

# Documentation d'accompagnement du schéma de données

GT SIAF/AAF/Dataactivist

2021-05-17

## Contents

<b>Introduction</b>	<b>1</b>
<b>Composition du groupe de travail</b>	<b>2</b>
<b>Phasage de la création du schéma</b>	<b>2</b>
Investigation . . . . .	2
Concertation . . . . .	5
<b>- Remarques générales sur les schémas de données</b>	<b>7</b>
Quand utiliser ce schéma? . . . . .	7
Suis-je obligé d'utiliser ce schéma? . . . . .	7
Recommandations pour le formatage des fichiers . . . . .	7
Recommandations pour le nommage des fichiers . . . . .	7
Recommandations pour la mise en conformité . . . . .	7
Recommandations pour la découvrabilité . . . . .	8
<b>Les étapes de publication d'un jeu de données</b>	<b>8</b>
Identifier les données . . . . .	8
Préparer les données . . . . .	8
Documenter les données . . . . .	8
Choisir une licence? . . . . .	9
Où publier ses données? . . . . .	9
<b>- Le schéma registre des entrées</b>	<b>9</b>
Registres des entrées . . . . .	9

## Introduction

En 2018, 86 kilomètres de documents et plus de 16000Go de données sont venus enrichir les collections des services publics d'archives français (Données du service interministériel des archives de France).

Dans une optique de gestion et de suivi de leurs fonds, les services d'archives consignent chaque entrée d'archives à partir de renseignements sommaires permettant d'identifier et de retracer toutes les arrivées de nouveaux documents (provenance, date d'arrivée dans le service, volume...) : c'est le registre des entrées.

Il est généralement directement tenu au sein d'un système d'information dédié aux archives (le SIA), mais il peut également être réalisé avec un tableur voir sur un cahier.

Dans son Cadre stratégique commun de modernisation des archives, la Délégation interministérielle aux archives de France, mentionne dans les objectifs 2020-2024 :

### **Objectif 1.4 : Numériser les archives et rendre les archives numériques aisément consultables et exploitables**

Il conviendra également de généraliser l'open data dans les services d'archives de l'État et de recommander ce modèle aux services territoriaux d'archives. La création d'un schéma de données permet de standardiser la diffusion des données produites par les services d'archives. Cette standardisation est impérative pour les ré-utilisateurs car elle permet de faciliter la découverte des jeux de données et de les utiliser, notamment en simplifiant l'alignement à différentes échelles (utilisation sur des échelles locales ou nationales).

## **Composition du groupe de travail**

- Service Interministériel des Archives de France
- Association des Archivistes Français
- Dataactivist
- Archives départementales de l'Eure
- Archives départementales de l'Oise
- Archives municipales d'Avignon
- Archives nationales
- Archives régionales des Pays-de-la-Loire
- INHA, Université de Paris Nanterre
- Programme Vitam

## **Phasage de la création du schéma**

Cette partie reprend l'organisation du guide publié par Etalab. Elle vient enrichir le contenu méthodologique développé par Etalab d'un retour d'expérience complémentaire.

### **Investigation**

#### **organisation**

L'idée de création d'un schéma de données pour les entrées d'archives vient à la fois du Service Interministériel des Archives de France (SIAF) et de l'Association des Archivistes Français (AAF).

La première étape a été de constituer un groupe de travail composé de ses deux acteurs mais également d'archivistes qui produisent les registres des entrées. L'objectif a été d'avoir un groupe assez représentatif des différents services qui peuvent exister, mais également des outils utilisés par les services. Le groupe a aussi associé un expert chargé de représenter les ré-utilisateurs et usagers extérieurs aux archives.

Dans le cadre de ce projet, les commanditaires (SIAF et AAF) ont également joué le rôle de sponsors et ont réellement permis de faire avancer et de faciliter le travail.

#### **Etat des lieux**

Un état des lieux des services ayant déjà publié en open data leur registre d'entrées a ouvert la démarche. Pour cela, une recherche sur data.gouv a permis d'identifier 9 jeux de données en ligne.

