidemment suivre le standard EAD-2002, mais doit pouvoir aussi être interopérable à terme avec des notices d’autres formats qui décrivent les producteurs (EAC\_CPF), les notices en EAG qui décrivent les institutions et les notices en EAF qui décrivent les fonctions.

# Indexation :

France Archives encourage vivement l'indexation des instruments de recherche selon les directives établies dans la note DITN/RES/2007/008.

Les éléments d'indexation sont principalement contenus dans l'élément <controlaccess> des fichiers XML-EAD. Cela améliore considérablement le référencement des instruments de recherche sur le moteur de recherche du portail.

De plus, l'auto-complétion proposée par le moteur de recherche s'appuie exclusivement sur ces termes d'indexation, ce qui signifie que seuls les instruments de recherche correctement indexés apparaîtront dans les résultats de recherche.

# Alignements sur les référentiels

Un effort est consenti pour enrichir les données collectées en alignant les termes d'indexation sur des référentiels externes.

Cette opération permet de "désambiguïser" les termes en leur associant des concepts univoques, compréhensibles à la fois par les humains et les machines.

Les alignements prioritaires concernent les noms géographiques (éléments <geogname>), les personnes physiques (éléments <persname>) et les institutions (éléments <corpname>).

Les référentiels utilisés incluent GeoNames ([**https://www.geonames.org/**](https://www.geonames.org/)

), data.bnf ([**https://catalogue.bnf.fr/recherche-autorite.do?pageRech=rat**](https://catalogue.bnf.fr/recherche-autorite.do?pageRech=rat)**) ou** wikidata (<https://www.wikidata.org/wiki/Wikidata:Main_Page>).

Des mots-matières contenus dans la balise <subject> accompagnent la description du contenu des archives.

# **Typologie documentaire**

La typologie documentaire est utilisée pour identifier les différents types d'objets archivés.

La balise <genreform> est employée pour distinguer principalement deux types de documents : un type "genre" qui correspond à une typologie de concept intellectuel (par exemple, une correspondance) et un type "form" qui décrit l'aspect physique du document (par exemple, une lettre ou un journal).

Cette typologie est proposée par le SIAF pour les archives locales.

# Validation du document

Précisons également que pour valider cet instrument de recherche, nous avons utilisé le logiciel Oxygen en ayant sélectionné la DTD de l’EAD-2002, fichier nommé ead.dtd, comme fichier de validation. Le logiciel a conclu à la validation de l’ensemble du code.

# LIMITES – LIEN AVEC D’AUTRES FORMATS

Certaines institutions européennes ont déjà mis en place la description des producteurs selon la norme ISAAR-CPF au format EAD-CPF.

Par exemple, une notice pour le peintre Henri Matisse a été créée au format EAD-CPF dans le dossier d’évaluation joint.

Cependant, le lien entre cette notice et l'instrument de recherche en EAD n'a pas encore été formellement intégré dans le document EAD, principalement en raison du manque d'identifiant pour la notice sur le serveur de l'institution concernée.

Pour pallier cela, une ligne a été insérée dans la balise <controlaccess> pour référencer la notice du producteur, bien que la source "eac-cpf" de la balise «<persname> n’ait pas été entièrement validée. Le code de la notice est similaire à celui de l'instrument de recherche, avec un suffixe contenant le nom du producteur.

<persname source="eac-cpf" authfilenumber="FRAD002\_NP\_Matisse" normal="Matisse, Henri (1869-1954)" role="040">Henri Matisse</persname>