



**MINISTÈRE  
DE L'ÉDUCATION  
NATIONALE,  
DE L'ENSEIGNEMENT  
SUPÉRIEUR  
ET DE LA RECHERCHE**

*Liberté  
Égalité  
Fraternité*

**Secrétariat général**

Direction du numérique  
pour l'éducation  
Sous-direction des services  
numériques  
Bureau des services et outils  
numériques pour l'éducation  
(DNE SN1)

99, rue de Grenelle  
75357 Paris SP 07

Secrétariat général  
Service de l'action  
administrative et des  
moyens  
Sous-direction des achats  
(SAAM B)  
Bureau de la stratégie  
et de l'ingénierie des achats  
(SAAM B1)

61-65, rue Dutot  
75732 Paris Cedex 15

# **CAHIER DES CLAUSES TECHNIQUES PARTICULIÈRES**

**ANNEXE 04.9:** Spécifications Moissonneur OAI-PMH

**Procédure :** MEN-SG-AOO-24002

**Objet :** Prestations de prise en charge de la solution du gestionnaire d'accès aux ressources (GAR), d'hébergement, d'exploitation, de maintenance, de support et de développement de ladite solution pour le compte du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.




# RENATER - GAR - Moissonnage

## DSFD

**Le 22/12/2023**

Référence :	Version	En date du
GAR-S2.DSFD.0007.Spécifications_du_moissonneur	15.00	22/12/2023

Version	Rédigé par	Objet	Vérfié		Validé	
			Par	Le	Par	Le
15.00	PASS-TECH	Mantis : 1395 - Lot7.1_PFV_post Moissonnage - Licence orpheline	Worldline	22/12/2023	RENATER  Re_ _EB 474_ Spécifications foncti	22/12/2023

## Évolutions successives

Version	Date	Description	Auteur(s)
1	13/09/2016	Création	Worldline
1.1	23/09/2016	Mise à jour: suite à l'atelier du 22/09 suite à la réception du document « SpecifDoc_V13_15092016.docx » Ajout précisions concernant le traitement des notices rejetées	Worldline
1.2	03/10/2016	Mise à jour suite à la réception du document « Moissonnage-identifiants-V2_29092016.docx » intégration du retour DNE du 29/09/2016	Worldline
1.3	05/10/2016	Mise à jour suite à la réception du retour DNE du 04/10/2016, à l'atelier validation du 05/10/2016 et à réception du document AttributsAvec_code_vM2m2c0_04102016.xlsx	Worldline
1.4	07/10/2016	Mise à jour suite à la réception du retour DNE du 06/10/2016	Worldline
1.5	25/11/2016	Francisation des intitulés de tables et de colonnes. Mise à jour suite à la première livraison du module de moissonnage, modifications mineures des modèle de notification et prise en compte du traitement des vignettes.	Worldline
1.6	15/12/2016	prise en compte des retours concernant le traitement des vignettes.	Worldline
1.7	25/01/2017	Prise en compte des réserves concernant le traitement des vignettes Évolutions du moissonnage : Désactivation du test de la date de validation GAR des notices lors du moissonnage Validation des éditeurs de notices lors du moissonnage Gestion d'un lock pour le lancement manuel du moissonnage	Worldline
1.8	02/02/2017	Précision concernant la Validation des éditeurs de notices lors du moissonnage	Worldline
1.9	21/02/2017	Ajout de précision demandées par Renater	Worldline
2.0	04/04/2017	[Defect 15] Ajout précision sur la gestion des erreurs pour les éléments de notice supérieurs à la taille des champs BDD (cf §3.2)	Worldline
2.1	07/06/2017	[Defect 143] Ajout de précisions sur l'identification des éléments de classification.	Worldline
2.2	06/07/2017	Prise en compte de la demande d'évolution « Implémentation de ScoLOMfr 3.2 dans le GAR (TUMTECSGEV-27) »	Worldline
2.3	04/08/2017	Ajout de précisions.	Worldline
2.4	15/09/2017	Mise à jour dans le cadre de l'évolution Id complémentaire : R.DDC.0016.1.1.SV_Id_complementaire	Worldline

Version	Date	Description	Auteur(s)
2.5	21/09/2017	Prise en compte des retours dans le cadre de l'évolution Id complémentaire : R.DDC.0016.1.1.SV_Id_complementaire	Worldline
2.6	17/11/2017	Mise à jour dans le cadre de l'évolution « Modification de contrôles sur une notice » : R.DDC.0021.1.0.SV-Modification_controles_notice	Worldline
2.7	28/11/2017	Prise en compte des retours Renater dans le cadre de l'évolution « Modification de contrôles sur une notice » : R.DDC.0021.1.0.SV-Modification_controles_notice	Worldline
2.8	27/02/2018	Defect 562 – rejet à tort des notices sur vignette incorrecte (§3.2.17)	Worldline
2.9	30/04/2018	Précision concernant les adresses de notification du DTR Précision concernant les séparateurs possibles pour la liste d'attributs GAR	Worldline
2.10	22/06/2018	évolution 58 Contrôle unicité URL des ressources dans la BDD	Worldline
2.11	12/07/2018	Prise en compte des retours concernant l'évolution 58 Contrôle unicité URL des ressources dans la BDD	Worldline
3.0	21/11/2018	Montée de version de ScoLOMFR3.2 vers ScoLOMFR4.0	Worldline
3.1	06/12/2018		
3.2	28/01/2019	Évolution 89 : Ajout de l'identifiant de plateforme dans la VCard DTR	Worldline
3.3	20/02/2019		
3.4	05/03/2019		
3.5	05/06/2019	Evolution 99 : Affichage picto mère de famille	Worldline
3.6	10/07/2019		
3.7	15/07/2019	Montée de version de ScoLOMFR4.0 vers ScoLOMFR5.0	Worldline
3.8	15/07/2020	Montée de version de ScoLOMFR5.0 vers ScoLOMFR6.0	Worldline
3.9	26/07/2021	Montée de version de ScoLOMFR6.0 vers ScoLOMFR7.0	Worldline
3.10			
4.0			
04.01	13/07/2022	Evolution 375 (WL : 177) : Montée de version de ScoLOMFR7.0 vers ScoLOMFR8.0	Worldline
04.02	12/08/2022	Prise en compte des retours sur l'évolution 375 (WL : 177) : Montée de version de ScoLOMFR7.0 vers ScoLOMFR8.0	Worldline
05.00	16/08/2022	Version validée dans le cadre de l'évolution 375 (WL : 177) : Montée de version de ScoLOMFR7.0 vers ScoLOMFR8.0	Worldline
05.01	15/07/2022	Evolution 239 (WL 167) : Gestion des applications natives	Worldline
05.02	10/08/2022	Evolution 239 (WL 167) : Gestion des applications natives et initialisation du formatage du document	Worldline
05.03	22/08/2022	Evolution 239 (WL 167) : Corrections globales et ajout de la prise en compte de la suppression des extendedLocation	Worldline
05.04	02/09/2022	Evolution 239 (WL 167) : Corrections globales et réorganisation partielle du sous-chapitre extendedLocation	Worldline
06.00	13/09/2022	Evolution 239 (WL 167) : version validée	Worldline
06.01	26/09/2022	Evolution 394 (WL 182) : Lien ressource backend ressource appelante	Worldline

Version	Date	Description	Auteur(s)
06.02	04/10/2022	Evolution 394 (WL 182) : Lien ressource backend ressource appelante	Worldline
07.00	05/10/2022	Evolution 394 (WL 182) : Version validée par Renater + report Scolomfr 8.0 (Evolution 375 (WL 177))	Worldline
07.01	08/11/2022	Corrections suite développement	Worldline
07.02	29/11/2022	Prise en compte des remarques de Renater	Worldline
07.03	06/12/2022	Prise en compte des remarques de Renater	Worldline
08.00	07/12/2022	Version validée par Renater	Worldline
08.01	14/12/2022	Evolution 392 (WL 161) : Suppression des ressources orphelines dans le GAR	Worldline
08.02	13/01/2023	Prise en compte des remarques de Renater	Worldline
08.03	20/01/2023	Prise en compte des remarques de Renater	Worldline
09.00	25/01/2023	Version validée dans le cadre de la release 7.1	Worldline
09.01	14/03/2023	Report Release 7.0 EB 412 : Ressource Backend en Ressource Technique Commune ALM 1803 – Retrait du type de présentation RTC Ajout de la Règle 44 sur les attributs obligatoires UAI et IDO	Worldline
10.00	20/03/2023	Version validée par Renater	Worldline
10.01	09/05/2023	EB 413 : Echapper les caractères spéciaux dans les notices	Worldline
10.02	17/05/2023	Prise en compte des retours de RENATER	Worldline
10.03	26/05/2023	Prise en compte des retours de RENATER	Worldline
11.00	30/05/2023	Version validée de l'évolution 413 dans la release 7.2	Worldline
11.01	31/05/2023	EB 437 : Contrôle d'unicité des informations OIDC	Worldline
12.00	16/06/2023	Version validée de l'évolution 437 dans la release 7.2	Worldline
12.01	13/07/2023	EB 452 : Migration ScoLOMFR de la version 8.0 à la version 9.0	Worldline
12.02	09/08/2023	EB 452 : Prise en compte des retours de RENATER	Worldline
13.00	11/08/2023	Version validée de l'évolution 452 par RENATER	Worldline
13.01	13/10/2023	EB 462 : Modification des DCP pour les RTC	Worldline
13.02	31/10/2023	Prise en compte des retours de RENATER	Worldline
13.03	07/11/2023	Prise en compte des retours RENATER	Worldline
14.00	07/11/2023	Version validée de l'évolution 462 par RENATER	Worldline
14.01	12/12/2023	Mantis : 1395 - Lot7.1_PFV_post Moissonnage - Licence orpheline	Worldline

Version	Date	Description	Auteur(s)
15.00	22/12/2023	Version validée	Worldline

## Table des matières

<b>Table des matières .....</b>	<b>6</b>
<b>1 Introduction.....</b>	<b>9</b>
1.1 Objet du document .....	9
1.2 Responsabilités liées au document.....	9
1.3 Documents de référence .....	9
1.4 Autres documents utilisés.....	9
1.5 Abréviations .....	9
1.6 Définitions .....	10
<b>2 Notes générales .....</b>	<b>11</b>
2.1 Présentation générale et contexte .....	11
2.1.1 Module de moissonnage.....	11
2.1.2 Traitement des vignettes .....	13
2.2 Objectif et engagement .....	13
2.3 Configuration du module de moissonnage.....	14
2.3.1 Normalisation de certains caractères par d'autres caractères.....	14
2.4 Versions de ScoLOMFR supportées .....	15
2.5 Planification des moissons.....	15
2.6 Moisson d'un entrepôt.....	16
2.6.1 Moissonnage OAI .....	16
2.6.2 Indexation ScoLOMFR .....	17
2.6.3 Traitements de vérification.....	17
2.6.4 Stockage et mise à jour en base de données.....	17
2.7 Envoi de notifications.....	18
2.8 Différentes variantes d'implémentations des ressources .....	18
2.9 Notion de familles de ressources.....	23
2.10 Organisation du système de fichiers et accès aux vignettes .....	24
2.10.1 Organisation du système de fichiers.....	24
2.10.2 Accès aux vignettes GAR et MÈRE DE FAMILLE .....	24
<b>3 Notes techniques .....</b>	<b>25</b>
3.1 Appel du module de moissonnage.....	25
3.1.1 Configuration du module .....	25
3.1.2 Traitement des notices orphelines .....	27
3.2 Validation et indexation des notices.....	28
3.2.1 Identifiant.....	28
3.2.2 Titre .....	29
3.2.3 URL de la vignette .....	30
3.2.4 Label GAR .....	30
3.2.5 Date de validation GAR .....	32
3.2.6 VCard de l'éditeur .....	33

3.2.7	VCard du distributeur technique .....	34
3.2.8	VCards des distributeurs commerciaux .....	36
3.2.9	VCard du valideur technique .....	37
3.2.10	Informations des éléments extendedLocation .....	38
3.2.11	Type pédagogique de la ressource .....	49
3.2.12	Typologie du document .....	50
3.2.13	Niveau éducatif détaillé .....	50
3.2.14	Domaine d'enseignement.....	51
3.2.15	Type de présentation .....	52
3.2.16	Description de la notice .....	52
3.2.17	Remarques générales.....	53
3.3	Stockage des données.....	54
3.4	Cloisonnement des données.....	55
3.5	Notifications.....	56
3.5.1	Bilan de la moisson .....	56
3.5.2	Erreurs et avertissements rencontrés dans les notices .....	57
3.6	Données loguées.....	58
3.7	Informations potentielles de niveau erreur ou avertissement.....	59
3.7.1	Table des erreurs .....	59
3.7.2	Table des erreurs et des avertissements concernant les extendedLocation.....	60
<b>4</b>	<b>Annexe : modèle de données .....</b>	<b>63</b>
4.1	Configuration des entrepôts à moissonner .....	63
4.2	Données issues de la moisson.....	64
4.2.1	Table « notice ».....	64
4.2.2	Table « notice_oidc » .....	69
4.2.3	Table « notice_descripteur » .....	70
4.2.4	Table « notice_vcard » .....	71
4.2.5	Table « notice_liaison » .....	72
<b>5</b>	<b>Annexe : GAR – Type de présentation.....</b>	<b>73</b>
<b>6</b>	<b>Annexe : GAR – Liste des attributs valides .....</b>	<b>74</b>
<b>7</b>	<b>Annexe : GAR – Liste des types d'identifiants autorisés.....</b>	<b>76</b>
<b>8</b>	<b>Annexe : GAR – Codes SIREN des éditeurs ayant adhéré au cadre de confiance.....</b>	<b>77</b>
<b>9</b>	<b>Annexe : Script de standardisation des vignettes.....</b>	<b>78</b>
9.1	Préambule .....	78
9.2	Traitement à apporter à chaque image .....	78
9.2.1	Création du fond de la vignette pour usage par le GAR .....	78
9.2.2	Mise à l'échelle de la vignette éditeur .....	79
9.2.3	Intégration de la vignette éditeur retaillée.....	80
9.2.4	Intégration de la pastille GAR et de la pastille MÈRE DE FAMILLE.....	81
9.2.5	Génération du fichier PNG.....	82
9.3	Gestion des erreurs.....	82
9.4	Lancement du script.....	83
9.5	Prérequis techniques .....	83



<b>10</b>	<b>Annexe : Résultat de l'étude du passage de ScoLOMFR 8.0 à ScoLOMFR 9.0.....</b>	<b>84</b>
10.1	Résultat de l'étude .....	84
10.2	Présentation de la solution .....	85
10.2.1	Vérification de la compatibilité ascendante des schémas 9.0.....	85
10.2.2	Intégration du format ScoLOMFR 9.0 .....	86
10.2.3	Mise à jour des contrôles effectués sur les notices .....	86

# 1 Introduction

## 1.1 Objet du document

Ce document est le Document de Spécifications Fonctionnelles pour le module de moissonnage dans le cadre du projet Gestionnaire d'Accès aux Ressources (GAR).

Il a pour objectif de définir le comportement attendu du module de moissonnage ainsi que son interaction avec son environnement.

## 1.2 Responsabilités liées au document

Le chef de projet Worldline est responsable de la rédaction du Dossier de Spécifications fonctionnelles, RENATER et le Ministère en charge de l'Éducation Nationale (MEN) sont responsables de sa validation.

## 1.3 Documents de référence

Numéro	Réf. Document	Type
DR1	DSFD Architecture Base de données	DSFD
DR2	R.NOE.0003.02.SV Matrice de notifications du GAR.xlsx	NOE
DR3	GAR-S2.PAR.0001.ParamétrageEnvironnements.xlsx	PAR

## 1.4 Autres documents utilisés

Numéro	Réf. Document	Type
ADU1	D07-2-Annexe 2 du Marché Subsequent n°2 – Cahier des charges du GAR.pdf	
ADU2	GAR_DT_ScoLOM_Interface_Donnees_V1.0.pdf	
ADU3	Dossier ScoLOM-GAR\SpecifDoc_GAR-ScoLOM_V1.0.pdf	
ADU4	GAR_Gestion_attributsGAR_DCP_V1.0.pdf	
ADU5	SpecifDoc_V14_30092016.docx	
ADU6	Moissonnage-identifiants-V2_29092016.docx	
ADU7	AttributsAvec_code_vM2m2c0_04102016.xlsx	
ADU8	scolomfr_familles_M1m2c0.docx	
ADU9	ScoLOMFR-manuel-reference-v9-0.pdf	

## 1.5 Abréviations

Abréviation	Signification
DSF	Document de Spécifications Fonctionnelles
GAR	Gestionnaire d'Accès aux Ressources
MEN	Ministère en charge de L'Éducation Nationale

## 1.6 Définitions

Terme	Définition
Vignette éditeur	Vignette associée à une ressource et fournie par l'éditeur. Elles sont téléchargées lors de la moisson.
Vignette normalisée (ou standardisée)	Vignette issue de la transformation d'une vignette éditeur pour un usage dans le GAR. Elle est une composition de la vignette éditeur, éventuellement retaillée, d'un fond et d'une ou plusieurs pastilles en fonction du type de ressource. L'ensemble des vignettes éditeurs est concerné par le processus de normalisation. Ce processus est indépendant de la moisson.
Vignette GAR	Vignette normalisée avec apposition d'une pastille GAR. L'ensemble des vignettes normalisées a une pastille GAR.
Vignette MÈRE DE FAMILLE	Pour les notices « mère de famille », c'est une vignette normalisée avec apposition d'une pastille GAR et d'une pastille MÈRE DE FAMILLE. C'est de fait une vignette GAR avec une pastille MÈRE DE FAMILLE.
Notice orpheline	Notice existante dans la base de données du GAR et plus dans l'entrepôt OAI correspondant.
Moissonnage OAI	Processus de collecte des enregistrements d'un entrepôt de notices respectant le protocole OAI-PMH : <a href="https://www.openarchives.org/OAI/openarchivesprotocol.html">https://www.openarchives.org/OAI/openarchivesprotocol.html</a>
Entrepôt OAI	Entrepôt de notices respectant le protocole OAI-PMH et ainsi interrogeable par un moissonneur OAI. Il dispose de métadonnées globales et d'un ou plusieurs ensembles d'enregistrements OAI.
Enregistrement OAI	Un enregistrement OAI ( <i>Record</i> ) représente le conteneur d'une notice respectant un format précis au sein d'un entrepôt OAI et dispose de métadonnées. Un enregistrement OAI dont la notice est supprimée présente la métadonnée suivante dans l'entête : <record><header status="deleted"> ... </header></record>

Glossaire projet : [GlossaireGAR.V06.00](#)

## 2 Notes générales

Ce chapitre présente la description fonctionnelle des services attendus sous forme de sous-chapitres :

Préambule :

1. **Présentation générale et contexte** : couvre la description générale du module de moissonnage et du module de traitement des vignettes

Module moissonneur :

2. **Objectif et engagement** : couvre la déclaration de l'objectif principal du moissonneur et l'engagement associé
3. **Configuration du module de moissonnage** : définit les informations nécessaires pour la déclaration des entrepôts à moissonner
4. **Versions de ScoLOMFR supportées** : liste les versions de ScoLOMFR supportées par le moissonneur
5. **Planification des moissons** : décrit le fonctionnement attendu pour le lancement automatique du moissonneur
6. **Moisson d'un entrepôt** : décrit les valeurs extraites des notices moissonnées d'un entrepôt OAI
7. **Envoi de notifications** : décrit le processus d'envoi d'email d'alertes et d'erreurs à la fin de la moisson d'un entrepôt
8. **Différentes variantes d'implémentations des ressources** : décrit les différents moyens d'accès aux ressources utilisateur web avec ou sans variantes natives ; ou aux ressources techniques communes (RTC)
9. **Notion de familles de ressources** : décrit la possibilité de lier des ressources utilisateur entre elles par des ressources « mère de famille » et des ressources « membre de famille »

Module traitement des vignettes :

10. **Organisation du système de fichiers et accès aux vignettes** : décrit cette organisation et le moyen d'accès aux vignettes spécifiques GAR et mère de famille

### 2.1 Présentation générale et contexte

#### 2.1.1 Module de moissonnage

Le module de moissonnage a pour rôle de collecter les notices ScoLOMFR de description de ressources auprès des fournisseurs de ressources affiliés au GAR. Il maintient une base de référence des notices moissonnées.