Un premier constat a pu être établi :

- des services ont déjà publié ce type de données
- il n'y a pas un modèle unique de registre des entrées
- la publication reste malgré tout marginale
- la granularité et le périmètre de publication ne sont pas toujours les mêmes

Pour compléter cet échantillon, un appel à transmission d'exemples de registres d'entrées a été lancé. Afin d'être suffisamment représentatif, des services utilisant différents logiciels métiers.

Ces versions de registres sont donc, dans la plupart des cas, des extractions des outils de gestion (Système d'Information archivistique), ils reflètent la structure même des bases de données dans lesquelles les données sont produites. Le nombre de SIA sur le marché étant assez limité, nous avons pu accéder à une version du registre pour une majorité d'éditeurs.

L'adéquation entre SIA et registre d'entrées est forte, le travail à effectuer sur le fichier en sortie de SIA sera sans doute important (changement de nom de colonne, enrichissement, nettoyage de données...). une piste pour lutter contre ce problème serait que les éditeurs implémentent le schéma de données comme modèle de données pour les registres des entrées des SIA.

### **Identification des cas d'usages des données**

La compréhension des usages potentiels des jeux de données est importante car elle permet d'évaluer les besoins en données. Pour cela, un atelier sur des cas d'usages et sur les (ré)utilisateurs a été réalisé. Il posait les questions suivantes :

- *En tant que*
- *Je veux*
- *Afin de*

Cette démarche permet d'identifier le type d'utilisateurs, leurs besoins et leurs objectifs.

En tant que	Je veux	Afin de ...
Historien-ne <small>(En tant que)</small>	connaître quel service collecte quoi...	Repérer des fonds pour proposer des sujets de recherche à mes étudiants
Généalogiste professionnel <small>(En tant que)</small> notaires <small>(En tant que)</small>	Connaître les derniers versements des administrations fiscales <small>(Je veux)</small>	Régler des successions <small>(Afin de)</small>
Rédacteur/curieux <small>(En tant que)</small>	avoir un panorama de ce que collecte un service d'archives <small>(Je veux)</small> pouvoir fusionner x fichiers des entrées de x institutions sans difficultés <small>(Je veux)</small>	avoir une vue d'ensemble (enfin le + possible) des entrées à l'échelle nationale <small>(Afin de)</small> faire des stats (et des petits graphiques...) <small>(Afin de)</small> voir l'évolution des entrées de sources numériquement <small>(Afin de)</small>
Journaliste <small>(Afin de)</small>	Connaître les fonds des élus <small>(Afin de)</small>	Faire des enquêtes <small>(Afin de)</small> vérifier si des versements postérieurs ont bien eu lieu :) <small>(Afin de)</small>
Archiviste	avoir des informations sur un fonds mal identifié entre il y a longtemps (connaître sa provenance etc) <small>(Je veux)</small> En tant qu'archiviste, je veux générer un registre des entrées exploitables afin de définir mon PSCE ou ma politique de collecte <small>(Je veux)</small> bénéficier des ventes privées <small>(Je veux)</small> comparer les volumes d'entrées des services d'archives du même type que le mien <small>(Je veux)</small> En temps que archiviste, je veux connaître la collecte des autres services d'archives afin de pouvoir travailler en réseau <small>(Je veux)</small> En tant qu'archiviste, je veux renseigner et exporter facilement mon registre des entrées, conforme au format (logiciel conforme au standard) <small>(Je veux)</small>	Faire un état des versements thématiques <small>(Afin de)</small> contribuer à la transparence du service public <small>(Afin de)</small> orienter les lecteurs <small>(Afin de)</small> Planifier l'occupation de mes locaux <small>(Afin de)</small>
chercheur/étudiant <small>(En tant que)</small>	savoir avoir une vision d'ensemble des versements d'un service <small>(Je veux)</small> savoir quels sont les fonds privés qui ont été récemment entrés <small>(Je veux)</small> En tant que chercheur, je veux savoir ce qui a été collecté sur mon domaine (y compris si ce n'est pas classé) pour organiser mes recherches <small>(Je veux)</small>	pour réfléchir à un nouveau objet de recherche (article, etc.) <small>(Afin de)</small>
Agent service archives en charge des entrées	Avoir une liste de toutes les entrées sur une période donnée <small>(Je veux)</small> des informations de gestion = outil de pilotage des entrées (nli, domaine, fonctionne, producteur, statut archives, statut du process d'entrée...) <small>(Je veux)</small>	Suivre la programmation de la collecte <small>(Afin de)</small> planifier les collectes sur plusieurs années <small>(Afin de)</small> Avoir des indicateurs d'activité <small>(Afin de)</small>
Service producteur d'archives <small>(En tant que)</small>	savoir ce que mon service a déjà versé par le passé <small>(Je veux)</small> Connaître les anciens versements de mon service <small>(Je veux)</small>	Pouvoir exploiter les archives de mon service <small>(Afin de)</small>
Wikimédien <small>(En tant que)</small>	Améliorer les sources dans wiki/dats/média <small>(Je veux)</small>	Améliorer les communes grâce aux archives <small>(Afin de)</small>
Archiviste en charge de la visualisation <small>(En tant que)</small>	Ministère dans la politique CO de ma structure/organisation <small>(Je veux)</small> donner un aperçu général des fonds collectés <small>(Je veux)</small>	Améliorer l'image de mon service <small>(Afin de)</small> faire connaître les fonds de mon service (par ex. s'il n'a pas de guide ou bdd en ligne) <small>(Afin de)</small> image de mon service en interne (vis-à-vis de la DSI par ex, qui gêne parfois l'opérateur) <small>(Afin de)</small>

