SARDE : Données issues de Légifrance

Manuel Utilisateur Public

|  |
| --- |
| Rédaction |
| Propriétaire | Département Validant | Utilisable au | Confidentialité |
| Dila | DIT : Département de l’Informatique et des télécommunications | 17/11/2014 | 🗹 Public🞎 Interne🞎 Restreint |
|  |  |  |  |

Sommaire

[1. Vue d’ensemble 2](#_Toc450892322)

[1.1 Objet et scope 2](#_Toc450892323)

[1.2 Organisation 2](#_Toc450892324)

[2. Structure technique détaillé des schémas 3](#_Toc450892325)

[2.1 Fichier Parent « SARDE » 3](#_Toc450892326)

[2.2 Fichier enfant « BLOC\_SARDE » 4](#_Toc450892327)

[Annexe 1: Transmission des Flux « SARDE » 5](#_Toc450892328)

[Annexe 2: Structure Interne à l’archive 5](#_Toc450892329)

# Vue d’ensemble

Ce document contient les éléments de documentation fonctionnelle des données XML ouverte en provenance de Légifrance et concernant les flux de données SARDE.

1.1 Objet et scope

Sarde est un thésaurus conçu pour indexer les textes parus au Journal officiel (JO). Il est composé d’un ensemble de blocs documentaire dénommé « bloc fils » ou « BLOC\_SARDE » dans la suite du document, rattaché à un et un seul « bloc père » dénommé élément « SARDE » dans la suite du document.

1.2 Organisation
Le fonds Sarde est composé de 2 types de fichiers XML, conforme à un seul schéma de référence.

Ce schéma de référence nommé Sarde\_sor\_1.0.xsd est composé des types remarquables représentés dans le diagramme ci-dessous. Les types gris et jaunes, permettent de valider le premier type de fichier, tandis que les types en vert, permettent de valider le second type de fichier.

Lecture du diagramme: Le type de donnée Sarde est composé d’une donnée nommé META de type Meta. Le type Meta est composé d’une donnée nommée META\_SPEC de type MetaSpec, et d’une donnée nommée META\_COMMUN de type MetaCommun, etc…

Le type de donnée MetaCommun est composé d’attributs informatifs élémentaires :

- ID : Identifiant unique de l’élément,

- ANCIEN\_ID : Champ vide pour SARDE, mais utilisé pour d’autres fonds comme JURI, CETA, JORF, LEGI,

- ORIGINE : prend toujours la valeur « SARDE »,

- URL : Emplacement du fichier à partir du dossier global,

- NATURE : valeurs possibles strictement défini dans le type EnumNature,

- I\_LIBELLE : contient la lettre qui indique dans quelle liste est rangée la donnée pour la recherche via,



Voir [L’écran de recherche](https://www.legifrance.gouv.fr/initRechSarde.do) dans SARDE.

Ex la valeur <I\_LIBELLE>E</I\_LIBELLE> indique que le sujet d’un XML est dans la liste qui répond à la requête <https://www.legifrance.gouv.fr/rechSarde.do?reprise=true&page=1&lettre=E>





 BLOC\_SARDE constitue le cœur informatif de SARDE. Il se raccroche au premier type de fichier à l’aide de l’identifiant parent qu’il porte en attribut (idparent).

Ainsi pour un bloc parent, identifié par son ID, on peut avoir plusieurs BLOC\_SARDE enfants référençant leur bloc père.

# Structure technique détaillé des schémas

## 2.1 Fichier Parent « SARDE »



Cette brique de schéma permet de valider les fichiers XML parent dont voici un exemple :