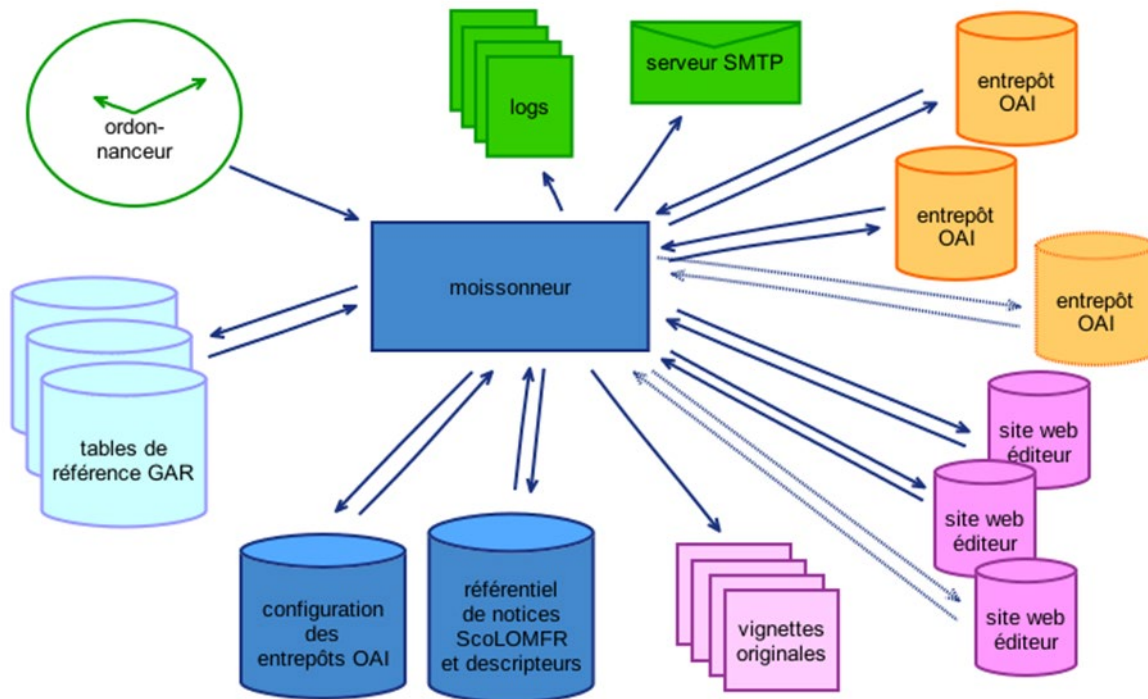


Diagramme global du module de moissonnage

Le module de moissonnage interroge les entrepôts OAI (en orange à droite dans le diagramme) mis à disposition par les fournisseurs de ressources. La liste des entrepôts à moissonner ainsi que leurs caractéristiques respectives (dont les coordonnées du gestionnaire d'entrepôt) sont décrites dans une base de configuration des entrepôts OAI (en bleu foncé en bas du diagramme).

Le déclenchement automatique des moissons est assuré par l'ordonnanceur (en haut à gauche du diagramme). La planification des moissons (fréquence, heure dans la journée) tient compte des souhaits du responsable de l'entrepôt.

Durant la moisson d'un entrepôt, le module de moissonnage consigne les différentes informations et erreurs de traitement dans un fichier de log (en vert en haut du diagramme). À la fin de la moisson, un récapitulatif est envoyé par courriel en utilisant le serveur SMTP du GAR (en vert en haut du diagramme). Le récapitulatif est envoyé au DTR indiqué dans les notices ScoLOMFR moissonnées et/ou, au gestionnaire de l'entrepôt. Le choix de l'envoi à l'ensemble des destinataires est paramétrable.

Durant la moisson, le module de moissonnage stocke les ressources moissonnées non rejetées dans une base de notices (base « référentiel de notices ScoLOMFR et descripteurs » en bleu foncé en bas du diagramme) en extrayant un certain nombre d'indicateurs utiles aux traitements suivant des ressources. Le module de moissonnage assure que toutes les notices stockées respectent les différents critères de validité (Voir 3.2 - Validation et indexation des notices).

Les tables de référence GAR (en bleu clair à gauche du diagramme) sont consultées pour la gestion des attributs GAR (liste contrôlée et correspondance avec les catégories et type de DCP), les types de présentation et les types d'identifiants ScoLOMFR acceptés avec leur priorité.

Lors de la moisson et l'analyse des notices ScoLOMFR, le module de moissonnage extrait les URLs des vignettes des notices, interroge les sites web des éditeurs (en rose à droite sur le diagramme) et stocke les images originales sur le disque (vignettes originales, aussi appelées « vignettes éditeur » dans le document).

Note : le CCTP évoquait la possibilité d'utiliser différents protocoles de moisson mais en ne citant que OAI-PMH. En l'absence d'information complémentaire, nous avons retenu OAI-PMH comme unique protocole de moisson, l'ajout ultérieur d'un autre protocole fera l'objet d'une spécification et d'une évolution ultérieure.

## 2.1.2 Traitement des vignettes

Afin d'illustrer les notices présentes dans le GAR, le module de moissonnage est chargé de récupérer les vignettes éditeurs liées aux notices ScoLOMFR durant la moisson afin de les traiter et de les mettre à disposition dans un format normalisé pour l'ensemble des modules du GAR qui le souhaitent.

De fait, le traitement des vignettes est réparti en deux phases :

- Téléchargement des vignettes éditeur associées aux notices durant la moisson d'un entrepôt
- Normalisation des vignettes pour leur utilisation par le GAR en 2 phases :
  - génération des vignettes GAR pour l'ensemble des vignettes téléchargées.
  - puis génération des vignettes MÈRE DE FAMILLE pour les notices « mère de famille »

Les images (originales et standardisées) sont stockées sur le disque et le chemin d'accès de chaque vignette GAR est indiqué dans la table des notices en base de données.

### 2.1.2.1 Téléchargement des vignettes éditeur

À la fin de la moisson d'un entrepôt, l'ensemble des vignettes éditeur associées aux notices de cet entrepôt sont téléchargées dans un dossier dédié. Le script de normalisation des vignettes traite le contenu de ce dossier pour créer les vignettes GAR et les vignettes MÈRE DE FAMILLE correspondantes.

### 2.1.2.2 Standardisation des vignettes

Le script de standardisation des vignettes est indépendant du module de moissonnage. Il est lancé après chaque moisson pour traiter les images éditeurs téléchargées par le module de moissonnage.

Le script de standardisation accède cependant au fichier de configuration du module de moissonnage afin de déterminer le chemin des vignettes éditeur à traiter et les différents paramètres de traitement à prendre en compte. Son fonctionnement est détaillé en annexe de ce document.

## 2.2 Objectif et engagement

Le module de moissonnage est chargé d'interroger les entrepôts OAI déclarés par les fournisseurs de ressources. Il maintient à jour une base de ressources moissonnées et extrait de chaque notice un ensemble de descripteurs à destination des autres modules du GAR.

Tout problème rencontré durant la moisson et l'analyse des notices est remonté au distributeur technique de la notice et/ou au gestionnaire de l'entrepôt et consigné dans les logs. Une notice rejetée du fait de l'absence de données ou d'un problème de validité n'est pas stockée et n'entraîne pas de mise à jour des données existantes.

De ce fait, le module de moissonnage assure que toutes les notices sauvegardées en base de données respectent les critères de validité du GAR et peuvent être utilisées par les autres modules du GAR.

Une exception toutefois : Il est possible de désactiver, par configuration, le test de date de validité GAR sur une plate-forme donnée. Dans ce cas, une notice qui ne comportant pas de date de validité GAR n'est pas rejetée pour ce motif. Ce paramètre particulier permet de tester (hors production) des notices d'éditeur n'ayant pas encore obtenu le label GAR.

## 2.3 Configuration du module de moissonnage

L'administrateur du GAR déclare les entrepôts à moissonner dans la base locale de configuration. Les informations nécessaires pour que le module de moissonnage puisse traiter chaque entrepôt sont :

- Identifiant de l'entrepôt, attribué automatiquement par le système ;
- URL d'accès à l'entrepôt ;
- Format des fiches : valeur à utiliser pour le paramètre « metadataPrefix » de requête de l'entrepôt OAI. Quelle que soit la valeur utilisée, il est entendu que les notices doivent être conformes aux schémas ScoLOMFR supportés ;
- Type de moisson à mettre en œuvre : totale ou partielle
- Type de granularité : date ou date et heure
- Type de gestion des suppressions de notices (champ « deletedRecord ») :
  - Non géré (no) : Les notices supprimées disparaissent de l'entrepôt OAI sans notification
  - Temporaire (transient) : Les notices sont marquées comme supprimées pendant une durée limitée avant de disparaître
  - Permanent (persistant) : Les informations de suppression sont conservées sans limitation de durée
- Nom du set à utiliser pour la requête (optionnel)
- Coordonnées du gestionnaire de l'entrepôt (courriel, nom, fonction, description)
- Préférences en termes de planification de la moisson
- Liste de correspondance des champs et caractères à normaliser (la liste est contenue dans un fichier qui est référencé sur le document DR3)

Notes :

- Certaines informations peuvent être obtenues par appel de l'entrepôt : Le verbe « Identify » permet d'interroger un entrepôt pour récupérer certaines informations techniques, notamment la façon dont sont gérées les notices supprimées.
- Les informations de périodicité de moisson sont présentes à titre d'information. L'ajout d'une tâche planifiée au sein de l'ordonnanceur est décrit dans la section « Planification des moissons »

La configuration du module de moissonnage est complétée par plusieurs directives nécessaires au traitement des images (chemin de stockage des vignettes, paramètres de configuration de la composition des vignettes GAR et vignettes MÈRE DE FAMILLE).

### 2.3.1 Normalisation de certains caractères par d'autres caractères

Afin d'éviter les problèmes liés à des caractères particuliers (tabulation, saut de ligne, etc.) ou tout autre caractère spécifique dans un des champs des notices, une normalisation de certains caractères par d'autres caractères est effectuée sur la chaîne de caractères indexée du champ concerné. Par exemple :

- Dans le champ Titre :
  - Remplacement des CR, TAB, LF par des espaces simples ;

- Suppression des espaces en début et en fin de chaîne de caractères ;
- Remplacement de séquences d'espaces successifs par des espaces simples.

Ce remplacement s'appuie sur les caractères Unicode concernés et listés dans le fichier de configuration qui suivra l'exemple ci-dessous.

La structure du fichier de configuration prend en charge la combinaison pour chaque champ à modifier et la liste des caractères à normaliser.

Voici un exemple de paramétrage :

champ	caractère à remplacer	Unicode à remplacer	Caractère de remplacement	Unicode remplacement
notice_titre	TAB (tabulation)	0009	Espace	0020
notice_titre	LF (retour à la ligne)	000A	Espace	0020
notice_titre	CR (retour chariot)	000D	Espace	0020
notice_titre	LS (espace(s) de début)	0020	Rien	
notice_titre	TS (espace(s) de fin)	0020	Rien	
notice_titre	MULTI (séquence d'espaces)	0020	Espace	0020
notice_description	TAB	0009	Espace	0020

Tout autre champ à part notice\_titre et notice\_description sont d'ores et déjà normalisés et de fait sont exclus.

## 2.4 Versions de ScoLOMFR supportées

Le module de moissonnage est compatible avec les versions suivantes de ScoLOMFR :

- ScoLOMFR V3.0
- ScoLOMFR V3.2
- ScoLOMFR V4.0
- ScoLOMFR V5.0
- ScoLOMFR V6.0
- ScoLOMFR V7.0
- ScoLOMFR V8.0
- ScoLOMFR V9.0

La compatibilité ascendante en terme de validation XSD est assurée entre les versions 3.0, 3.1, 3.2, 4.0, 5.0, 6.0, 7.0, 8.0 et 9.0 sauf dans le cas peu probable où une notice contient un élément de classification avec « enseignement » en valeur de <lom:purpose>. Dans ce cas la notice sera rejetée et le message suivant : « Notice invalide au regard du schéma XSD ScoLOMFR [version ScoLOMFR actuelle] : [Nom de la notice] » sera ajouté à la notification.

## 2.5 Planification des moissons

Le processus de moisson est lancé de manière automatique par l'ordonnanceur pour chacun des entrepôts déclarés pour le GAR. Lors du référencement d'un nouvel entrepôt à moissonner, le gestionnaire de l'entrepôt fournit au titulaire du marché l'ensemble des informations nécessaires à la moisson de son entrepôt.



La planification des moissons est effectuée, autant que possible, en concertation avec le gestionnaire de l'entrepôt.

## 2.6 Moisson d'un entrepôt

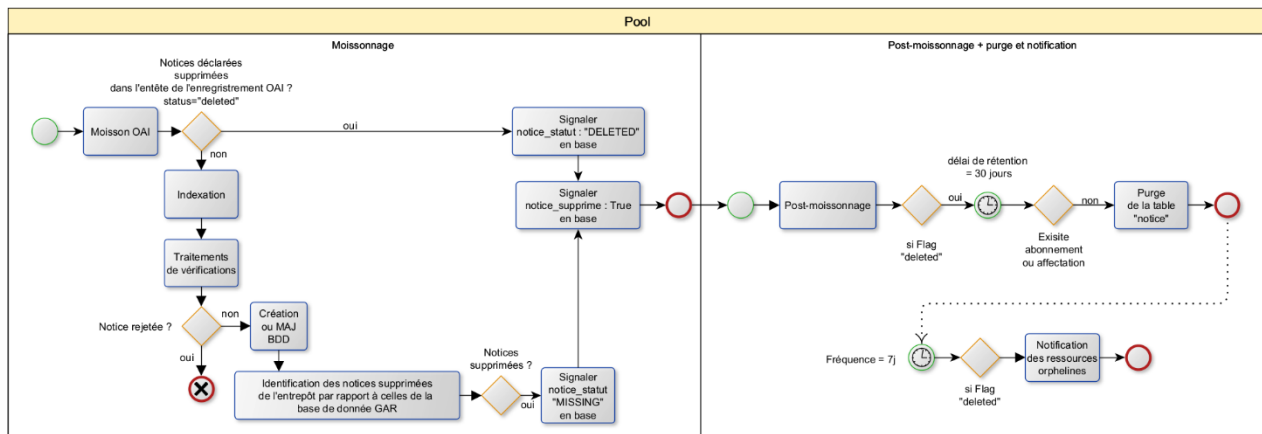


Figure 1 - Schéma général

### 2.6.1 Moissonnage OAI

Au déclenchement d'une moisson, le module de moissonnage lit les informations nécessaires à l'interrogation OAI en base de données. Il peut ainsi interroger l'entrepôt correspondant de manière à récupérer les notices.

La récupération peut s'effectuer en mode « FULL » (récupération de l'ensemble des notices) ou « INCREMENTAL » (seulement les notices créées, modifiées ou supprimées depuis la dernière moisson). À cet effet, le module de moissonnage consigne la date et heure de dernière moisson de chaque entrepôt.

Préalablement au lancement d'une moisson (manuel ou planifiée), le module de moissonnage récupère le statut de moissonnage de l'entrepôt. Si celui-ci est libre (pas de moisson en cours), le module inscrit la date de début de moisson et passe le statut à « en cours » et lance la moisson. Si le statut est « en cours », le module de moissonnage vérifie la date de début de dernière moisson :

- si le délai écoulé depuis la date de dernière moisson est supérieur au délai maximal choisi (cf. paramètre « delai.maximal.moisson », la nouvelle moisson est lancée, le module inscrit la date de début de la nouvelle moisson et passe le statut à « en cours ». Un avertissement est enregistré dans les logs
- si le délai écoulé depuis la date de la dernière moisson est inférieure au délai maximal choisi, le module de moissonnage s'arrête avec le message d'erreur suivant : « Une moisson est déjà en cours sur cet entrepôt. ». Aucun mécanisme de rejeu de la moisson n'est mis en place, en effet une moisson aura déjà été faite pour l'entrepôt concerné sur le créneau demandé.