klaxoon

A l'issu de cet atelier, nous pouvons distinguer 2 grandes catégories d'usagers :

- **Usagers internes** : archivistes, services producteurs
- **Usagers externes** : étudiants, érudit, curieux et professionnels (journalistes, historiens, généalogistes professionnels, notaires)

Les premiers ayant des besoins plutôt orientés sur le pilotage, le suivi et la gestion et les seconds des besoins beaucoup plus vastes et difficiles à anticiper et prévoir.

De ce travail , il est possible d'identifier les intérêts de la publication des données des registres des entrées :

- apporte de la visibilité sur le travail de collecte
- permet de s'orienter dans les fonds
- offre une vision agrégée sur les collectes réalisées
- avoir des indicateurs d'activité

## Analyse des données

Une fois ce premier travail effectué, il a fallu analyser les registres collectés et transmis par les services, pour cela, les différents éléments ont été regroupés dans une base de données relationnelles en procédant au :

- recensement des sources de données, producteurs, types de producteurs, source des données
- dépouillement des fichiers de données : compilation de tous les éléments présents dans les différents fichiers et tels qu'ils existent (par ex. "cote" et "cotation", même s'ils représentent la même chose, constituent chacun une entrée). Ajout d'une description du champ, du type (même si dans la très grande majorité des cas il s'agit de champ de type "chaîne de caractères")
- regroupement par la création de familles et de sous-familles. Ces éléments apparaissent assez rapidement une fois le travail d'identification effectué. Les familles et sous-familles permettent de voir les éléments importants du schéma.

## Concertation

*Les citations proviennent de [schema.data.gouv.fr](http://schema.data.gouv.fr)*

### Profiter de l'existant

*De nombreux standards existent déjà, qu'ils concernent des formats de données ou des formats de champs. Certains standards sont devenus incontournables aujourd'hui, comme ISO-8601 pour les dates ou WGS 84 pour les coordonnées géographiques.*

Profiter de l'existant c'est aussi le travail effectué lors de la phase d'identification. Partir de ce qui existe, en tirer le commun puis ajuster/adapter.

### Identifier et associer l'écosystème

*Les personnes/organisations que vous associez sont la meilleure garantie d'un schéma de données efficace et largement adopté, permettant d'aboutir à un véritable standard. Les producteurs d'une part qui connaissent la réalité de leurs données, de la collecte, etc. et qui ont leurs propres usages. Les usagers d'autre part, leurs besoins et leurs difficultés d'autres part, qu'ils soient déjà connus, « sous le radar » ou en devenir.*

Dans le cas des registres des entrées, l'association de l'écosystème a été simple car il s'agit d'une commande par le réseau et pour le réseau. C'est sans doute plus difficile d'associer des usagers mais cela a tout de même été fait en associant un expert qui avait beaucoup commenté le registre des Archives Nationales lors de sa publication.