<?xml version="1.0" encoding="UTF-8"?>

<SARDE>

<META>

<META\_COMMUN>

<ID>SARDOBJT000007104067</ID>

<ANCIEN\_ID/>

<ORIGINE>SARDE</ORIGINE>

<URL>sarde/element/SARD/OBJT/00/00/07/10/40/SARDOBJT000007104067.xml</URL>

<NATURE>RUBRIQUE\_BASE</NATURE>

</META\_COMMUN>

<META\_SPEC>

<META\_SARDE>

<LIBELLE>MINISTERE DE LA JUSTICE : ORGANISATION</LIBELLE>

<I\_LIBELLE>M</I\_LIBELLE>

<CATEGORIE/>

<IDENTIFIANT>11110000</IDENTIFIANT>

</META\_SARDE>

</META\_SPEC>

</META>

</SARDE>

## 2.2 Fichier enfant « BLOC\_SARDE »



Cette brique de schéma permet de valider les fichiers XML enfant dont voici un exemple :

<?xml version="1.0" encoding="UTF-8"?>

<BLOC\_SARDE

idparent="SARDOBJT000007104067"

libelleparent="MINISTERE DE LA JUSTICE : ORGANISATION">

<ID>SARDBLOC000007125827</ID>

<TITRE>Arrêté du 7 juillet 2006 relatif à la création, à la composition et au fonctionnement de la commission d'appel d'offres pour le marché relatif à la mission d'assistance à la maîtrise d'ouvrage (AM ...</TITRE>

<NATURE\_TEXTE>Arrêté</NATURE\_TEXTE>

<ORDRE>708669588</ORDRE>

<CONTENU\_SARDE>

<CONTENU>

<p>Arrêté du 7 juillet 2006 relatif à la création, à la composition et au fonctionnement de la commission d'appel d'offres pour le marché relatif à la mission d'assistance à la maîtrise d'ouvrage (AMOA), nécessité par la conception et le déploiement de l'infocentre de contrôle de gestion des services judiciaires, passé par la direction des services judiciaires du ministère de la justice</p>

<blockquote>

<font color="#ff0000">NOR <a alt="JUSB0610462A" href="/affichTexte.do?cidTexte=JORFTEXT000000818378&amp;categorieLien=cid">JUSB0610462A</a>

</font>

<br/>

<font color="#ff0000">J.O. du 26/07/2006 texte : <a alt="JO - 2006-07-26 Texte n°- page 0" href="/jo\_pdf.do?cidTexte=JPDF2607200600000026&amp;categorieLien=id"> n° 26</a>(page 11137)</font>

</blockquote>

</CONTENU>

</CONTENU\_SARDE>

<LIENS\_SARDE/>

</BLOC\_SARDE>

# Annexe 1: Transmission des Flux « SARDE »

Le site FTP accueille un ensemble d’archives compressées au format tar.gz, nommées selon le pattern suivant :

NomBase\_<yyyymmdd>-<hhmmss>.tar.gz

Ce timestamp correspond au moment de l’extraction et permet de connaître l’ordre dans lequel les flux ont été mis à disposition.

Les archives sont conservées 62 jours sur l'espace FTP.

# Annexe 2: Structure Interne à l’archive

Un élément enfant < BLOC\_SARDE > dont l’ID est SARDBLOC000030469109 sera stocké dans :

global/sarde/bloc/SARD/BLOC/00/00/30/46/91/SARDBLOC000030469109.xml

Un élément père <Sarde> dont l’ID est SARDOBJT000030435989 sera stocké dans :

global/sarde/element/SARD/OBJT/00/00/30/43/59/SARDOBJT000030435989.xml

Les répertoires sont générés automatiquement en fonction des identifiants fournis par la base de données.

Les séquences par 2 numéros, permettent d’avoir des répertoires n’excédant pas une taille de 100 fichiers, pour des considérations d’optimisation technique.

Structure d’un ID

|  |  |  |  |
| --- | --- | --- | --- |
| Niveau | Pattern | Description | Exemple |
| 1 | FFFF | Identifiant du Fonds sur 4 caractères | SARD |
| 2 | TTTT | Type de document sur 4 caractères | BLOC ou OBJT |
| 3 | xx | Séquence Base | 00 |
| 4 | xx | Séquence Base | 00 |
| 5 | xx | Séquence Base | 07 |
| 6 | xx | Séquence Base | 43 |
| 7 | xx | Séquence Base | 47 |
| 8 |  | Fichier XML | SARDOBJT000007434712.xml |

Fin du document