Le statut de l'entrepôt repasse à "libre" à la fin de l'exécution du batch de moissonnage pour un entrepôt.

Dans le cas d'un moissonnage en mode « FULL » d'un entrepôt qui ne gère pas l'information sur la suppression (deletedRecord=no ou deletedRecord=transient), le moissonnage identifie les notices supprimées depuis le dernier moissonnage et l'indique en base de données.

Les notices marquées comme « supprimées » sont étiquetées telles que dans le référentiel des notices ScoLOMFR.

## 2.6.2 Indexation ScoLOMFR

Chaque notice moissonnée est testée dans un premier temps pour vérifier qu'elle est bien conforme au format ScoLOMFR V9.0 version « lax ». Si ce n'est pas le cas, elle est rejetée.

Pour chaque notice valide, l'indexeur ScoLOMFR extrait les informations suivantes :

- Identifiant de la ressource
- Type d'identifiant
- Type de besoin de données à caractère personnel des utilisateurs de la ressource et liste des attributs GAR associés
- Identifiant de l'éditeur de la ressource (composé ainsi : [SIREN]\_[ISNI], le SIREN étant obligatoire et tronqué à 9 caractères, et le ISNI remplacé par des 0 si absent)
- Libellé de l'éditeur de la ressource
- Libellé de la ressource
- Vignette (URL d'accès à une image associée à cette ressource)
- L'URL d'accès à la ressource chez le distributeur technique
- Label GAR
- Date de validation pour le rôle de valideur technique (2.3)
- Les éléments suivants peuvent être multivalués :
  - URI du type pédagogique de la ressource (5.2)
  - URI de la typologie du document (1.10)
  - URI du niveau éducatif détaillé (9.1)
  - URI du domaine d'enseignement (9.1)
  - Code du type de présentation de la ressource dans le médiacentre
- Le libellé des différents éléments multivalués de type URI est extrait des vocabulaires ScoLOMFR correspondants.

## 2.6.3 Traitements de vérification

Une fois l'indexation réalisée, le processus s'assure de la validité des informations récupérées :

- Données liées au GAR (Label GAR, Date de validité du test d'accès) ;
- Présence des VCard de l'éditeur, du distributeur technique, du ou des distributeurs commerciaux et du valideur technique
- Présence au minimum d'un élément de localisation étendue « extendedLocation » pour la plateforme GAR
- Validité des attributs GAR demandés
- Valeur du type de présentation
- Normalisation de certains caractères par d'autres caractères

## 2.6.4 Stockage et mise à jour en base de données

Seules les notices pour lesquelles l'ensemble des informations vérifiées est correct seront conservées en base de données. Les notices qui sont présentes en base GAR et, qui ont été signalées supprimées dans

l'entrepôt ou qui sont absentes de l'entrepôt, seront traitées en tant que suppression et identifiées en base (Cf. §3.1.2.3).

Lorsqu'il s'agit d'une mise à jour de notice, les données extraites remplacent les données existantes.

Durant la moisson d'un entrepôt, les vignettes éditeur dont l'URL est présente dans les notices moissonnées sont systématiquement téléchargées et stockées sur le disque. Ceci permet de s'assurer que les vignettes du GAR sont toujours à jour.

Si le téléchargement d'une vignette éditeur échoue, la version précédente de la vignette (si elle existait) n'est pas supprimée. À chaque téléchargement réussi, le chemin de la vignette GAR correspondante est sauvegardé en base de données.

## 2.7 Envoi de notifications

À la fin de chaque moisson, un courriel est envoyé au gestionnaire de l'entrepôt pour donner un bilan du moissonnage.

De plus, la liste détaillée des erreurs rencontrées lors du traitement des notices, est envoyée aux DTR de ces notices et/ou au gestionnaire de l'entrepôt lorsque les notices sont illisibles ou ne contiennent pas l'information de « diffuseur / distributeur ».

Un paramètre permettra dans un premier temps d'activer l'envoi simultané des notifications au gestionnaire de l'entrepôt et au DTR.

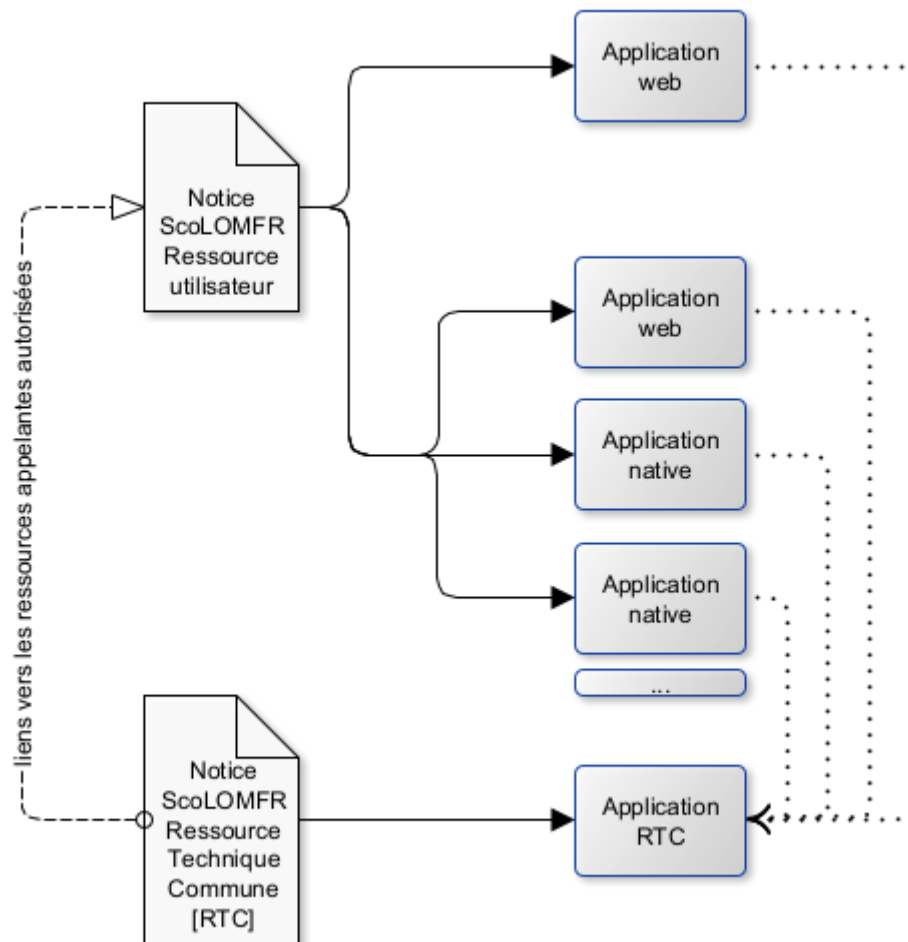
A chaque analyse des notices, les distributeurs techniques sont identifiés par adresse email dans le VCARD.

Le distributeur technique est sauvegardé dans la table de données notice\_vcard.

Si l'adresse email du DTR est manquante, le mail est envoyé uniquement au gestionnaire d'entrepôt.

## 2.8 Différentes variantes d'implémentations des ressources

Les notices peuvent contenir plusieurs variantes d'implémentation d'une ressource : web app, natives ou RTC.



*Schéma présentant les types de ressources et leurs différentes variantes d'implémentations*

Au sein du GAR, une notice ScoLOMFR représente une **ressource utilisateur** ou une **RTC**.

- Une **ressource utilisateur** peut être :
  - Une ressource « autonome » à accès web uniquement

#### 2.8.1.1 Une ressource à accès web avec une ou plusieurs applications Natives (§1.6 Définitions), pour plus de détails voir (§3.2.10.2 Information d'accès aux ressources (utilisateur ou RTC) et informations sur les familles

Pour l'accès aux ressources (utilisateur ou RTC), les éléments concernés sont :

- URL d'accès à la ressource (uniquement pour les ressources utilisateur)
  - Dans le cas d'une RTC, la valeur par défaut suivante sera renseignée en base et non contrôlée sur l'unicité : <http://data.education.fr/gar/rtc>
- URI plateforme pour identifier :
  - Les ressources utilisateur web uniquement
  - Les ressources utilisateur web ayant des variantes natives
  - Les ressources techniques communes (RTC)
- Type de données à caractère personnel
- Attributs à transmettre au DTR lors de l'accès à la ressource
- pour les informations sur les familles :

- Ressource affichable ou non
- Liste des membres de la famille (si la notice est mère de famille)
- Attributs de la famille
- liste des ressources appelantes
- liste des ressources appelées
- pour les informations sur les variantes natives :
  - Lien profond d'accès à l'application
  - Identifiant unique du client
  - Nom du client

### 2.8.1.2 Descriptif de l'élément <extendedLocation>

Les informations d'accès à la ressource utilisateur ou RTC sont regroupées dans un élément « 4.3.1 Localisation étendue » (cf. document ADU9) <extendedLocation> dédié au GAR et reconnaissable par le fait que l'élément « 4.3.1.2 Plateforme » (cf. document ADU9) <scolomfr:platform> a une des valeurs autorisées (cf. 3.2.10.1).

S'il n'existe pas au moins un élément <extendedLocation> répondant à ces exigences, la notice est rejetée.

### 2.8.1.3 Ressources utilisateur web mère de famille et membre de famille

Les informations sur les familles sont optionnelles. Les notices ne comportant pas d'informations famille sont dites « autonomes ».

Les informations sur les familles, lorsqu'elles sont présentes, sont enregistrées également dans l'élément « 4.3.1.4 Description » (cf. document ADU9) <lom:description>. Pour ne pas risquer de collision entre les informations d'accès (attributs GAR) et de familles, celles-ci seront définies dans deux balises <lom:string> distinctes.

La balise <lom:string> contenant les informations de famille doit débuter par la ligne « GAR:Famille ». Si tel n'est pas le cas, les lignes suivantes seront ignorées.

## Spécificité

Les lignes suivantes se présentent sous la forme « Propriété = valeur ». La valeur peut être vide. Les propriétés sont définies dans le tableau suivant :

Propriété	Description	Valeur attendue	Valeur par défaut
GAR:RessourceAffichable	Indique si la ressource liée est affichable	true ou false	False
GAR:MembresFamille	Identifiants des membres de la famille. La présence de cette propriété indique que la ressource est mère de famille	Liste d'identifiants séparés par un point-virgule (« ; »)	Liste vide
GAR:AttributsFamille	Attributs de la famille	Liste d'attributs séparés par un point-virgule (« ; »)	Liste vide

GAR:IdRessourcesAppelantes	Identifiants des ressources appelantes	Liste d'identifiants de ressource séparés par un point-virgule (« ; »)	Liste vide
GAR:IdRessourcesAppelées	Identifiants des ressources appelées	Liste d'identifiants de ressource séparés par un point-virgule (« ; »)	Liste vide

## Notes :

- Les attributs de famille (propriété « GAR:AttributsFamille ») sont présents sous la forme « GAR:AttributsFamille = [IDC] Identifiant complémentaire : [LRA] Liste des ressources appelées ». Le module de moissonnage vérifie que tous les attributs famille mentionnés sont bien présents dans la table de référence « Attributs GAR famille » par extraction des codes des attributs placés entre crochets. Si ce n'est pas le cas, la notice est rejetée.
- L'attribut [IDC] ne peut figurer dans la liste des attributs famille que si la notice est une « mère de famille ». Dans le cas contraire, la notice est rejetée.
- Les propriétés du bloc famille doivent être uniques. En cas de doublon (par exemple 2 fois la propriété GAR:IdRessourcesAppelantes), la notice est rejetée.
- Tous les séparateurs peuvent être utilisés pour séparer la liste des Attributs GAR. La recherche des attributs se fait sur la détection des crochets ouvrant et fermant ([ ])

## Exemple :

```
<lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI]
code établissement ---- [IDO] id opaque
</lom:string>
```

donnera une liste d'attributs [UAI, IDO], le libellé de l'attribut n'étant pas utilisé pour l'analyse.

## Exemple de Mère de famille

```
<scolomfr:extendedLocation>
  <scolomfr:location>https://www.edumedia-sciences.com/sso-gar/</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv3.0</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier directement l'individu et/ou comportent des éléments de son état civil ou de sa vie privée</scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-1/fr"> [Toute information de description de l'accès spécifique à la ressource] </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque</lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">
GAR:Famille
GAR:RessourceAffichable = true
GAR:MembresFamille = id1 ; id2 ; id3
GAR:AttributsFamille = [IDC] Identifiant complémentaire ; [LRA] Liste des ressources appelées
GAR:IdRessourcesAppelantes = id1
GAR:IdRessourcesAppellees = id2
  </lom:string>
</lom:description>
</scolomfr:extendedLocation>
```

## Exemple de Membre d'une famille

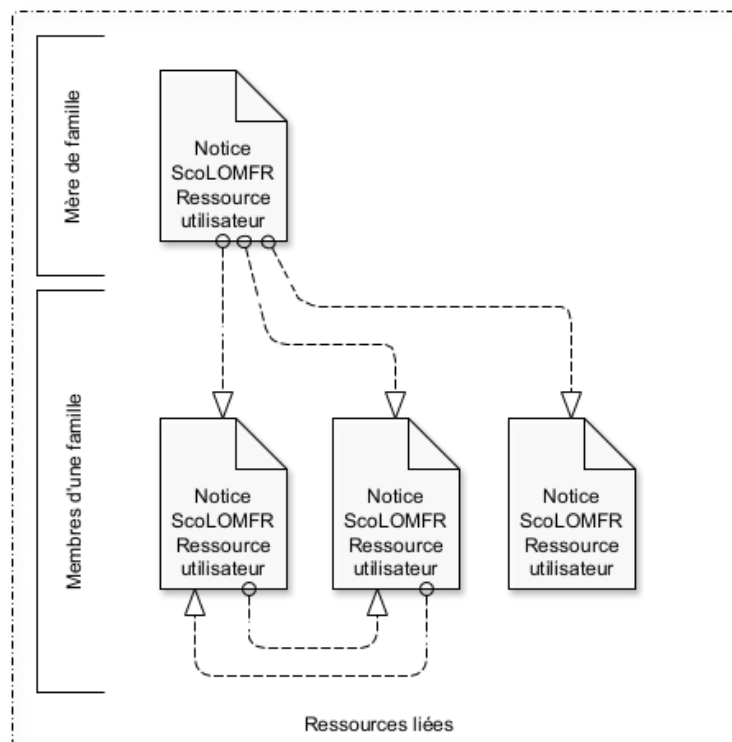
```
<scolomfr:extendedLocation>
  <scolomfr:location>https://www.edumedia-sciences.com/sso-gar/</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv3.0</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier directement l'individu et/ou comportent des éléments de son état civil ou de sa vie privée</scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-1/fr"> [Toute information de description de l'accès spécifique à la ressource] </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">
GAR:Famille
GAR:RessourceAffichable = false
GAR:AttributsFamille = [LRA] Liste des ressources appelées
GAR:IdRessourcesAppelantes = id1 ; id2
GAR:IdRessourcesAppellees = id2 ; id3
  </lom:string>
</lom:description>
</scolomfr:extendedLocation>
```

- Ressources utilisateur web et web ayant une ou plusieurs variantes native)

- Une ressource liée « mère de famille » ou « membre de famille » (cf.2.9)
- Une **RTC** (§1.6 Définitions) pour plus de détails voir §3.2.10.6

## 2.9 Notion de familles de ressources

La notion de famille de ressources permet à un éditeur de déclarer dans le GAR des liens entre les ressources.



*Les liens entre les ressources mère de famille et membres d'une famille*

Chaque famille dispose d'une ressource mère de famille qui porte les informations principales de la famille. Chez le fournisseur de ressources, elle offre un service transverse à plusieurs ressources. Les autres ressources composant la famille sont les membres.

Les membres d'une famille et la mère de famille sont appelés des ressources liées. Les ressources qui n'appartiennent pas à une famille sont des ressources autonomes.

Les ressources d'une famille peuvent s'appeler entre elles, c'est-à-dire que l'accédant aura, par exemple, la possibilité de passer de l'une à l'autre. Chaque notice déclare les ressources qu'elle appelle (ressources appelées) et celles qui peuvent l'appeler (ressources appelantes).

Une ressource technique dans une famille peut ne pas être visuellement présentable à l'utilisateur, il s'agit alors d'une ressource « non affichable ».

L'ID complémentaire (IDC) est un identifiant qui, s'il est fourni lors de l'accès aux ressources, permet au DTR de faire le lien entre les ids opaques de l'accédant pour les ressources de la famille.

La liste des ressources appelées (LRA) est une liste de ressources demandées par l'éditeur qui est filtrée grâce aux affectations de l'utilisateur. Cela permet par exemple à l'éditeur de proposer un menu de ressources sur ses interfaces ou de créer des passerelles entre ressources.



Exemples :

1. Portail de recherche fédérée, fonctionnant sur un ensemble défini de ressources de contenus.  
Dans le GAR, le portail est la mère de la Famille, il appelle les différents membres, il est affichable.  
L'id complémentaire ne semble pas nécessaire à ce cas.
2. Service de type « espace personnel » transverse à plusieurs ressources de contenus permettant de conserver des travaux élèves/enseignant et/ou des données d'usage.  
Dans le GAR, ce service est la mère de famille, il est appelé par les différents membres, il est affichable. L'id complémentaire est nécessaire à ce cas, pour afficher des informations provenant de différentes ressources pour un même accédant.
3. Environnement de stockage commun à plusieurs versions d'une ressource. Exemple, service d'arrière-plan commun à des versions différentes d'un manuel numérique (web, android, ios, etc.).  
Dans le GAR, ce service est la mère de famille, il est appelé par les différents membres, il n'est pas affichable. L'id complémentaire est nécessaire à ce cas.
4. Dictionnaire utilisable dans un ensemble de manuel (lien souligné).  
Dans le GAR, le dictionnaire est la mère de famille, il est appelé par les différents membres, il est affichable. L'id complémentaire ne semble pas nécessaire à ce cas.