### Prendre le temps

*Un schéma de données est susceptible de concerner beaucoup de producteurs et d'usagers. Sa modification peut avoir un impact important. Il est donc crucial de prendre le temps d'obtenir tous les retours avant de publier un schéma utilisable par le plus grand nombre. Un schéma de données devrait être publié quand il est prêt, non pas en fonction d'un impératif de délai.*

Les pratiques de chaque service impactent assez fortement les fichiers produits. Le choix a été fait dès le départ de faire un schéma le plus exhaustif complet et avec de nombreux éléments optionnels afin de pouvoir prendre en compte la plupart des pratiques.

### Cas de l'identifiant pivot

*Utiliser des données pivot relevant d'un référentiel ouvert pour relier les données à d'autres données, par exemple l'utilisation du numéro SIREN pour identifier des organisations.*

Il est à noter que l'importance du recours à des référentiels partagés figure également dans le Cadre stratégique des Archives :

**Objectif 6.1** : Promouvoir l'utilisation d'autorités et de référentiels partagés

La recherche d'une donnée pivot a été une opération assez délicate. L'utilisation du SIREN de la collectivité à laquelle appartient le service des archives ne s'avérait pas totalement satisfaisante, la notion de service public d'archives étant gommée au profit de l'institution d'appartenance. Par ailleurs, dans certains cas, cela pouvait entraîner une confusion plus préjudiciable : cas des versements réalisés par les services déconcentrés de l'Etat et conservés par les services départementaux d'archives.

Il pouvait y avoir 2 candidats à cet identifiant pivot :

- de nombreux services diffusant leurs fonds numérisés en ligne possèdent des identifiants pérennes ARK. L'ARK se compose d'une partie, le Name Assigning Authority Number (NAAN), qui permet d'identifier une institution. Mais, même si le registre des NAAN est public, il n'est pas diffusé de façon structurée. Par ailleurs, tous les services n'ont pas un tel identifiant et cela semble complexe de le rendre obligatoire pour les services.
- les archivistes utilisent depuis de nombreuses années un standard de description des archives, l'*Encoded Archives Description* (EAD), l'EAD comporte une balise qui possède l'attribut MAINAGENCYCODE. Cela fait référence à une autre norme, l'ISAD-G, qui dans son 3.1.1, définit bien les éléments d'identification d'un fonds.
  - le code du pays conformément à la dernière version de l'ISO 3166, *Codes pour la représentation des noms de pays* (FR);
  - le code du service d'archives conformément à la norme nationale de codification des services d'archives, ou tout autre élément d'identification du lieu de conservation (AD01);
  - la cote ou tout autre élément d'identification spécifique (1W).

L'EAD précise pour sa part que le MAINAGENCYCODE doit être conforme à la norme ISO 15511, qui pour la France correspond à FR-[identifiantRCR]. Le Répertoire des centres de ressources (RCR) est lui-même structuré de la façon suivante : sur 9 chiffres reprenant : Code géog. postal, n° INSEE, numéro séquentiel. Une recherche simple et rapide permet de vite se rendre compte que ce n'est pas le cas.

- Pour les archives départementales, le code est de type : FRAD\_0COG (ex. FRAD\_001)
- Pour les archives communales, le code est de type : FRAC\_COG (ex. FRAC\_06088) ou FRAC\_CodePostal (FRAC\_95500)
- Comment gérer les cas des services d'archives métropolitaines et des services mutualisés?
- Comment gérer les cas des entrées venant des services de l'Etat?

L'une des limites est que la liste des identifiants n'est pas publiée en open data et donc ne permet pas de réellement jouer son rôle de pivot.