## 2.10 Organisation du système de fichiers et accès aux vignettes

### 2.10.1 Organisation du système de fichiers

Les vignettes sont organisées dans le système de fichiers comme suit :

- un dossier racine contenant un sous-dossier par entrepôt nommé en utilisant le code de l'entrepôt
  - dans chaque dossier d'entrepôt :
    - un dossier « original » contenant les images téléchargées par le module de moissonnage (i.e les vignettes éditeurs), les images y sont nommées « {iddelanotice}.{extension} »
      - les vignettes éditeurs des ressources « mère de famille » sont téléchargées dans un sous-répertoire « mere » du dossier « original ».
    - un dossier « gar » contenant les vignettes GAR générées
    - le script de normalisation des vignettes, les images standardisées sont nommées « {iddelanotice}.png »
      - Les vignettes MÈRE DE FAMILLE sont générées dans un dossier « mere » sous répertoire du dossier « gar ».

### 2.10.2 Accès aux vignettes GAR et MÈRE DE FAMILLE

Un module GAR souhaitant accéder à la vignette GAR ou à une vignette MÈRE DE FAMILLE associée à une notice consultera la table des notices (cf. §4.2) pour en extraire le chemin de la vignette GAR ou MÈRE DE FAMILLE permettant de composer l'url d'accès à la vignette GAR ou à la vignette MÈRE DE FAMILLE (si celles-ci sont présentes et ont été standardisées pour le GAR).

## 3 Notes techniques

Le module de moissonnage des ressources est un outil « batch » en ligne de commande sans interface Web. Il est lancé automatiquement par tâche de type cron. Il interagit avec :

- la base de données MySQL pour :
  - lire la configuration des moissons à effectuer et tenir à jour la date et l'heure de dernière moisson ;
  - stocker les données issues de la moisson ;
  - récupérer les informations relatives au gestionnaire d'un entrepôt pour l'envoi d'une notification par courriel en fin de moisson
  - récupérer les informations relatives aux attributs GAR pour calculer le type de DCP correspondant aux attributs demandés
- le système de fichiers pour sauvegarder les vignettes des notices
- le système de log pour tracer l'ensemble des étapes de fonctionnement nécessaires à la mise en place d'indicateurs statistiques et alertes opérationnelles
- le serveur d'envoi de mail pour l'envoi des rapports de moisson aux gestionnaires d'entrepôts et/ou aux DTR concernés

### 3.1 Appel du module de moissonnage

Le module de moissonnage est disponible sous forme d'une archive jar exécutable nommée garharvest.jar. Il peut être lancé en ligne de commande, via un script shell ou par une tâche cron. Il nécessite la présence du JDK et la définition d'une variable d'environnement (cf. §Configuration du module).

Le lancement du module s'effectue comme suit :

```
java -jar garharvest.jar <commande> [ <paramètre> ]
```

La commande est obligatoire, le paramètre est optionnel, il dépend de la commande choisie. Le tableau ci-dessous indique l'ensemble des commandes et paramètres possibles :

Commande	Paramètre	Effet
--version	Non	Affiche la version du module de moisson
--harvest	code entrepôt	Moissonne l'entrepôt identifié
--reset	code entrepôt	Supprime les données de moisson relatives à l'entrepôt identifié
test	Non	Effectue un diagnostic de l'installation du module de moissonnage

Notes :

- Lorsque le module de moissonnage est lancé sans paramètres de ligne de commande, une notice d'usage est affichée.
- La notion de « code entrepôt » renvoie à la définition de l'ensemble des paramètres nécessaires à la moisson d'un entrepôt donné (cf. §Configuration des entrepôts à moissonner)

#### 3.1.1 Configuration du module

### 3.1.1.1 Paramètres techniques

Le module de moissonnage interagit avec la base de données MySQL ainsi que le serveur mail de la plateforme. La configuration de l'accès à ces services est stockée dans un fichier de type « *.properties* » dont l'ensemble des directives attendues est indiqué dans le tableau ci-dessous :

Directive	Valeur attendue
<b>db.hibernate.dialect</b>	Dialecte Hibernate pour la base de données
<b>db.jdbc.driver.class</b>	Driver JDBC à utiliser pour la connexion
<b>db.jdbc.connection.url</b>	URL JDBC d'accès à la base de données
<b>db.host</b>	Nom réseau du serveur de base de données
<b>db.name</b>	Nom de la base de données
<b>db.user</b>	identifiant de connexion
<b>db.pass</b>	mot de passe de connexion
<b>mail.default.from</b>	expéditeur des courriels de notification
<b>mail.host</b>	nom du serveur mail
<b>mail.port</b>	port du serveur mail
<b>mail.user</b>	identifiant de connexion au serveur (peut être laissé vide)
<b>mail.pass</b>	mot de passe de connexion au serveur (peut être laissé vide)
<b>log.dir</b>	Dossier de log du module de moissonnage
<b>label.gar.duree.validite</b>	durée en année de validité de la date de validation GAR
<b>notifier.tous</b>	les mails de notification sur le contenu des notices sont envoyés aussi au gestionnaire d'entrepôt lorsque la directive est à true. Les mails de notification sur le contenu des notices sont envoyés uniquement au DTR de la notice concernée lorsque la directive est à false.
<b>img.root.dir</b>	Chemin complet de la racine des dossiers contenant les vignettes
<b>img.thn.size</b>	Dimension en pixels de la vignette GAR à créer, ex : « 300x260 »
<b>img.thn.bb.size</b>	Dimension en pixels pour la mise à l'échelle de la vignette éditeur, ex : « 250x250 »
<b>img.thn.bb.pos</b>	Position en pixels du coin haut gauche de la boîte englobante contenant la vignette éditeur retaillée, ex : « 5x5 »
<b>img.label.pos</b>	Position en pixels du coin haut gauche de la pastille GAR
<b>img.mere.pos</b>	Position en pixels du coin haut gauche de la pastille MÈRE DE FAMILLE
<b>pastille.file</b>	Chemin du fichier « pastille.png » utilisé pour la génération des vignettes GAR
<b>pastille_mere.file</b>	Chemin du fichier « mere.png » utilisé pour la génération des vignettes MÈRE DE FAMILLE
<b>controle.gar.date.validite</b>	Le contrôle sur la date de validité GAR est désactivé lorsque la directive vaut false.
<b>delai.maximal.moisson</b>	Délai maximal (exprimé en minutes) à partir duquel on considère qu'une moisson est en erreur. Utilisé pour débloquer et pour permettre une nouvelle moisson sur un entrepôt dont la dernière moisson est bloquée.

Note : les directives dont le nom commence par « img. » et les directives « pastille.file » et « pastille\_mere.file » sont utilisées exclusivement par le script de standardisation des vignettes. Seule exception, la directive « img.root.dir » est également utilisée par le module de moissonnage pour l'emplacement de stockage des vignettes éditeur téléchargées.

Le module de moissonnage identifie le fichier de configuration à utiliser via la variable d'environnement GARHARVEST\_CONF. Ainsi, si les informations de configuration sont contenues dans le fichier /opt/moisson/conf/moisson\_prod.properties, il pourra être lancé comme suit :

```
export GARHARVEST_CONF=/opt/moisson/conf/moisson_prod.properties
java -jar garharvest.jar [...]
```

### 3.1.1.2 Configuration des entrepôts

L'appel du module de moissonnage se base sur la notion de « code d'entrepôt ». Un code d'entrepôt identifie de manière unique un ensemble de paramètres décrivant l'entrepôt à moissonner. Le paramétrage d'un entrepôt en base de données est le préalable à l'appel du module de moissonnage (cf. §Stockage des données).

## 3.1.2 Traitement des notices orphelines

Les notices orphelines représentent les notices identifiées en base de données GAR et qui ne sont plus présentes dans l'entrepôt OAI correspondant.

### 3.1.2.1 Lors du moissonnage OAI

La solution repose sur un moissonnage en mode complet d'un entrepôt. Les entrepôts dont le moissonnage serait en mode partiel et les entrepôts en mode complet qui gèrent l'information sur la suppression seront ignorés.

Précisions techniques :

- ⇒ Le processus identifie pour chaque entrepôt, à chaque moisson si celle-ci est complète ou partielle afin d'ignorer cette moisson dans le traitement des notices orphelines en cas de moisson partielle (interrogation base de données – table entrepôt)
- ⇒ Le processus identifie pour chaque entrepôt en mode moissonnage complet, à chaque moisson si celle-ci gère la suppression en mode « deletedRecord = persistent » afin de l'ignorer (interrogation OAI « verb=Identify » de l'entrepôt)
- ⇒ Le processus permet de conserver une copie en mémoire des identifiants OAI des notices de l'entrepôt

### 3.1.2.2 Lors de l'indexation

L'indexation prendra en compte l'identification des notices supprimées qu'elles soient signalées dans l'entrepôt ou manquantes dans l'entrepôt par rapport à la base.

Les notices qui disparaissent de l'entrepôt sans signalement, seront traitées en tant que suppressions.

Précisions techniques :

- ⇒ Pour chaque identifiant OAI de la base, le processus vérifie son existence dans ceux de l'entrepôt
- ⇒ Si l'identifiant OAI n'existe pas dans l'entrepôt, le processus permet de constituer un tableau des notices supprimées de l'entrepôt par leur identifiant interne en base de données.

### 3.1.2.3 Lors de l'enregistrement en base

Le processus d'identification des notices à supprimer en base, hors celles identifiées dans les entrepôts, marquera celles existant en base et manquant dans l'entrepôt, en notice manquante « MISSING ».

Précisions techniques :

- ⇒ Le processus permet de tagguer chaque notice supprimée de l'entrepôt dans la table « notice ». Cela concerne :
  - le champ « notice\_statut » est valorisé avec la valeur « MISSING »
  - le champ « notice\_supprime » est valorisé avec la valeur « true »

## 3.2 Validation et indexation des notices

Lors de la moisson, toutes les notices récupérées sont analysées. Lors de cette analyse, le module de moissonnage extrait les indicateurs (titre, identifiant...) nécessaires à l'exploitation des ressources. Si un problème est détecté, la notice est rejetée et ne fait l'objet d'aucun stockage en base de données.

Toutes les erreurs ou avertissements produits par la phase d'analyse des notices sont consignés dans les logs de la moisson et inclus dans le mail récapitulatif à destination du responsable de l'entrepôt et/ou du DTR de la notice.

Les notices dont l'identifiant (« [catalogue]:[entrée] ») serait déjà utilisé sont rejetées. Les DTR des notices concernées ainsi que les gestionnaires des entrepôts concernés sont notifiés du problème.

Les paragraphes ci-dessous décrivent les éléments extraits de chaque notice ScoLOMFR et les contrôles afférents.

Note : A partir de la version ScoLOMFR 6.0, il existe une déclaration et documentation de l'ontologie ScoLOMFR :

<https://www.reseau-canope.fr/scolomfr/ontologie/comprehension-de-lontologie.html>

Note : Dans les sous-chapitres qui suivent les exemples de code mettent en évidence les contenus recherchés par une mise en surbrillance **jaune** et les éléments contrôlés en surbrillance **magenta**.

### 3.2.1 Identifiant

#### 3.2.1.1 Descriptif de l'élément

Les éléments concernés sont :

- Type d'identifiant
- Identifiant de notice

L'élément « 1.1 Identifiant » (cf. document ADU9) est répétable. Parmi les identifiants présents, seuls ceux associés à un catalogue reconnu sont pris en compte (voir liste en §7). Parmi ceux-ci, l'identifiant retenu est celui appartenant au premier catalogue dans l'ordre de priorité. Si plusieurs identifiants appartiennent à ce catalogue, la notice est rejetée.

Si une modification de l'identifiant est détectée par rapport au champ « notice\_id » d'une même ressource identifiée par le champ « notice\_id\_oai », la notice est rejetée et une alerte est envoyée au gestionnaire de l'entrepôt et/ou au DTR de la notice.

### 3.2.1.2 Exemple

```
<lom:identifiant>
  <lom:catalog>ARK</lom:catalog>
  <lom:entry>ark:/19153/b2nk3c28</lom:entry>
</lom:identifiant>
```

## 3.2.2 Titre

### 3.2.2.1 Descriptif de l'élément

L'élément « 1.2 Titre » (cf. document ADU9) contient le titre de la notice.

Ce titre est extrait du contenu de la balise <lom:string> fille sélectionnée parmi les balises <lom:string> répondant aux critères suivants (par ordre de priorité décroissant) :

- balise avec attribut @language = « http://id.loc.gov/vocabulary/iso639-2/fre »
- ou balise avec attribut @language = « http://id.loc.gov/vocabulary/iso639-2/fra »
- ou balise avec attribut @language = « fre »
- ou balise avec attribut @language = « fra »
- ou balise sans attribut @language

Si plusieurs balises répondent au même critère, c'est la première balise rencontrée qui est utilisée.

Si aucune balise <lom:string> ne répond aux critères, le titre ne peut être extrait et la notice est rejetée.

Dans le cas d'une notice **mère de famille**, le titre devrait contenir exactement « **mère de famille** ».

Dans le cas d'une **RTC**, le titre devrait commencer exactement par « **[RTC]** ». (cf. Règle 58 )

Dans le cas contraire un message d'alerte est envoyé au DTR et Gestionnaire d'entrepôt

### 3.2.2.2 Exemple

```
<lom:title>
  <lom:string language="http://id.loc.gov/vocabulary/iso639-2/spa">Diccionario
  técnico</lom:string>
  <lom:string>Dictionnaire technique</lom:string>
  <lom:string language="http://id.loc.gov/vocabulary/iso639-2/deu">Technische
  Wörterbuch</lom:string>
</lom:title>
```

Le contenu extrait sera « Dictionnaire technique » car c'est la seule balise <lom:string> répondant à un des critères listés (balise sans attribut @language).

```

<lom:title>
  <lom:string>Dictionnaire technique</lom:string>
  <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Mon dictionnaire
technique</lom:string>
  <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Dictionnaire avec des
illustrations</lom:string>
</lom:title>

```

Le contenu extrait sera « Mon dictionnaire technique » car le critère « balise avec attribut @language = « http://id.loc.gov/vocabulary/iso639-2/fre » » est prioritaire sur le critère « balise sans attribut @language » et que c'est la première balise à apparaître avec ce critère.

### 3.2.3 URL de la vignette

#### 3.2.3.1 Descriptif de l'élément

L'élément « 7. Relation » (cf. document ADU9) est facultatif et répétable. Si aucune relation de type « a pour vignette » n'est trouvée dans la notice, ou si cet élément ne contient pas d'URL valide vers une image, le DTR en sera notifié.

Si plusieurs relations « a pour vignette » sont trouvées, aucune n'est retenue et le DTR en est notifié.

Si une seule relation est trouvée et que l'URL est correcte, l'image en question est téléchargée, stockée sur le disque (avec l'identifiant de la notice) et son chemin est sauvegardé en base de données. Si le téléchargement échoue, le DTR en est notifié.

#### 3.2.3.2 Exemple

```

<lom:relation>
  <lom:kind>
    <lom:source>SCOLOMFRv2.1</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-009-num-
021</lom:value>
    <lom:label>a pour vignette</lom:label>
  </lom:kind>
  <lom:resource>
    <lom:identifiant>
      <lom:catalog>URI</lom:catalog>
      <lom:entry>http://exemple.com/notice_vignette.png</lom:entry>
    </lom:identifiant>
    <lom:description/>
  </lom:resource>
</lom:relation>

```

### 3.2.4 Label GAR

L'encodage du label GAR a changé entre ScoLOMFR v3.0 et ScoLOMFR v3.2 (jusqu'à la version actuelle). Par souci de compatibilité, le module de moissonnage accepte les deux formes.

### 3.2.4.1 Version ScoLOMFR v3.0

#### Descriptif de l'élément

Pour être labellisé GAR, la notice doit comprendre un élément « 9.1 Classification » (cf. document ADU9) lié au vocabulaire 045 « Label » et contenant le concept « FARE » dont l'identifiant est :  
« <http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-045-num-001> »

Si cet élément n'est pas trouvé dans la notice, cette dernière est rejetée.

#### Exemple

```
<lom:classification>
  <lom:purpose>
    <lom:source>SCOLOMFRv3.0</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-028-num-013</lom:value>
    <lom:label>label</lom:label>
  </lom:purpose>
  <lom:taxonPath>
    <lom:source>
      <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Label - scolomfr-voc-045</lom:string>
    </lom:source>
    <lom:taxon>
      <lom:id>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-045-num-001</lom:id>
      <lom:entry>
        <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">FARE</lom:string>
      </lom:entry>
    </lom:taxon>
  </lom:taxonPath>
</lom:classification>
```

### 3.2.4.2 Versions ScoLOMFR v3.2 et ultérieures

#### Descriptif de l'élément

Pour être labellisé GAR, la notice doit comprendre un élément « 9.1 Classification » (cf. document ADU9) lié au vocabulaire 045 « Label » et contenant le concept « GAR » dont l'identifiant est :  
« <http://data.education.fr/gar> ».



## Exemple

```
<lom:classification>
  <lom:purpose>    <lom:source>SCOLOMFRv3.2-skos</lom:source>
<lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-028-num-013</lom:value>

<lom:label>label</lom:label>
</lom:purpose>
<lom:taxonPath>
  <lom:source>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Label - scolomfr-voc-
045</lom:string>
  </lom:source>
  <lom:taxon>
    <lom:id>http://data.education.fr/gar</lom:id>
    <lom:entry>
      <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">GAR</lom:string>
    </lom:entry>
  </lom:taxon>
</lom:taxonPath>
</lom:classification>
```

## 3.2.5 Date de validation GAR

### 3.2.5.1 Descriptif de l'élément

La date de validation GAR est stockée dans l'élément « 2.3.3.1 Date » (cf. document ADU9). Cette date est identifiée comme la date de validation du GAR par la présence de la chaîne « GAR : validation technique » dans l'élément « 2.3.3.2 Description » (cf. document ADU9)

Cette date de validation est présente dans la contribution avec le rôle « valideur technique »

Si cette information n'est pas présente, la notice est rejetée.

Si la date est considérée comme ancienne (Voir §3.1.1.1), une alerte est envoyée au DTR et/ou au gestionnaire de l'entrepôt.

Note : le contrôle sur la date de validité peut être désactivé en configuration (cf. §Paramètres techniques).

### 3.2.5.2 Exemple

```
<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.0</lom:source>
    <lom:value
uniqueElementName="value">http://data.education.fr/voc/scolomfr/concept/technical_validator</lom:
value>
    <lom:label>valideur technique</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:5.0
...
END:VCARD]]></lom:entity>
  <lom:date>
    <lom:dateTime>2016-10-01</lom:dateTime>
    <lom:description>
      <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">GAR : validation
technique</lom:string>
    </lom:description>
  </lom:date>
</lom:contribute>
```

## 3.2.6 VCard de l'éditeur

### 3.2.6.1 Descriptif de l'élément

La section « 2. Cycle de vie » (cf. document ADU9) doit contenir un ou plusieurs vcards « éditeur ».

Parmi ces vcards, une seule doit contenir un numéro SIREN présent dans la table de référence des éditeurs ayant adhéré au cadre de confiance du GAR.

Cette Vcard est stockée pour mise à disposition des autres services du GAR (Voir table « notice\_vcard »).

Les autres vcards « éditeur » ne sont pas stockées.

La notice est rejetée, dans les autres cas :

- Aucune vcard « éditeur » dont le numéro SIREN est trouvé dans la table de référence des éditeurs
- Plusieurs vcards « éditeur » dont le numéro SIREN est trouvé dans la table de référence des éditeurs

Note : le numéro de SIREN contenu dans la vcard peut contenir des espaces, ceux-ci sont alors automatiquement retirés.

### 3.2.6.2 Exemple

```

<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.0</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/publisher</lom:value>
    <lom:label>éditeur</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:5.0
...
END:VCARD]]></lom:entity>
  ...
</lom:contribute>

```

## 3.2.7 VCard du distributeur technique

### 3.2.7.1 Descriptif de l'élément

Cette Vcard correspond à celle du DTR.

Ces coordonnées sont celles à qui doivent être envoyées les erreurs rencontrées dans la notice.

Si l'information de SIREN ou de courriel est absente de la VCard du distributeur technique, la notice est rejetée.

Il ne doit y avoir qu'une seule VCard « DTR » par notice.

Cette Vcard est stockée pour mise à disposition des autres services du GAR (Voir table « notice\_vcard »).

En ScoLOMFRv3.0, le rôle associé au distributeur technique est « diffuseur / distributeur », depuis ScoLOMFRv3.2, un nouveau rôle « distributeur technique » a été ajouté. Par souci de compatibilité, le module de moissonnage accepte les deux formes.

Dans la Vcard du DTR est ajouté le champ NOTE:X-PLATEFORME-ID précisant la plateforme technique, cet identifiant est facultatif.

S'il est présent, sa valeur doit comporter 2 chiffres

S'il est absent la valeur par défaut 00 est utilisée.

Si la valeur n'est pas conforme ou que le champ est multivalué, la notice est rejetée

### 3.2.7.2 Exemple de Vcard DTR avec Plateforme

```

BEGIN:VCARD
VERSION:4.0
FN:PERRIER
N:PERRIER;OLIVIER;;Mr;
ORG:Diffu-G
ADR:CS 50109;;131 RUE AUGUSTE RENOIR;BOURG EN BRESSE CEDEX;01003;;France
TEL:0474504170
EMAIL:elisa@le-genie.com
NOTE:SIREN=343605952
NOTE:X-PLATEFORME-ID=15
REV:05/06/2018
END:VCARD

```

### 3.2.7.3 Version ScoLOMFR v3.0

#### Spécificité

Cette Vcard est reconnue comme celle du DTR lorsque le rôle vaut « diffuseur / distributeur » (identifiant : « <http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-017> ») et que la description associée contient « GAR : distributeur technique ».

#### Exemple

```

<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.0</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-017</lom:value>
    <lom:label>diffuseur / distributeur</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:4.0
...
END:VCARD]]></lom:entity>
  <lom:date uniqueElementName="date">
    <lom:dateTime uniqueElementName="dateTime">2016-07-01</lom:dateTime>
    <lom:description uniqueElementName="description">
      <lom:string>GAR : distributeur technique</lom:string>
    </lom:description>
  </lom:date>
</lom:contribute>

```

### 3.2.7.4 Version ScoLOMFR v3.2 et ultérieures

#### Spécificité

Cette Vcard est reconnue comme celle du DTR lorsque le rôle vaut « distributeur technique » (identifiant « <http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-026> ») et que la description associée contient « GAR : distributeur technique ».

#### Exemple

```
<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.2-skos</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-026</lom:value>
    <lom:label>distributeur technique</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:4.0
..
END:VCARD]]></lom:entity>
  <lom:date uniqueElementName="date">
    <lom:dateTime uniqueElementName="dateTime">2016-07-01</lom:dateTime>
    <lom:description uniqueElementName="description">
      <lom:string>GAR : distributeur technique</lom:string>
    </lom:description>
  </lom:date>
</lom:contribute>
```

## 3.2.8 VCards des distributeurs commerciaux

### 3.2.8.1 Descriptif de l'élément

La notice doit contenir au moins une VCard de distributeur commercial, elle peut en contenir plusieurs.

Ces Vcards sont stockées pour mise à disposition des autres services du GAR (Voir table « notice\_vcard »).

En ScoLOMFRv3.0, le rôle associé au distributeur commercial est « diffuseur / distributeur », depuis ScoLOMFRv3.2, un nouveau rôle « diffuseur » a été ajouté. Par souci de compatibilité, le module de moissonnage accepte les deux formes.

### 3.2.8.2 Version ScoLOMFR v3.0

#### Spécificité

Une Vcard est reconnue comme distributeur commercial lorsque le rôle vaut « diffuseur / distributeur » (<http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-017>) et que la description associée contient « GAR : distributeur commercial ».

## Exemple

```
<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.0</lom:source>
  <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-017</lom:value>
  <lom:label>diffuseur / distributeur</lom:label>
</lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:4.0
...
END:VCARD]]]></lom:entity>
  <lom:date uniqueElementName="date">
    <lom:dateTime uniqueElementName="dateTime">2016-07-01</lom:dateTime>
    <lom:description uniqueElementName="description">
      <lom:string>GAR : distributeur commercial</lom:string>
    </lom:description>
  </lom:date>
</lom:contribute>
```

### 3.2.8.3 Version ScoLOMFR v3.2 et ultérieures

## Spécificité

Une Vcard est reconnue comme distributeur commercial lorsque le rôle vaut « diffuseur » (<http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-027>) et que la description associée contient « GAR : distributeur commercial ».

## Exemple

```
<lom:contribute>
  <lom:role>
    <lom:source> SCOLOMFRv3.2-skos</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-003-num-027</lom:value>
    <lom:label>diffuseur</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:4.0
...
END:VCARD]]]></lom:entity>
  <lom:date uniqueElementName="date">
    <lom:dateTime uniqueElementName="dateTime">2016-07-01</lom:dateTime>
    <lom:description uniqueElementName="description">
      <lom:string>GAR : distributeur commercial</lom:string>
    </lom:description>
  </lom:date>
</lom:contribute>
```

## 3.2.9 VCard du valideur technique

### 3.2.9.1 Descriptif de l'élément

Cette Vcard est stockée pour mise à disposition des autres services du GAR (Voir table « notice\_vcard»).

Note : la date de validation GAR est stockée dans cet élément <contribute> comme précisé dans cf. 3.2.5 Date de validation GAR.

La section « 2. Cycle de vie » doit contenir une et une seule vcard « valideur technique ». Si ce n'est pas le cas, la notice est rejetée.

Toutes les Vcards correspondantes aux autres rôles pouvant être rencontrés dans une notice (auteur, créateur...) sont ignorées.

### 3.2.9.2 Exemple

```
<lom:contribute>
  <lom:role>
    <lom:source>SCOLOMFRv3.0</lom:source>
    <lom:value>http://data.education.fr/voc/scolomfr/concept/technical_validator</lom:value>
    <lom:label>valideur technique</lom:label>
  </lom:role>
  <lom:entity><![CDATA[BEGIN:VCARD
VERSION:4.0
...
END:VCARD]]></lom:entity>
  ...
</lom:contribute>
```

## 3.2.10 Informations des éléments extendedLocation

Une **ressource utilisateur** peut comporter plusieurs éléments « extendedLocation » avec l'obligation de disposer à minima d'un élément « extendedLocation » couvrant l'accès à l'application web correspondante.

Une **ressource technique commune (RTC)** comporte un seul élément « extendedLocation ».

Une ressource utilisateur peut également être une ressource mère de famille ou membre d'une famille de ressources.

Chaque type d'accès est décrit dans un élément « extendedLocation ». Les éléments « extendedLocation » sont groupés dans l'élément « technical » de la notice.

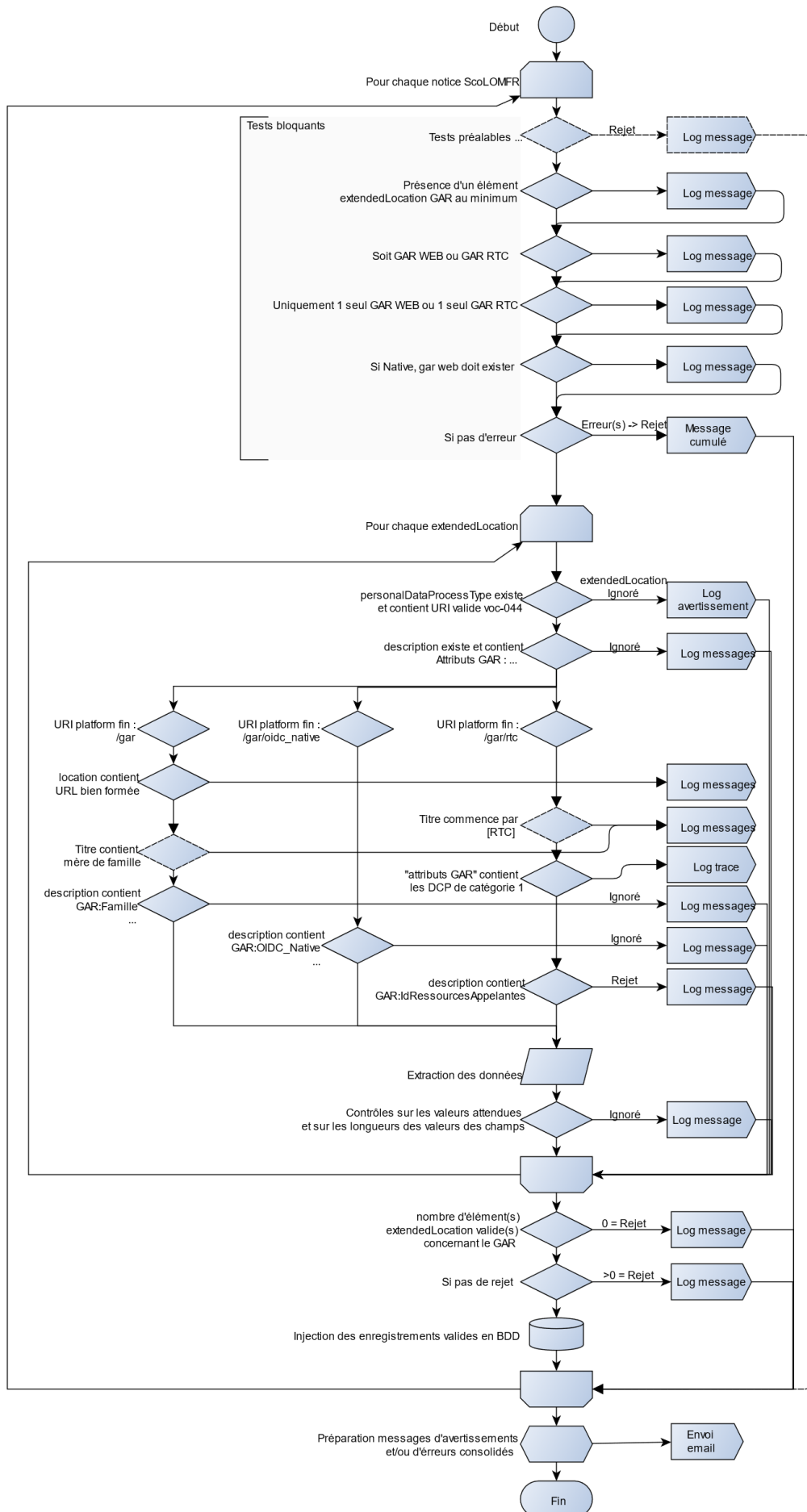
Au sein de l'élément « extendedLocation », 4 éléments sont contrôlés :

- La localisation : élément <scolomfr:location>
- La plateforme : élément <scolomfr:platform>
- Les données à caractère personnel : élément <scolomfr:personalDataProcessType>
- La description : élément <lom:description>

Les extendedLocation qui concernent le GAR sont ceux contrôlés sur l'élément <scolomfr:platform> qui contient bien un URI autorisé GAR – cf.3.2.10.1).

Les autres éléments « extendedLocation » qui peuvent concerner d'autres sujets que le GAR, sont ignorés.

Le schéma suivant présente le processus de traitement des extendedLocation d'une notice ScoLOMFR :





### 3.2.10.1 Valeurs autorisées pour l'élément « scolomfr:platform »

Pour qu'un élément « extendedLocation » soit concerné par le GAR, son sous-élément « platform » doit avoir un des URI suivants :

- <http://data.education.fr/gar> (variante web, mère de famille ou membre de famille)
- [http://data.education.fr/gar/oidc\\_native](http://data.education.fr/gar/oidc_native) (variante native)
- <http://data.education.fr/gar/rtc>(RTC)

Règle 40 : Le contenu de l'élément « platform » de chaque élément « extendedLocation » devrait contenir une des valeurs autorisées, sinon l'élément extendedLocation est ignoré et un message d'avertissement rappelle la raison : « extendedLocation ignoré car le contenu de l'élément « platform » ne contient pas une des valeurs [liste\_valeurs] ».

### 3.2.10.2 Information d'accès aux ressources (utilisateur ou RTC) et informations sur les familles

Pour l'accès aux ressources (utilisateur ou RTC), les éléments concernés sont :

- URL d'accès à la ressource (uniquement pour les ressources utilisateur)
  - Dans le cas d'une RTC, la valeur par défaut suivante sera renseignée en base et non contrôlée sur l'unicité : <http://data.education.fr/gar/rtc>
- URI plateforme pour identifier :
  - Les ressources utilisateur web uniquement
  - Les ressources utilisateur web ayant des variantes natives
  - Les ressources techniques communes (RTC)
- Type de données à caractère personnel
- Attributs à transmettre au DTR lors de l'accès à la ressource
- pour les informations sur les familles :
  - Ressource affichable ou non
  - Liste des membres de la famille (si la notice est mère de famille)
  - Attributs de la famille
  - liste des ressources appelantes
  - liste des ressources appelées
- pour les informations sur les variantes natives :
  - Lien profond d'accès à l'application
  - Identifiant unique du client
  - Nom du client

### 3.2.10.3 Descriptif de l'élément <extendedLocation>

Les informations d'accès à la ressource utilisateur ou RTC sont regroupées dans un élément « 4.3.1 Localisation étendue » (cf. document ADU9) <extendedLocation> dédié au GAR et reconnaissable par le fait que l'élément « 4.3.1.2 Plateforme » (cf. document ADU9) <scolomfr:platform> a une des valeurs autorisées (cf. 3.2.10.1).

S'il n'existe pas au moins un élément <extendedLocation> répondant à ces exigences, la notice est rejetée.

### 3.2.10.4 Ressources utilisateur web mère de famille et membre de famille

Les informations sur les familles sont optionnelles. Les notices ne comportant pas d'informations famille sont dites « autonomes ».

Les informations sur les familles, lorsqu'elles sont présentes, sont enregistrées également dans l'élément « 4.3.1.4 Description » (cf. document ADU9) <lom :description>. Pour ne pas risquer de collision entre les informations d'accès (attributs GAR) et de familles, celles-ci seront définies dans deux balises <lom:string> distinctes.

La balise <lom:string> contenant les informations de famille doit débiter par la ligne « GAR:Famille ». Si tel n'est pas le cas, les lignes suivantes seront ignorées.

#### Spécificité

Les lignes suivantes se présentent sous la forme « Propriété = valeur ». La valeur peut être vide. Les propriétés sont définies dans le tableau suivant :

Propriété	Description	Valeur attendue	Valeur par défaut
GAR:RessourceAffichable	Indique si la ressource liée est affichable	true ou false	False
GAR:MembresFamille	Identifiants des membres de la famille. La présence de cette propriété indique que la ressource est mère de famille	Liste d'identifiants séparés par un point-virgule (« ; »)	Liste vide
GAR:AttributsFamille	Attributs de la famille	Liste d'attributs séparés par un point-virgule (« ; »)	Liste vide
GAR:IdRessourcesAppelantes	Identifiants des ressources appelantes	Liste d'identifiants de ressource séparés par un point-virgule (« ; »)	Liste vide
GAR:IdRessourcesAppelées	Identifiants des ressources appelées	Liste d'identifiants de ressource séparés par un point-virgule (« ; »)	Liste vide

#### Notes :

- Les attributs de famille (propriété « GAR:AttributsFamille ») sont présents sous la forme « GAR:AttributsFamille = [IDC] Identifiant complémentaire : [LRA] Liste des ressources appelées ». Le module de moissonnage vérifie que tous les attributs famille mentionnés sont bien présents dans la table de référence « Attributs GAR famille » par extraction des codes des attributs placés entre crochets. Si ce n'est pas le cas, la notice est rejetée.
- L'attribut [IDC] ne peut figurer dans la liste des attributs famille que si la notice est une « mère de famille ». Dans le cas contraire, la notice est rejetée.
- Les propriétés du bloc famille doivent être uniques. En cas de doublon (par exemple 2 fois la propriété GAR:IdRessourcesAppelantes), la notice est rejetée.

- Tous les séparateurs peuvent être utilisés pour séparer la liste des Attributs GAR. La recherche des attributs se fait sur la détection des crochets ouvrant et fermant ([ ])

Exemple :

```
<lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI]
code établissement ---- [IDO] id opaque
</lom:string>
```

donnera une liste d'attributs [UAI, IDO], le libellé de l'attribut n'étant pas utilisé pour l'analyse.