Il existe d'autres identifiants (VIAF, ISNI...) mais ceux-ci ne sont pas utilisés (connus) par les services archives, qui ne les utilisent pas, en l'état actuel de diffusion de ces identifiants, ils ne constituent donc pas un pivot intéressants pour faire le lien entre un service et des IR par exemple.

### Concertation en dehors du groupe de travail

Le schéma a été ajouté à [schema.data.gouv](https://schema.data.gouv.fr) comme étant en investigation le 15 novembre 2020.

Un appel à commentaires a eu lieu du 04 janvier au 12 février 2021. Un webinaire de présentation du schéma a ouvert cet appel à commentaire le 04 janvier 2021. Sur la période, une cinquantaine de retours a été réalisé.

Le groupe de travail a analysé les commentaires et à l'issue de ce travail, une V1.0 a été réalisée.

Le schéma sera produit au format json.

## - Remarques générales sur les schémas de données

### Quand utiliser ce schéma?

Ce schéma peut être utilisé lors de la publication d'un jeu de données de registre des entrées. Il a été créé par des services publics d'archives mais il peut tout à fait convenir à des services d'archives privées ou à des établissements spécialisés.

### Suis-je obligé d'utiliser ce schéma?

Non, aucune obligation car ce schéma n'a pas de caractère réglementaire. En revanche, si vous souhaitez publier des données l'utilisation de ce schéma est très fortement recommandée car il peut permettre une réutilisation plus facile de vos données car il est documenté et que vos données sont standardisées et peuvent être agrégées et comparées avec d'autres.

### Recommandations pour le formatage des fichiers

Le format de fichier retenu pour la publication des données est le CSV (Comma Separated Values, valeurs séparées par des virgules).

Les fichiers doivent, sauf exception et autant que possible, respecter les règles de formatage suivantes :

- l'encodage des caractères est UTF-8,
- le séparateur des colonnes est la virgule,
- le séparateur des nombres décimaux est le point,
- le séparateur de valeurs multiples dans un champ est le point-virgule,
- si un champ contient une virgule, il doit être entouré de guillemets doubles,
- chaque ligne doit avoir le même nombre de champs,
- le type MIME ou Content-Type est text/csv.

### Recommandations pour le nommage des fichiers

Les fichiers doivent, sauf exception et autant que possible, respecter les règles de nommage suivantes :

`AAAAMMJJ_idProducteur_registre_des_entrees_millesime.csv`

- AAAAMMJJ : Date de création du fichier
- idServArch : identifiant du service d'archives (ex. FRAD\_001)
- registre\_des\_entrees : nom du fichier, en minuscules non accentuées
- millésime : année des entrées
- extension : Si les règles de formatage sont respectées, l'extension est .csv

**Exemple :** `20201012_FRAD_001_registre_des_entrees_2020.csv`

### Recommandations pour la mise en conformité

Ces conseils reprennent ceux du Schéma des données locales publié par Open Data France

Les fichiers doivent comporter :

- Toutes les colonnes, y compris celles dont les cellules ne sont pas renseignées, dans le bon ordre, et avec des en-têtes correctement nommées sur la première ligne (nom correspondant strictement au schéma)

- Autant de lignes que nécessaire comprenant des cellules dont les valeurs peuvent être obligatoires (elles doivent être impérativement renseignées) ou optionnelles (elles sont seulement recommandées ou soumises à condition de disponibilité / pertinence)
- Traitement des cellules vides (absence de valeur ou valeur équivalente à 0) : ces cellules doivent être laissées vides. Dans le cas où une valeur numérique est égale à zéro elle doit être écrite 0.0 (zéro [point] zéro), et, dans le cas où des caractères spéciaux sont utilisés pour remplacer des valeurs manquantes (ex. “-” ou “NaN”), cela doit être mentionné dans les métadonnées.

## Recommandations pour la découvrabilité

Le respect du nommage de fichier et des métadonnées du schéma doivent permettre de découvrir l’ensemble des jeux de données répondant au schéma.