## Exemple de Mère de famille

```
<scolomfr:extendedLocation>
  <scolomfr:location>https://www.edumedia-sciences.com/sso-gar/</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv3.0</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-
004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier
directement l'individu et/ou comportent des éléments de son état civil ou de sa vie
privée</scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-1/fr"> [Toute information de
description de l'accès spécifique à la ressource] </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI]
Code établissement ; [IDO] Id opaque</lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">
GAR:Famille
GAR:RessourceAffichable = true
GAR:MembresFamille = id1 ; id2 ; id3
GAR:AttributsFamille = [IDC] Identifiant complémentaire ; [LRA] Liste des ressources appelées
GAR:IdRessourcesAppelantes = id1
GAR:IdRessourcesAppellees = id2
    </lom:string>
  </lom:description>
</scolomfr:extendedLocation>
```

## Exemple de Membre d'une famille

```
<scolomfr:extendedLocation>
  <scolomfr:location>https://www.edumedia-sciences.com/sso-gar/</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv3.0</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier directement l'individu et/ou comportent des éléments de son état civil ou de sa vie privée</scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-1/fr"> [Toute information de description de l'accès spécifique à la ressource] </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Attributs GAR : [UAI] Code établissement ; [ID0] Id opaque </lom:string>
    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">
GAR:Famille
GAR:RessourceAffichable = false
GAR:AttributsFamille = [LRA] Liste des ressources appelées
GAR:IdRessourcesAppelantes = id1 ; id2
GAR:IdRessourcesAppellees = id2 ; id3
  </lom:string>
</lom:description>
</scolomfr:extendedLocation>
```

### 3.2.10.5 Ressources utilisateur web et web ayant une ou plusieurs variantes natives

Dans le cas où l'élément « platform » contient la valeur « <http://data.education.fr/gar/> », la notice est considérée de type utilisateur.

Les règles suivantes s'appliquent alors.

## Spécificité

### Règles pour la localisation : élément <scolomfr:location>

L'URL d'accès à la ressource est trouvée dans l'élément « 4.3.1.1 Localisation » (cf. document ADU9) pour l'accès à la ressource web et dans l'élément « 4.3.1.4 Description » pour l'accès à une variante native.

Règle 31 : Si la ressource est une ressource GAR WEB ou WEB/NATIVE alors l'élément « location » doit comporter une URL d'accès web dont la forme est valide sinon la notice est rejetée avec le message « l'élément « location » ne contient pas une URL web dont la forme est valide pour la notice [notice\_id\_oai] ».

### Règles pour la plateforme : élément <scolomfr:platform>

Pour une notice utilisateur, il doit y avoir un unique élément extendedLocation contenant un élément « platform » égal à « <http://data.education.fr/gar/> » dans la notice. Dans le cas contraire, la notice est rejetée avec un message d'erreur.

Il peut y avoir 0 à n éléments extendedLocation contenant un élément « platform » égal à « [http://data.education.fr/gar/oidc\\_native](http://data.education.fr/gar/oidc_native) ».

**Règles pour les données à caractère personnel : élément <scolomfr:personalDataProcessType>**

Les données à caractère personnel sont extraites de la valeur de l'élément « 4.3.1.3 Type des données à caractère personnel nécessaires au fonctionnement de la ressource » (cf. document ADU9). Pour les règles voir les « Règles pour la Description » ci-dessous.

**Règles pour la Description : élément <lom:description>**

Les attributs à transmettre au DTR lors de l'accès à la ressource sont encodés dans l'élément « 4.3.1.4 Description » (cf. document ADU9) sous la forme « Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque »

Règle 44 : Les attributs GAR « [UAI] Code établissement » et « [IDO] Id opaque » sont obligatoires dans l'élément « 4.3.1.4 Description ». S'ils ne sont pas présents, la notice est rejetée avec le message « Les attributs obligatoires [UAI] et [IDO] sont absents ».

Règle 45 : Pour une Ressource utilisateur web seule ou web + variantes natives, les données et attributs GAR doivent être présents dans l'élément « extendedLocation » correspondant à l'application web sinon la notice est rejetée avec le message « Les attributs GAR ne sont pas présents dans l'élément « extendedLocation » de l'application web pour la notice [notice\_id\_oai] ».

Règle 47 : Pour les ressources utilisateur avec une ou plusieurs variantes natives, les paramètres suivants seront définis dans la notice au sein de l'élément description (extendedLocation/description) dans un élément « string » sous forme d'une chaîne de caractères multi-lignes commençant par une première ligne contenant « GAR:OIDC\_Native » suivie, sur des lignes séparées, par :

- GAR:RedirectUri : un Uniform Ressource Identifier de redirection vers laquelle la réponse à la demande d'authentification doit être envoyée, c'est l'url d'accès à la ressource native et également url de Callback lors du processus d'authentification;
- GAR:ClientId : un identifiant unique au sein du GAR<sup>1</sup>. Le client id sera un UUID version 4 qui est représenté sous forme de 32 chiffres hexadécimaux en minuscules séparés par 4 tirets, exemple : 123e4567-e89b-12d3-a456-426614174000 soit 8Caractères-4caractères-4caractères-4caractères-12caractères, les groupes de caractères étant des nombres hexadécimaux 0123456789ABCDEF.
- GAR:ClientName : un nom reconnaissable identifiant l'application native.

Règle 50 : Si un paramètre est manquant la notice est rejetée avec le message indiquant le ou les paramètre(s) manquant(s) « 'paramètre1' et 'paramètre2' .... est / sont manquant(s) pour la notice [notice\_id\_oai] ».

Règle 51 : Si la valeur de l'un d'entre eux existe déjà dans la base de données du GAR la notice est rejetée avec le message indiquant le paramètre existant avec sa valeur « 'paramètre = valeur' est déjà présent dans le GAR pour la notice [notice\_id\_oai] ». Il y aura autant de messages que de paramètres avec une valeur existante.

Règle 52 : Si la valeur du paramètre GAR:ClientId n'est pas un UUID valide la notice est rejetée avec le message « GAR:ClientId n'est pas un UUID valide pour « [paramètre]=[valeur] » dans la notice [notice\_id\_oai] ». Il y aura autant de messages que de paramètres avec une valeur inexacte.

---

<sup>1</sup> Il est important de proposer au DTR un type d'identifiant unique, une solution est de fournir un uuid version 4

Règle 53 : Si la valeur du paramètre GAR:RedirectUri n'est pas une URI valide (selon la norme scheme+nom d'hôte) la notice est rejetée avec le message « GAR:RedirectUri » n'est pas un URI valide pour « [paramètre]=[valeur] » dans la notice [notice\_id\_oai] ». Il y aura autant de messages que de paramètres avec une valeur inexacte.

Règle 54 : Pour contrôler l'unicité des informations OIDC, si la valeur de GAR:ClientId ou GAR:RedirectUri existe déjà dans la table « plateforme » de la base de données du GAR, la notice est rejetée avec le message indiquant le paramètre existant avec sa valeur : « 'paramètre = valeur' est déjà présent dans le service de données d'initialisation du GAR pour la notice [notice\_id\_oai] ». Il y aura autant de messages que de paramètres avec une valeur existante.

## Exemple pour une ressource ayant des variantes natives

```

<scolomfr:extendedLocation>
  <scolomfr:location>https://acces.ressource.fr</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source uniqueElementName="source">SCOLOMFRv2.1</scolomfr:source>
    <scolomfr:value
uniqueElementName="value">http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-
004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier
directement l'utilisateur </scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description uniqueElementName="description">
    <lom:string>Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque ; [PRO] Profil ;
[NOM] Nom ; [PRE] Prénom </lom:string>
  </lom:description>
</scolomfr:extendedLocation>
<scolomfr:extendedLocation>
  <scolomfr:location>https://info.ressource.fr/android</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar/oidc_native</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv2.1</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-
004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier
directement l'utilisateur </scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description uniqueElementName="description">
    <lom:string>Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque ; [PRO] Profil ;
[NOM] Nom ; [PRE] Prénom </lom:string>
    <lom:string>GAR:OIDC_Native
GAR:RedirectUri = Deep Link _Ressource NATIVE 1
GAR:ClientId = unique_clientId_1
GAR:ClientName = unique_clientName_1</lom:string>
  </lom:description>
</scolomfr:extendedLocation>
<scolomfr:extendedLocation>
  <scolomfr:location>https://info.ressource.fr/ios</scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar/oidc_native</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source>SCOLOMFRv2.1</scolomfr:source>
    <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-
004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier
directement l'utilisateur </scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description uniqueElementName="description">
    <lom:string>Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque ; [PRO] Profil ;
[NOM] Nom ; [PRE] Prénom </lom:string>
    <lom:string>GAR:OIDC_Native
GAR:RedirectUri = Deep Link _Ressource NATIVE 2
GAR:ClientId = unique_clientId_2
GAR:ClientName = unique_clientName_2</lom:string>
  </lom:description>
</scolomfr:extendedLocation>

```

### 3.2.10.6 Ressources techniques communes (RTC)

Dans le cas où l'élément « platform » contient la valeur « <http://data.education.fr/gar/rtc> », la notice est considérée de type RTC.

Les règles suivantes s'appliquent alors.

#### Spécificité

Règle 46a : Pour une ressource technique commune (RTC), les attributs de catégorie 1 devraient être présents dans l'élément « extendedLocation » correspondant à la RTC.

L'absence d'un ou de plusieurs attributs de la catégorie 1, sera tracée dans les logs avec le message suivant : « [Règle 46a] Les attributs GAR : [listeAttributsManquants] ne sont pas présents dans l'élément « extendedLocation » de la RTC pour la notice [notice\_id\_oai] ».

La variable [listeAttributsManquants] contiendra la liste des attributs GAR de catégorie 1 manquants dans la notice.

Règle 46b : Pour une ressource technique commune (RTC), les attributs autres que ceux de la catégorie 1 ne devraient pas être présents dans l'élément « extendedLocation » correspondant à la RTC.

La présence d'un ou de plusieurs attributs GAR ne faisant pas partie de la catégorie 1 sera tracée dans les logs avec le message suivant : « [Règle 46b] Les attributs GAR : [listeAttributsSup] ne doivent pas être présents dans l'élément « extendedLocation » de la RTC pour la notice [notice\_id\_oai] ».

La variable [listeAttributsSup] contiendra la liste des attributs GAR qui ne sont pas de catégorie 1 présents dans la notice.

Ces deux règles 46a et 46b ne concernent que les logs. Aucun message n'est envoyé aux fournisseurs de ressources.

Dans tous les cas, seuls les attributs GAR de catégorie 1 seront enregistrés en base pour la RTC.

Règle 55 : Dans le cas d'une ressource technique commune (RTC), un seul élément « extendedLocation » GAR doit être présent sinon la notice est rejetée avec le message « Un seul élément « extendedLocation » doit être présent pour la RTC de la notice [notice\_id\_oai] ».

Règle 56 : Une ressource technique commune (RTC) ne peut pas être membre d'une famille. Si une ressource membre de famille est déclarée comme RTC, la notice est rejetée avec le message « Un membre de famille ne peut pas être une RTC de la notice [notice\_id\_oai] ».

NB : Une ressource technique commune (RTC) ne peut pas avoir des variantes applications natives. La règle 55 n'autorise pas d'autres éléments « extendedLocation » GAR que celui déclarant la RTC.

Règle 58 : Dans le cas d'une notice de ressource technique commune, le titre devrait commencer exactement par « [RTC] ». Dans le cas contraire, un message d'avertissement est envoyé au DTR et Gestionnaire d'entrepôt.

Une information est introduite dans les notices des ressources techniques communes pour lister les ressources utilisateur pouvant interagir avec elles. Le lien se base sur l'identifiant de la ressource, de façon similaire à ce qui est fait pour les ressources liées au sein des familles.



Règle 59 : L'élément « extendedLocation » d'une ressource technique commune (RTC) doit inclure la liste des ressources appelantes liées. A l'instar des ressources liées (mère de famille), une RTC doit contenir, dans un élément « extendedLocation/description/string », l'attribut **GAR:IdRessourcesAppelantes** identifiant la liste des identifiants des ressources appelantes, sinon la notice est rejetée.

L'attribut **GAR:IdRessourcesAppelantes** a la forme suivante : **GAR:IdRessourcesAppelantes = id1 ; id2 ; id3**

Les identifiants attendus sont les identifiants de ressources utilisateur.

Par exemple : **ark:/19153/b2n877zp**

Règle 60 : La propriété **GAR:IdRessourcesAppelantes** doit être présente une seule fois, sinon la notice est rejetée.

L'indexeur prend en compte l'extraction de l'attribut **GAR:IdRessourcesAppelantes** lors de l'indexation d'une ressource technique commune et l'analyse de la conformité de la chaîne de caractères au modèle proposé. Si la chaîne est correctement structurée, la liste des identifiants est extraite en préparation de l'injection en base.

Le module de moissonnage ne fait toutefois aucun contrôle sur la validité des informations passées pour les identifiants des ressources appelantes.

### Exemple d'extendedLocation pour une ressource technique commune (RTC)

```
<scolomfr:extendedLocation>
  <scolomfr:location></scolomfr:location>
  <scolomfr:platform>http://data.education.fr/gar/rtc</scolomfr:platform>
  <scolomfr:personalDataProcessType>
    <scolomfr:source uniqueElementName="source">SCOLOMFRv2.1</scolomfr:source>
    <scolomfr:value
uniqueElementName="value">http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044-num-
004</scolomfr:value>
    <scolomfr:label>personnalisation forte : les données transférées permettent d'identifier
directement l'utilisateur</scolomfr:label>
  </scolomfr:personalDataProcessType>
  <lom:description uniqueElementName="description">
    <lom:string>Attributs GAR : [UAI] Code établissement ; [IDO] Id opaque ; [PRO] Profil ;
[NOM] Nom ; [PRE]
    </lom:string>

    <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">
GAR:IdRessourcesAppelantes = id1 ; id2 ; id3
    </lom:string>
  </lom:description>
</scolomfr:extendedLocation>
```

#### 3.2.10.7 Indexation des Eléments ExtendedLocation

Les informations portées par les éléments « extendedLocation » dont l'élément « platform » a comme valeur un URI identifié (cf.3.2.10.1) pour les variantes natives des ressources sont stockés dans une table additionnelle « notice\_oidc » (voir annexe 0 Table « notice\_oidc »).

Le modèle de données est modifié en conséquence, une relation « un à plusieurs » est définie dans la notice, une notice pouvant avoir plusieurs éléments « extendedLocation ».

L'alimentation de `notice_url_acces` de la table « notice » se fera avec :

- l'URL de l'élément « *location* » appartenant à l'élément « *extendedLocation* » dont l'élément « *platform* » prend la valeur « `http://data.education.fr/gar` » des ressources utilisateur.
- Pour le cas des ressources techniques communes, l'élément « *location* » précédemment décrit n'est pas contrôlé (peut être vide ou renseigné) et une valeur par défaut est proposée dans la table.

Dans le cas d'une modification d'un élément « *extendedLocation* » existant ou l'ajout d'un élément « *extendedLocation* » des applications natives, le champ « `notice_oidc_updated` » de la table « notice\_oidc » sera positionné à « `true` », afin d'indiquer au Post-Moissonnage que l'élément a été modifié. Le Post-Moissonnage après traitement repositionnera le champ à « `false` ».

Dans le cas d'une suppression d'un élément « *extendedLocation* » préexistant, le champ « `notice_oidc_deleted` » de la table « notice\_oidc » sera positionné à « `true` », afin d'indiquer au Post-Moissonnage que l'élément a été supprimé. Le Post-Moissonnage après traitement supprimera la ligne de la base.

### 3.2.10.8 Vérification des codes DCP et des attributs GAR

Le module de moissonnage vérifie que tous les attributs mentionnés sont bien présents dans la table de référence « Classification par type de DCP et catégorie d'attributs GAR » par extraction des codes des attributs placés entre crochets. Si ce n'est pas le cas, la notice est rejetée.

Le module de moissonnage calcule le type de DCP maximum en fonction des attributs demandés. Cette valeur peut ensuite être utilisée par le GAR pour comparer le type de DCP déclaré dans la notice et celui calculé d'après les attributs demandés.

La liste des attributs autorisés est donnée en annexe.

Cas particulier des notices RTC : Les attributs GAR de catégorie 1 étant enregistrés dans la table notice, le calcul est également forcé à 1 (maximum de la catégorie des attributs autorisés).

### 3.2.10.9 Unicité de l'url d'accès à la ressource

L'url d'accès à la ressource doit être unique pour toutes les ressources utilisateur dans le GAR.

Le contrôle se fait sur le champ « `notice_url_acces` » dans la table « notice » pour tous les enregistrements dont le champ « `notice_rtc` » est positionné à « `false` ».

Si l'url existe pour un autre Identifiant d'enregistrement OAI [`notice_id_oai`], que la notice traitée, alors la notice est rejetée.

Dans le cas d'un rejet, la notification contient le message suivant :

« La notice a été rejetée car l'url d'accès [`notice_url_acces`] est déjà utilisée par la notice [`notice_id_oai`] »

## 3.2.11 Type pédagogique de la ressource

### 3.2.11.1 Descriptif de l'élément

Le type pédagogique de la ressource est extrait de l'élément « 5.2 Type pédagogique de la ressource » (cf. document ADU9), cet élément est associé au vocabulaire scolomfr-voc-010.

Cet élément est répétable, l'ensemble des valeurs est pris en compte.

### 3.2.11.2 Exemple

```
<lom:learningResourceType>
  <lom:source>SCOLOMFRv3.0</lom:source>
  <lom:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-010-num-001</lom:value>
  <lom:label>animation</lom:label>
</lom:learningResourceType>
```

## 3.2.12 Typologie du document

### 3.2.12.1 Descriptif de l'élément

La typologie du document est extraite de l'élément « 1.10 Typologie générale de documents » (cf. document ADU9), cet élément est associé au vocabulaire scolomfr-voc-005.

Cet élément est répétable, l'ensemble des valeurs est pris en compte.

### 3.2.12.2 Exemple

```
<scolomfr:generalResourceType>
  <scolomfr:source>SCOLOMFRv2.1</scolomfr:source>
  <scolomfr:value>http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-005-num-006</scolomfr:value>
  <scolomfr:label>banque de vidéos</scolomfr:label>
</scolomfr:generalResourceType>
```

## 3.2.13 Niveau éducatif détaillé

### 3.2.13.1 Descriptif de l'élément

Le niveau éducatif détaillé est extrait de l'élément « 9 Classification » (cf. document ADU9) lié au vocabulaire scolomfr-voc-022.