## Les étapes de publication d’un jeu de données

Pour préparer la publication de vos données, vous pouvez trouver des informations sur les fiches pratiques réalisées par OpenDataFrance

### Identifier les données

Lors de la phase de création du standard, nous avons identifié deux sources de productions des données, elles peuvent :

- provenir des systèmes d’information des services d’archives (SIA)
- provenir de fichiers tabulaires (excel, calc)

### Préparer les données

Il va falloir préparer les données afin qu’elles correspondent au modèle défini par le standard. Il peut être nécessaire de changer le nom des colonnes, de supprimer ou ajouter certaines colonnes, de modifier des valeurs etc.

### Documenter les données

Pour obtenir des **informations sur la publication** d’un jeu de données, vous pouvez consulter : l’aide de [data.gouv.fr](http://data.gouv.fr).

#### Métadonnées descriptives

- Nom du producteur du jeu de données : Le nom du producteur permet d’indiquer le service qui a produit et publié le jeu de données.

```
`Archives municipales de la commune de XXX`
`Archives départementales de XXX`
```

- Titre du jeu de données : il s’agit du nom du jeu de données et pas du nom du fichier

```
`Registre des entrées de [entité productrice], année [XXXX]`
```

- Description du jeu de données :

```
`Ce jeu de données contient les entrées du services des archives de [entité] pour l’année [XXXX].`
```

[Toutes les précisions complémentaires concernant le contenu peuvent être apportées si le jugé nécessaire]

- Fréquence de mise à jour :

```
`Annuelle` (à modifier le cas échéant)
```



- Catégories (tag) :  
`"Archives", "Entrées d'Archives", "Archives publiques"`
- Licence : [Dans le cas où une licence est associée au jeu de données, elle est à choisir dans le respect du cadre réglementaire]
- Contact :Indiquer un mail de contact pour les réutilisateurs  
`(archives@archives.fr)`

## Choisir une licence?

La licence de réutilisation n'est pas obligatoire, dans le cas où il n'y en a pas, c'est le cadre réglementaire du CRPA qui s'applique.

Si vous choisissez une licence, elle doit correspondre au cadre légal. Vous pouvez choisir soit la Licence Ouverte d'Etalab, soit la licence ODbL.

## Où publier ses données?

Suivant les procédures mises en oeuvre dans votre structure sur l'open data, vous pouvez être amenés à publier vos données soit :

- Sur le portail open data de votre structure, si celui-ci est moissonné par data.gouv, vos jeux de données remonteront automatiquement sur la plateforme nationale. Dans ce cas, vous devez vous rapprocher du chargé open data de votre structure.
- Soit directement sur data.gouv si votre structure ne possède pas de portail open data, toutes les informations nécessaire se trouvent là

## - Le schéma registre des entrées

Spécification du modèle de données relatif aux registres des entrées des services publics d'archives.

### Registres des entrées

#### Modèle de données

La version validée 1.0 du schéma de registre des entrées se trouve ici.

Un gabarit au format tableur est également prévu pour faciliter la réalisation d'un registre des entrée au format du schéma.

#### ID

- **titre** : Identifiant unique de chaque entrée
- **description** : Identifiant unique de chaque entrée composés de ces trois éléments séparés par des tirets bas : l'identifiant du service archives, année d'entrée et un identifiant unique (numéro incrémental, identifiant technique...).
- **type** : Chaîne de caractères
- **exemple** : FRAC\_13001\_2020\_001, FRAN\_2020\_aeaqaaabehgm15fab7gialyq46s4jyaaaca ‘
- valeur obligatoire
- **commentaire** : L'identifiant du service d'archives peut être composé soit du numéro utilisé pour identifier les IR, ou pour des services d'entreprises le SIREN ou pour d'autres types de services le RCR, le numéro Muséophile. Dans ce cas, il faut renseigner cet éléments dans les métadonnées associées.

#### **nomArch**

- **titre** : Nom du service archives
- **description** : Nom sous forme textuelle du service d'archives dans lequel sont entrées les archives
- **type** : Chaîne de caractères
- **exemple** : Archives municipales d'Aix-en-Provence
- valeur obligatoire

#### **coteArch**

- **titre** : Cotation
- **description** : Identifiant de l'entrée d'archives et correspondant à sa place dans le cadre de classement (série et sous-série). Dans le cas où les articles sont connus dès l'entrée, ils peuvent être ajoutés à la cotation (mais cela reste optionnel).
- **type** : Chaîne de caractères
- **exemple** : 1238W, 1240W1-12
- valeur optionnelle

#### **dateEntree**

- **titre** : Date d'entrée
- **description** : Date à laquelle les documents sont entrés dans le service d'archives au format ISO8601
- **type** : Date
- **exemple** : 2020-08-26
- valeur obligatoire

#### **statutJur**

- **titre** : Statut juridique des documents entrés
- **description** : Permet d'identifier le statut juridique des archives
- **type** : Chaîne de caractères
- **exemple** : Archives publiques
- valeur obligatoire
- **valeurs autorisées** : ["Archives publiques", "Archives privées", "Archives publiques et privées"]
- **commentaire** : "Archives publiques et privées" est à réserver pour des entrées qui comporteraient à la fois du contenu public et privé.

#### **modeEntree**

- **titre** : Modalité d'entrée
- **description** : Permet d'identifier le mode d'entrée des archives
- **type** : Chaîne de caractères
- **exemple** : Versement
- valeur obligatoire

- **valeurs autorisées** : [“Versement”, “Don”, “Achat”, “Dévolution”, “Leg ou Dation”, “Copie”, “Réintégration”, “Protocole”, “Autre”]

#### orgaVers

- **titre** : Organisation qui verse l'entrée
- **description** : Nom de l'organisation qui a versé l'entrée, distincte du service qui a produit les entrées dans l'organisation. Dans le cas où des identifiants pérennes existent pour identifier les organisations versantes, elles peuvent être ajoutées en complément du nom de l'organisation : URI - Nom de l'organisation. Ce champ peut comporter plusieurs valeurs, dans ce cas, les éléments sont séparés par une barre verticale *pipe* “|”.
- **type** : Chaîne de caractères
- **exemple** : Ville d'Aix-en-Provence, FR78422804100033\_000000011 - Tribunal administratif
- valeur optionnelle

#### servVers

- **titre** : Service qui verse l'entrée
- **description** : permet d'identifier la personne physique ou morale qui a transféré les documents
- **type** : Chaîne de caractères
- **exemple** : Service de l'achat public
- valeur optionnelle

#### orgaProd

- **titre** : Organisation productrice de l'entrée
- **description** : nom de l'organisation qui a produit l'entrée, distincte du service qui a produit les entrées dans l'organisation. Dans le cas où des identifiants pérennes existent pour identifier les organisations versantes, elles peuvent être ajoutées en complément du nom de l'organisation : URI - Nom de l'organisation. Ce champ peut comporter plusieurs valeurs, dans ce cas, les éléments sont séparés par une barre verticale *pipe* “|”.
- **type** : Chaîne de caractères
- **exemple** : Ville d'Aix-en-Provence, FR78422804100033\_000000011 - Tribunal administratif
- valeur optionnelle

#### servProd

- **titre** : Service producteur
- **description** : Permet d'identifier la personne physique ou morale qui a produit les documents
- **type** : Chaîne de caractères
- **exemple** : Service de l'achat public, Service producteur inconnu
- valeur obligatoire
- **commentaire** : Si le producteur est inconnu, saisir Service producteur inconnu

#### typeProd

- **titre** : Fonction du producteur
- **description** : permet d'identifier la fonction du producteur, liste établie à partir des éléments du SIAF
- **type** : Chaîne de caractères
- **exemple** : Commune et établissement public communal
- valeur obligatoire
- **valeurs autorisées** : ["Présidence de la République", "Premier ministre", "Ministère (administration centrale)", "Assemblée parlementaire", "Grand organe de contrôle", "Service déconcentré et établissement public de l'État à compétence départementale ou locale", "Service déconcentré et établissement public de l'État à compétence régionale ou supra-départementale", "Etablissement public national", "Commune et établissement public communal", "Conseil départemental et établissement public départemental", "Conseil régional et établissement public régional", "Structure de coopération intercommunale ou interdépartementale", "Établissement public de santé", "Organisme de droit privé chargé d'une mission de service public", "Officier public ou ministériel (dont notaire)", "Producteur privé"]
- **commentaire** : La séparation entre type et activité répond à la nouvelle organisation définie par le SIAF et mise en place au sein de l'enquête annuelle.