Cet élément est répétable, l'ensemble des valeurs est pris en compte.

Notes :

- la liaison de l'élément Classification au vocabulaire 22 est déterminée par la valeur de la balise <lom:value> de la balise <lom:purpose> :« http://data.education.fr/voc/scolomfr/concept/educational\_level »
- pour l'ensemble des éléments de classification, au sein de la balise <lom:taxon>, c'est la valeur de la balise <lom:id> qui détermine le concept à utiliser. La balise <lom:entry> associée est là à titre d'information et n'est pas prise en compte. Il n'est pas effectué de vérification de compatibilité entre l'identifiant du concept (porté par <lom:id>) et son terme préférentiel (porté par <lom:entry>)

### 3.2.13.2 Exemple

```
<lom:classification>
  <lom:purpose>
    <lom:source>SCOLOMFRv2.1</lom:source>
    <lom:value>Erreur ! Référence de lien hypertexte non valide.</lom:value>
    <lom:label>niveau éducatif détaillé</lom:label>
  </lom:purpose>
  <lom:taxonPath>
    <lom:source>
      <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">Niveau éducatif détaillé
- scolomfr-voc-022</lom:string>
    </lom:source>
    <lom:taxon>      <lom:id>Erreur ! Référence de lien hypertexte non valide.</lom:id>
    <lom:entry>
      <lom:string language="http://id.loc.gov/vocabulary/iso639-2/fre">école
élémentaire</lom:string>
    </lom:entry>
  </lom:taxon>
  ...
</lom:taxonPath>
</lom:classification>
```

## 3.2.14 Domaine d'enseignement

### 3.2.14.1 Descriptif de l'élément

Sur le même principe, le domaine d'enseignement est extrait de l'élément « 9 Classification » (cf. document ADU9) lié au vocabulaire scolomfr-voc-015.

Dans la version 4.0 de ScoLOMFr le domaine d'enseignement était issu d'un des cinq sous-vocabulaires créés pour remplacer le vocabulaire 015 : 015A, 015G, 015P, 015T, 015X.

Dans la version 5.0 de ScolomFr les vocabulaires 015G, 015P, 015T et 015X ont été fusionnés dans un seul nouveau vocabulaire 015GTPX. Deux nouveaux vocabulaires sont apparus, 0151 et 0152 correspondant respectivement aux « Points de programme » et aux « Compétences travaillées »

Cet élément est répétable, l'ensemble des valeurs est pris en compte.

Note : la liaison de l'élément Classification au vocabulaire 15 est déterminée par la valeur de la balise <lom:value> de la balise <lom:purpose>

### 3.2.14.2 Exemple

```
<lom:classification>
  <lom:purpose uniqueElementName="purpose">
    <lom:source uniqueElementName="source">SCOLOMFRv2.1</lom:source>
    <lom:value uniqueElementName="value">
      http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-028-num-003</lom:value>
    <lom:label>domaine d'enseignement</lom:label>
  </lom:purpose>
  <lom:taxonPath>
    <lom:source uniqueElementName="source">
      <lom:string/>
    </lom:source>
    <lom:taxon>
      <lom:id uniqueElementName="id">
        http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-015-num-461</lom:id>
      <lom:entry uniqueElementName="entry">
        <lom:string>économie gestion organisation du travail</lom:string>
      </lom:entry>
    </lom:taxon>
  </lom:taxonPath>
</lom:classification>
```

## 3.2.15 Type de présentation

### 3.2.15.1 Descriptif de l'élément

Cette information est extraite de l'élément « 9.3 Description » (cf. document ADU9) inclus dans une section Classification dont l'objectif associé est « label » (<http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-028-num-013>)

La notice est rejetée si cette information est absente ou si le code de type de présentation est inconnu.

Le préfixe attendu est « GAR\_Présentation : » ou « GAR\_Presentation : » (sans accent).

Note : seule la validité du code est vérifiée, le libellé est ignoré. Les codes autorisés sont listés en annexe.

### 3.2.15.2 Exemple

```
<lom:description uniqueElementName="description">
  <lom:string>GAR_Présentation : [MAN] manuels numériques</lom:string>
</lom:description>
```

Ici, le code « MAN » sera extrait et utilisé comme type de présentation.

```
<lom:description uniqueElementName="description">
  <lom:string>GAR_Presentation : [PRO] ressources de production pédagogique</lom:string>
</lom:description>
```

Ici, le code « PRO » est extrait et utilisé comme type de présentation.

## 3.2.16 Description de la notice

### 3.2.16.1 Descriptif de l'élément

L'élément « 1.4 description » (cf. document ADU9) contient le descriptif de la notice.

Cette information est extraite de l'élément « 1.4 description » (cf. document ADU9). Au sein de cet élément, les balises <lom:string> retenues sont celles respectant le premier critère rencontré dans cette liste :

- balise avec attribut @language = « http://id.loc.gov/vocabulary/iso639-2/fre »
- balise avec attribut @language = « http://id.loc.gov/vocabulary/iso639-2/fra »
- balise avec attribut @language = « fre »
- balise avec attribut @language = « fra »
- balise sans attribut @language

S'il existe plusieurs balises <lom:string> avec le même critère, leur contenu est concaténé en utilisant le séparateur « || ».

### 3.2.16.2 Exemple

Le traitement de l'extrait de code ScoLOMFR suivant :

```
<lom:description>
  <lom:string>Des mots et leur définition</lom:string>
  <lom:string language='http://id.loc.gov/vocabulary/iso639-2/fra'>Les entrées sont classées
  par ordre alphabétique</lom:string>
  <lom:string language='fre'>De nombreuses illustrations sont proposées</lom:string>
  <lom:string language='http://id.loc.gov/vocabulary/iso639-2/fra'>Tout trouver de A à
  Z</lom:string>
</lom:description>
```

aboutira à la chaîne : « Les entrées sont classées par ordre alphabétique| |Tout trouver de A à Z »

### 3.2.17 Remarques générales

Le dépassement du nombre autorisé de caractères de la valeur d'un élément de la notice (cf taille des champs en BDD du §4.2) entraînera le rejet de celle-ci avec le message d'erreur suivant :

« La taille de l'élément XXX dépasse la limite prévue (YYY caractères) ».

Le module de moissonnage est chargé du téléchargement des vignettes associées aux notices. Il est attendu que l'URL de la vignette se termine par une extension correspondant à son format. Les extensions acceptées sont les suivantes :

- .png pour les images au format PNG (Portable Network Graphics)
- .gif pour les images au format GIF (Graphics Interchange Format)
- .jpg ou .jpeg pour les images au format JPEG (Joint Photographic Experts Group)

En cas d'extension non reconnue ou d'absence d'extension, le type MIME de contenu (HTTP Content-type) de la réponse est examiné. Les correspondances suivantes sont alors appliquées entre types mime et format :

- image/png -> format PNG
- image/gif -> format GIF
- image/jpeg -> format JPEG

Si le type de la vignette ne peut être déterminé (par son extension ou par le type de contenu de la réponse HTTP), alors le module de moissonnage émet une notification au DTR de la notice et tente le téléchargement de la vignette.

S'il est impossible de télécharger la vignette (erreur réseau, ...), le module de moissonnage émet un avertissement non bloquant.

### 3.3 Stockage des données

Table	Description	Responsabilité	Utilisation
gar_dcp_attribut	Table de correspondance entre les attributs GAR, leur catégorie et le type de DCP associé	Administrateur GAR Le contenu de cette table est fourni par RENATER	Le module de moissonnage interroge cette table pour calculer le type de DCP associé aux attributs GAR demandés dans la notice ScoLOMFR => accès lecture seule
gar_type_presentation	Table de référence des codes et libellés des types de présentation dans le médiacentre.	Administrateur GAR Le contenu de cette table est fourni par RENATER. Les premières valeurs de référence sont indiquées en annexe	Le module de moissonnage interroge cette table pour vérifier la conformité du code de type de présentation. Cette table est également utilisée par le GAR pour obtenir le libellé associé à un code => accès lecture seule
entrepot_conf	Table de description des entrepôts à moissonner La structure de cette table est décrite en annexe.	Administrateur GAR Le contenu de cette table est mis à jour lors de l'ajout/suppression d'un entrepôt à moissonner. L'ajout et la suppression d'un entrepôt doivent aller de pair avec la mise en place et le retrait de la tâche cron de lancement de la moisson	Le module de moissonnage interroge cette table au lancement d'une moisson pour connaître les caractéristiques de l'entrepôt à moissonner. Il met à jour la date de dernière moisson en fin de traitement. => accès lecture seule sauf la colonne « harvest_last_date » qui doit pouvoir être modifiée
notice	Table des ressources moissonnées La structure de cette table est décrite en annexe.	Module de moissonnage Le contenu de cette table est mis à jour lors d'une moisson.	Stockage des ressources moissonnées => accès lecture/écriture

Table	Description	Responsabilité	Utilisation
notice_descripteur	Table des concepts ScoLOMFR associés aux ressources. La structure de cette table est décrite en annexe.	Module de moissonnage Le contenu de cette table est mis à jour lors d'une indexation des ressources moissonnées	Stockage des valeurs des concepts ScoLOMFR associés aux ressources => accès lecture/écriture
notice_vcard	Table des vcards associées aux ressources	Module de moissonnage Le contenu de cette table est mis à jour lors d'une indexation des ressources moissonnées	Stockage des valeurs des identifiants (SIREN + ISNI) et des vcards des éditeurs, des distributeurs commerciaux et techniques et du valideur GAR associés aux ressources et de la valeur de la plateforme pour les DTR (paramètre X-PLATFORME-ID) => accès lecture/écriture
notice_liaison	Table des liaisons entre ressources	Module de moissonnage Le contenu de cette table est mis à jour lors d'une indexation des ressources moissonnées	Stockage des liaisons entre ressources => accès lecture/écriture
type_id_scolomfr	Table des types d'identifiants autorisés	Administrateur GAR Le contenu de cette table est fourni par RENATER. Les premières valeurs de référence sont indiquées en annexe	Le module de moissonnage interroge cette table pour vérifier la conformité et trier les identifiants. => accès lecture seule
editeur	Table des SIREN des éditeurs ayant adhéré au cadre de confiance	Administrateur GAR Le contenu de cette table est fourni par RENATER.	Le module de moissonnage interroge cette table pour vérifier qu'un éditeur décrit dans une notice a bien adhéré au cadre de confiance.

### 3.4 Cloisonnement des données

Les données de moisson stockées en base de données sont reliées à l'entrepôt OAI d'origine par l'identifiant de l'entrepôt (entrepot\_id) présent dans chaque table.



## 3.5 Notifications

Au sein des modèles de notification présentés ci-dessous, les indications « HEADER » et « FOOTER » contiennent respectivement l'entête et le pied de page définis pour le GAR.

### 3.5.1 Bilan de la moisson

À chaque moisson, un courriel est envoyé au gestionnaire de l'entrepôt concerné pour l'informer du fonctionnement de la moisson. Ce courriel comprend les informations suivantes :

- URL de l'entrepôt
- Type de moisson (complète ou partielle)
- Date et heure de début et de fin de la moisson
- Statut de la moisson (OK / erreur)
- Nombre de notices moissonnées
- Erreurs rencontrées

#### 3.5.1.1 Format de la notification

Exemple de notification par courriel pour une moisson qui s'est correctement déroulée :

**Objet : GAR : rapport de moisson (succès)**

**Corps :**

{HEADER}

{avertissement contrôle date validite}

La moisson de votre entrepôt OAI s'est correctement déroulée.

La moisson {complète|partielle} de l'entrepôt {url} (set {set}) a débuté le {date/heure-débutmoisson} et s'est terminée le {date/heure-finmoisson}.

{n} enregistrements ont été moissonnés.

{FOOTER}

Et un courriel pour une moisson qui a comporté des erreurs :

**Objet : GAR : rapport de moisson (erreur(s))**

**Corps :**

{HEADER}

{avertissement contrôle date validite}

La moisson de votre entrepôt OAI a rencontré des erreurs.

La moisson {complète|partielle} de l'entrepôt {url} (set {set}) a débuté à {date/heure-débutmois} et s'est terminée à {date/heure-finmois}. Elle s'est correctement déroulée.

{n} enregistrements ont été moissonnés.

Les erreurs suivantes ont été rencontrées :

[erreur 1]

...

{erreur n}

{FOOTER}

Règle : gabarit (template) : dans les modèles de notification présentés plus haut, la chaîne « {avertissement contrôle date validite} » :

- n'apparaît pas lors que le contrôle de la date de validité GAR est actif (cas standard)
- est remplacée par la mention « Attention : le contrôle de la date de validité GAR est désactivé » si le contrôle de la date de validité GAR est désactivé pour cette moisson.

### 3.5.2 Erreurs et avertissements rencontrées dans les notices

Les erreurs rencontrées dans le contenu des notices sont envoyées au distributeur de la notice correspondante et/ou au gestionnaire de l'entrepôt.

Il en est de même pour certains avertissements.

Le courriel envoyé contient la liste des notices concernées avec les erreurs rencontrées (ainsi que les éventuels avertissements qui doivent être reportés).

### 3.5.2.1 Format de la notification

**Objet : GAR : erreurs dans les notices moissonnées**

**Corps :**

{HEADER}

Lors de la moisson de l'entrepôt OAI {url}, des erreurs ont été détectées dans les notices que vous distribuez.

Notices concernées :

```
- {notice1}
    {erreur1}
    {erreur2}
    ...
- {notice2}
    ...
{FOOTER}
```

Cette notification pourra être envoyée en copie au gestionnaire de l'entrepôt.

## 3.6 Données loguées

Au cours de son fonctionnement le module de moissonnage logue un certain nombre d'informations relatives aux traitements effectués. Ces informations sont consignées dans des fichiers logs dont le contenu est ensuite exploité par les outils de statistiques et d'alerte opérationnelle du GAR.

Les informations identifiées comme pertinentes sont décrites dans le tableau suivant :

Niveau	Message	Information(s) complémentaire(s)
Info	Lancement du module de moissonnage	Code de l'entrepôt à moissonner avec détails (URL, set, metadataPrefix, type de moisson)
Erreur	Erreur au lancement	Code de l'erreur + détail
Debug	Traitement de l'enregistrement	Identifiant OAI
Erreur	La notice n'est pas valide ScoLOMFRv[X.X]	Type d'identifiant et identifiant de la notice
Erreur	Information non trouvée	Détail
Avertissement	Information non trouvée	Détail
Erreur	La ressource est rejetée	Type d'identifiant et identifiant de la notice

Niveau	Message	Information(s) complémentaire(s)
Info	Moisson terminée	Statistiques de moisson : <ul style="list-style-type: none"> <li>• durée totale</li> <li>• nombre d'enregistrements collectés</li> <li>• nombre d'enregistrements rejetés</li> <li>• nombre d'enregistrements créés</li> <li>• nombre d'enregistrements mis à jour</li> <li>• nombre d'enregistrements supprimés</li> </ul>

Notes :

- Tous les logs sont horodatés et indiquent l'identifiant de l'entrepôt moissonné
- Chaque exécution du module de moissonnage crée un nouveau fichier de log nommé comme suit « moisson-[codeentrepôt]-[date-heure-début] »

## 3.7 Informations potentielles de niveau erreur ou avertissement

Cette section décrit les différentes informations, de niveau erreur ou avertissement, potentiellement rencontrées par le module de moissonnage. Ces informations sont identifiées par un code préfixé par « MM » (module de moissonnage) et un intitulé complémentaire. Elles sont toutes consignées dans le système de log du GAR. Les informations relatives à la moisson d'un entrepôt ou à l'exploitation des notices associées sont également reportées dans le courriel envoyé au gestionnaire de l'entrepôt en question.

### 3.7.1 Table des erreurs

Code	Libellé(s) associé(s)	Information(s) complémentaire(s)
MM-00 *	La variable d'environnement GARHARVEST_CONF est absente ou vide	
MM-01 *	Le fichier de configuration est introuvable ou illisible	Chemin du fichier de configuration
MM-02 *	Le contenu du fichier de configuration est incorrect	Chemin du fichier de configuration erreur détaillée
MM-03 *	Impossible de se connecter à la base de données	Chemin du fichier de configuration détail de l'erreur rencontrée
MM-04 *	Entrepôt introuvable	Code de l'entrepôt à moissonner
MM-10	Erreur réseau lors de la connexion à l'entrepôt	Code de l'entrepôt URL de requête Détail de l'erreur rencontrée
MM-11	Erreur XML lors de la moisson	Code de l'entrepôt URL de requête Détail de l'erreur rencontrée

Code	Libellé(s) associé(s)	Information(s) complémentaire(s)
MM-21	La notice n'est pas valide ScoLOMFR [X.X]	Code de l'entrepôt Identifiant OAI de la notice détail de l'erreur
MM-22	Donnée manquante	Code de l'entrepôt Identifiant OAI de la notice détail de l'erreur
MM-23	Impossible de télécharger la vignette éditeur	Code de l'entrepôt Identifiant de la notice URL de la vignette
MM-24	Attention : concept [...] inconnu !	Code de l'entrepôt Identifiant OAI de la notice Identifiant de la notice Titre de la notice Identifiant du concept incriminé
MM-25	la notice a été rejetée car l'url d'accès [notice_url_acces] est déjà utilisée par la notice [notice_id_oai]	Code de l'entrepôt Identifiant OAI de la notice Identifiant de la notice Titre de la notice Url d'accès de la notice Identifiant OAI de la notice utilisant l'url
MM-26	La notice a été rejetée car la VCard du DTR contient un champ X-PLATEFORME-ID dont la valeur n'est pas deux chiffres	Code de l'entrepôt Identifiant OAI de la notice Identifiant de la notice Titre de la notice
MM-27	La notice a été rejetée car la VCard du DTR contient plusieurs valeurs pour le champ X-PLATEFORME-ID	Code de l'entrepôt Identifiant OAI de la notice Identifiant de la notice Titre de la notice

\* les erreurs techniques MM-00 à MM-04 sont à destination de l'administrateur système de la plate-forme GAR. Elles ne sont pas transmises aux gestionnaires des entrepôts.