#### activiteProd

- **titre** : Domaine ou thématique d'action du producteur
- **description** : Permet d'identifier le domaine d'activité du producteur, liste établie à partir des éléments du SIAF
- **type** : Chaîne de caractères
- **exemple** : Administration générale (fonctions transverses, RH) | Culture, jeunesse et sports
- valeur obligatoire
- **valeurs autorisées** : ["Instance de délibération", "Direction, cabinet", "Administration générale (fonctions transverses, RH)", "Finances, fiscalité", "Économie, industrie", "Agriculture", "Équipement, environnement", "Travail, emploi", "Affaires sociales, santé", "Justice", "Police, protection civile, intérieur", "Éducation, recherche", "Culture, jeunesse et sports", "Défense, anciens combattants", "Outre-mer", "Archives privées personnelles et familiales", "Archives privées culturelles", "Archives privées d'associations, de partis politiques, de syndicats", "Archives privées d'entreprises", "Archives privées professionnelles"]
- **commentaire** : La séparation entre type et activité répond à la nouvelle organisation définie par le SIAF et mise en place au sein de l'enquête annuelle. Les valeurs "Archives privées personnelles et familiales", "Archives privées culturelles", "Archives privées d'associations, de partis politiques, de syndicats", "Archives privées d'entreprises", "Archives privées de professionnels (architectes, photographes...)" sont à réserver aux éléments qui possèdent l'élément "Producteurs privés" en *fonction du producteur*. Ce champ peut comporter plusieurs valeurs, dans ce cas, les éléments sont séparés par une barre verticale *pipe* "|".

#### descContenu

- **titre** : Description du contenu
- **description** : Permet de décrire le contenu des archives de l'entrée
- **type** : Chaîne de caractère
- **exemple** : Marchés publics de prestations intellectuelles

- valeur obligatoire

#### **dateExD**

- **titre** : Date extrême de début
- **description** : Date du plus vieux document d'archives format AAAA
- **type** : Date
- **exemple** : 2014
- valeur optionnelle

#### **dateExF**

- **titre** : Date extrême de fin
- **description** : Date du plus récent document d'archives format AAAA
- **type** : Date
- **exemple** : 2020
- valeur optionnelle

#### **natureSupport**

- **titre** : Nature du support matériel des documents
- **description** : Permet d'indiquer de façon macro la typologie de support des documents
- **type** : Chaîne de caractères
- **exemple** : Support physique
- valeur obligatoire
- **valeurs autorisées** : ["Support physique", "Support électronique", "Support mixte"]

#### **mlEntree**

- **titre** : Métrage linéaire de l'entrée
- **description** : Volume en mètre linéaire de l'entrée. Le séparateur décimal doit être le point. L'unité sera rappelée en métadonnée.
- **type** : nombre réel
- **exemple** : 1.60
- valeur optionnelle
- **commentaire** : le séparateur décimal doit être le point . L'unité sera rappelée en métadonnée.

#### **nbreArt**

- **titre** : Nombre d'articles
- **description** : Permet d'indiquer le nombre d'articles d'une entrée
- **type** : nombre entier
- **exemple** : 56
- valeur optionnelle

#### volElec

- **titre** : Volume d'archives électroniques de l'entrée
- **description** : Correspond à la volumétrie de fichiers électroniques présents dans l'entrée en Go. Le séparateur décimal doit être le point. L'unité sera rappelée en métadonnée.
- **type** : nombre réel
- **exemple** : 2.30, 0.075
- valeur optionnelle
- **commentaire** : le séparateur décimal doit être le point . L'unité sera rappelée en métadonnée.

#### objElec

- **titre** : Nombre d'objets électroniques
- **description** : Permet d'indiquer le nombre d'objets électroniques d'une entrée
- **type** : nombre entier
- **exemple** : 234