### 3.7.2 Table des erreurs et des avertissements concernant les extendedLocation

Code	Type	Condition non respectée	Message
<b>MM-36</b>	ERREUR	Présence d'un élément extendedLocation GAR	Localisation étendue GAR non trouvée dans la notice
<b>MM-43</b>	ERREUR	Présence soit d'un extendedLocation Web ou un extendedLocation RTC	La notice contient un « extendedLocation » Web et un extendedLocation RTC
<b>MM-38</b>	ERREUR	Uniquement 1 seul GAR WEB	Plusieurs localisations étendues GAR trouvées dans la notice
<b>MM-55</b>	ERREUR	Uniquement 1 seul GAR RTC	Il y a plusieurs éléments extendedLocation pour la RTC
<b>MM-42</b>	ERREUR	Si Native, gar web doit exister	La notice contient un « extendedLocation » Native sans extendedLocation Web

Code	Type	Condition non respectée	Message
<b>MM-37</b>	ERREUR	personalDataProcessType existe et contient URI valide voc-044	La valeur du concept pour le personalDataProcessType doit appartenir au vocabulaire 044 et doit commencer par <a href="http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044">http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-044</a>
<b>MM-45</b>	ERREUR	Pour une ressource web, la description existe et contient des attributs GAR valides	Les attributs GAR ne sont pas présents dans l'élément « extendedLocation » de l'application web
<b>MM-46</b>	ERREUR	Pour une RTC, la description existe et des contient des attributs GAR valides	Les attributs GAR ne sont pas présents dans l'élément « extendedLocation » de la RTC
<b>MM-31</b>	ERREUR	URI platform se termine par : /gar et location contient URL bien formé	L'élément location ne contient pas une URL web dont la forme est valide
<b>MM-28</b>	AVERTISS- EMENT	URI platform se termine par : /gar, et le titre contient « mère de famille »	La notice est acceptée cependant le titre de la notice ne respecte pas le format défini pour les mères de famille (le titre doit contenir « mère de famille »)
<b>MM-47</b>	ERREUR	URI platform se termine par : /gar/oidc_native et description contient GAR:OIDC_Native	L'attribut GAR:OIDC_Native est manquant ou erroné dans un ou plusieurs éléments extendedLocation des plateformes natives
<b>MM-58</b>	AVERTISS- EMENT	URI platform se termine par : /gar/rtc et le titre commence par « [RTC] »	La notice est acceptée cependant le titre de la notice ne respecte pas le format défini pour les RTC (le titre doit commencer par « [RTC] »)
<b>MM-59</b>	ERREUR	URI platform se termine par : /gar/rtc et description contient GAR:IdRessourcesAppelantes	Une RTC doit référencer les ressources appelantes. L'attribut GAR:IdRessourcesAppelantes doit être présent dans la description
<b>MM-60</b>	ERREUR	La propriété GAR:IdRessourcesAppelantes est présente plusieurs fois	Une RTC doit contenir une seule balise GAR:IdRessourcesAppelantes
<b>MM-50</b>	ERREUR	Un ou plusieurs paramètres OIDC sont manquants	Le ou les paramètres sont manquants dans un ou plusieurs extendedLocation natives
<b>MM-51</b>	ERREUR	un ou plusieurs paramètres existants avec la ou les valeurs de la base de données du GAR	Le ou les paramètres existent dans la base de données GAR
<b>MM-52</b>	ERREUR	GAR:ClientId est un UUID valide	GAR:ClientId n'est pas un UUID valide
<b>MM-53</b>	ERREUR	GAR:RedirectUri est une URI valide	GAR:RedirectUri n'est pas une URI valide
<b>MM-56</b>	ERREUR	Une RTC ne peut pas être membre d'une famille	Un membre de famille ne peut pas être une RTC de la notice

Code	Type	Condition non respectée	Message
<b>MM-40</b>	AVERTISS- EMENT	Le contenu de l'élément « platform » de chaque élément « extendedLocation » doit contenir une des valeurs http://data.education.fr/gar, http://data.education.fr/gar/ rtc., <a href="http://data.education.fr/gar/oidc_native">http://data.education.fr/gar/ oidc_native</a>	La notice contient au moins un « extendedLocation » ignoré car le contenu de l'élément « platform » ne contient pas une des valeurs http ://data.education.fr/gar, http ://data.education.fr/gar/rtc, http://data.education.fr/gar/oidc_native

*Note : Les messages sont référencés dans l'ordre des contrôles présentés sur le schéma du §3.2.10*

## 4 Annexe : modèle de données

Cette annexe décrit la structure des tables gérées par le module de moissonnage : une table pour décrire la configuration des entrepôts à moissonner et deux tables pour stocker les données moissonnées.

### 4.1 Configuration des entrepôts à moissonner

Avant de pouvoir moissonner un entrepôt, il est nécessaire de décrire sa configuration en base de données dans la table `entrepot_conf` dont la structure est décrite dans le tableau ci-dessous :

Colonne	Description	Type MySQL	Obligatoire
<code>entrepot_id</code>	Identifiant numérique de l'entrepôt	INT(11) NOT NULL AUTO_INCREMENT	Oui
<code>entrepot_code</code>	Code court de l'entrepôt. Ce code est celui utilisé pour identifier un entrepôt lors de l'appel au module de moissonnage. Ce code doit être unique, respecter le formatage suivant : [a-z][a-z0-9]+, et être significatif, ex : « coredu »	VARCHAR(50)	Oui
<code>entrepot_nom</code>	L'intitulé de l'entrepôt, ex : « Entrepôt Corédu »	VARCHAR(255)	Oui
<code>entrepot_url</code>	URL de l'entrepôt OAI, ex : « http://v3.coredu.fr/repository/OAIHandler »	VARCHAR(1024)	Oui
<code>entrepot_metadataprefix_oai</code>	Format des notices à moissonner (paramètre URL « metadataPrefix »), ex : « lom »	VARCHAR(255)	Oui
<code>entrepot_set_oai</code>	Filtre sur les notices (paramètre « set » optionnel)	VARCHAR(255)	Peut être vide
<code>entrepot_type_moissonnage</code>	Type de moisson. Valeurs autorisées : « FULL » : la moisson est complète et récupère l'ensemble des enregistrements de l'entrepôt « INCREMENTAL » : la moisson est partielle et récupère seulement les enregistrements mis à jour depuis la moisson précédente.	VARCHAR(30) DEFAULT « FULL »	Oui
<code>entrepot_date_moissonnage</code>	Date de début de la dernière moisson effectuée. Cette colonne est mise à jour à chaque moisson. Note : c'est l'unique colonne qui est mise à jour par le processus de moissonnage à chaque moisson.	DATETIME DEFAULT NULL	Est null avant la première moisson
<code>entrepot_courriel_gestionnaire</code>	Adresse de courriel du gestionnaire de l'entrepôt	VARCHAR(255)	Oui



Colonne	Description	Type MySQL	Obligatoire
entrepot_nom_gestionnaire	Nom du gestionnaire de l'entrepôt ((nom de l'équipe, « centre de support » ou un nom et un prénom)	VARCHAR(255)	Oui
entrepot_telephone_gestionnaire	Numéro de téléphone du gestionnaire de l'entrepôt	VARCHAR(20)	Oui
entrepot_description_	Description du gestionnaire de l'entrepôt	VARCHAR(255)	Non
entrepot_type_granularite	Type de granularité dans l'utilisation des dates pour l'entrepôt à moissonner. « DATE » pour les dates au format « AAAA-MM-JJ » et « DATETIME » pour les dates au format « AAAA-MM-JJThh:mm:ssZ »	VARCHAR(30) DEFAULT « DATE »	Oui
entrepot_statut_moissonnage	Vaut « PENDING » si une moisson de cet entrepôt est en cours et « FREE » dans le cas inverse.	VARCHAR(10) DEFAULT « FREE »	Oui

Notes :

- Il n'est pas prévu d'interface de saisie pour renseigner cette table lors de l'ajout d'un nouvel entrepôt
- Dans le cas où un entrepôt doit être moissonné avec 2 sets différents, il sera nécessaire de dupliquer la configuration en attribuant 2 valeurs de « entrepot\_code » distincts.

## 4.2 Données issues de la moisson

Les données issues de la moisson d'un entrepôt sont les notices ScoLOMFR ainsi que l'ensemble des données extraites de l'analyse des notices pour utilisation ultérieure. Ces données sont réparties en 3 tables :

- « notice » : le contenu de l'enregistrement OAI moissonné
- « notice\_descripteur » : les descripteurs de vocabulaire associés à une notice. Cette table est utilisée pour la correspondance avec les valeurs de la BCN.
- « notice\_vcard » : le contenu des vcards associées à une notice. Ces vcards sont catégorisées par profil : éditeur, distributeur technique, distributeur commercial et valideur technique.

Avertissement : la description des tables ci-après indique la taille maximale réservée à chaque colonne (ex : il est prévu 1024 caractères pour la colonne « notice\_id » qui contient l'identifiant de la notice). Ces tailles ont été choisies en fonction des données à stocker et devraient suffire dans la plupart des cas. Le moissonneur vérifie la taille de chaque donnée avant de stocker cette dernière en base de données et rejette la notice en cas de dépassement.

### 4.2.1 Table « notice »

La table principale de sauvegarde des enregistrements OAI. Le tableau ci-dessous décrit sa structure. La clef primaire de cette table positionnée sur la colonne « notice\_id\_interne » ; le couple « entrepot\_id », « oai\_id » est une clé primaire logique.

Colonne	Description	Type MySQL	Obligatoire
notice_id_interne	Identifiant interne de l'enregistrement dans la base de données	INT(11)	Oui
notice_id_entrepot	Identifiant numérique de l'entrepôt, lien vers la table repository_conf	INT(11)	Oui
notice_id_oai	Identifiant de l'enregistrement OAI	VARCHAR (1024)	Oui
notice_date_modification	Date de l'enregistrement OAI	DATETIME	Oui
notice_supprime	Indique si l'enregistrement est noté comme « supprimé »	BOOLEAN	Oui
notice_statut	Indicateur si la ressource a été supprimée ou recrée Valeurs possibles : à traiter – supprimée : DELETED à traiter – recréé : CREATED traité : DONE manquante : MISSING Valeur par défaut : CREATED	VARCHAR (20)	Oui
notice_contenu	Contenu XML de la notice La notice XML est stockée dans la BDD pour des interrogations par les administrateurs : ex données de contact pour le support, données nécessaires pour les contrôles a posteriori du respect des règles des accords contractuels des partenaires avec le MEN (fédération), etc.	MEDIUMTEXT	Peut être vide si l'enregistrement est supprimé
notice_type_id	Type d'identifiant de la notice, ex : « ARK »	VARCHAR (50)	Oui
notice_id	Identifiant de la notice ScoLOMFR, ex : « ark:/19153/b2n877zp »	VARCHAR (1024)	Oui
notice_titre	Titre de la notice ScoLOMFR, ex : « Maxicours:Principes fondamentaux de l'Économie et de la Gestion pour la 2nde »	VARCHAR (255)	Oui
notice_url_image	URL de l'illustration associée à la notice, ex : « http://art.rmngp.fr/profiles/rmn/themes/custom/rmn_art/images/footer/agence_photo.png »	VARCHAR (1024)	Non mais recommandé
notice_label_gar	Label GAR. Cette colonne est à true si la notice contient le label GAR	BOOLEAN	Oui
notice_date_validation_gar	Date de dernière validation GAR, ex : « 2016-10-07 »	DATE	Vide si la notice n'est pas labellisée GAR

Colonne	Description	Type MySQL	Obligatoire
notice_type_presentation_code_gar	Code du type de présentation pour la notice, ex : « MAN » pour « [MAN] manuels numériques ». Le libellé associé au code peut être retrouvé via la table « gar_display_presentation »	VARCHAR (10)	Oui
notice_url_acces	URL d'accès à la ressource, ex : « https://fresques.ina.fr/jalons/ » Dans le cas d'une RTC, la valeur par défaut suivante sera utilisée et non contrôlée sur l'unicité : http://data.education.fr/gar/rtc	VARCHAR (1024)	Oui
notice_attributs_dcp	Les attributs à transmettre au DTR lors de l'accès à la ressource. Les noms d'attributs sont séparés par des «   », ex : « IDO UAI E_MS3 »	VARCHAR (1024)	La chaîne peut être vide
notice_categorie_gar_calcule	Calcul de la catégorie GAR associée à la notice, à partir de la liste des attributs demandés (catégorie maximale demandée)	TINYINT	Oui
notice_dcp_modifie	Colonne permettant d'indiquer que les DCP ont été modifiés. Le module de moissonnage passe cette colonne à VRAI lorsque la liste des attributs DCP est modifiée. Il ne passe jamais cette valeur à FAUX. Le module de gestion DCP passe cette colonne à FAUX lorsque la modification est prise en compte.	BOOLEAN	Oui
notice_chemin_vignette	Chemin relatif (au dossier racine des vignettes) de la vignette GAR de la notice. Cette information est à destination des modules du GAR qui souhaitent associer une illustration à une ressource	VARCHAR (1024)	Non
notice_description	Contenu texte de l'élément ScoLOMFR « 1.4 Description » constitué de la concaténation du contenu des balises <lom:string> dont la balise @language vaut « http://id.loc.gov/vocabulary/iso639-2/fre » ou « http://id.loc.gov/vocabulary/iso639-2/fra ». Le séparateur utilisé pour la concaténation est «    ».	TEXT	Non, chaîne vide par défaut

Colonne	Description	Type MySQL	Obligatoire
notice_liee	Indique si la ressource est liée (true) ou autonome (false). La valeur des colonnes « notice_affichable », « notice_mere_famille », « notice_attributs_famille », « notice_bloc_famille », « notice_famille_modifiee » n'est à prendre en compte que si la colonne « notice_liee » vaut true	BOOLEAN DEFAULT 0	Oui, faux par défaut (= ressource autonome)
notice_affichable	Indique si la ressource liée est affichable.	BOOLEAN DEFAULT 0	Oui, faux par défaut
notice_mere_famille	Indique si la ressource liée est « mère de famille ».	BOOLEAN DEFAULT 0	Oui, faux par défaut
notice_attributs_famille	Les attributs DCP de type famille. Les noms d'attributs sont séparés par des «   », ex : « IDC LRA »	VARCHAR (1024)	La chaîne peut être vide
notice_bloc_famille	Contenu du bloc de texte commençant par « GAR:Famille » (pour permettre de détecter les changements d'une moisson sur l'autre)	MEDIUMTEXT DEFAULT « »	La chaîne peut être vide
notice_famille_modifiee	Indique si le bloc famille a été modifié depuis la dernière moisson. Le module de moissonnage passe cette colonne à VRAI lorsque le contenu du bloc famille a été modifié. Il ne passe jamais cette valeur à FAUX. Le module de post-moissonnage passe la valeur de cette colonne à FAUX lorsque la modification est prise en compte.	BOOLEAN DEFAULT 0	Oui, faux par défaut
notice_rtc	Indique si c'est une ressource technique commune (RTC)	BOOLEAN DEFAULT 0	Oui, faux par défaut

## Notes :

- Dans le cas d'une notice mère de famille, le titre et par conséquent le champ titre\_id doivent contenir exactement « mère de famille ».
- Dans le cas d'une notice d'une Ressource Technique Commune (RTC), le titre et par conséquent le champ titre\_id doivent commencer exactement par « [RTC] » (cf. Règle 58 ).
- La colonne notice\_contenu contient l'intégralité de la notice ScoLOMFR. Disposer de ce contenu est précieux en phase de test pour s'assurer du bon fonctionnement de l'analyse et l'extraction des données et pouvoir diagnostiquer simplement certaines erreurs. Lorsque le fonctionnement sera stabilisé, on pourra se contenter du stockage des éléments issus de l'analyse de la notice pour alléger le poids de la table « notice ».

- Dans le cas où une notice est supprimée (colonne « notice\_supprime » à true), seules les informations contenues dans les colonnes « notice\_id », « notice\_date\_modification », « notice\_supprime » et « notice\_contenu » seront utilisables.
- Dans le cas où une notice est rejetée, aucune mise à jour des tables « notice » et « notice\_\* » n'est effectuée.

## 4.2.2 Table « notice\_oidc »

La table notice\_oidc permet de stocker les éléments liés au protocole OIDC qui sont dans le ou les éléments <extendedLocation/> de la notice.

Colonne	Description	Type MySQL	Obligatoire
notice_oidc_id_interne	Identifiant interne de l'enregistrement dans la base de données	INT(11)	Oui
notice_oidc_id_notice	Identifiant de la notice associée	INT(11)	Oui
notice_oidc_client_id	Identifiant unique de type uuid	TBD	Oui
notice_oidc_redirect_uri	Url de redirection (callback)	VARCHAR(1024)	Oui
notice_oidc_client_name	Nom du client	VARCHAR(1024)	Oui
notice_oidc_updated	Indique que l'élément extendedLocation a été modifié	BOOLEAN DEFAULT 0	Oui, faux par défaut
notice_oidc_deleted	Indique que l'élément extendedLocation a été supprimé	BOOLEAN DEFAULT 0	Oui, faux par défaut

Notes :

- Dans le cas où une notice est rejetée, aucune mise à jour (création ou modification) de cette table n'est effectuée.

### 4.2.3 Table « notice\_descripteur »

Colonne	Description	Type MySQL	Obligatoire
notice_desc_id_interne	Identifiant interne de la relation entre une notice et une valeur de descripteur	INT(11)	Oui
notice_desc_id_notice	Identifiant de la notice associée	INT(11)	Oui
notice_desc_id_element	Identifiant de l'élément ScoLOMFR concerné, ex : « 9.1 » (cf. document ADU9)	VARCHAR(10)	Oui
notice_desc_id_voc	Identifiant du vocabulaire ScoLOMFR associé à l'élément. Ex : « 040 »	VARCHAR(10)	Oui
notice_desc_uri_concept	Identifiant URI de concept, ex : « http://data.education.fr/voc/scolomfr/concept/scolomfr-voc-040-num-014 »	VARCHAR(512)	Oui
notice_desc_term_pref_concept	Terme préférentiel du concept, ex : « culture scientifique et technologique »	VARCHAR(512)	Oui

#### 4.2.4 Table « notice\_vcard »

Cette table contient les informations de Vcard extraites de chaque notice. La structure de cette table est définie ci-dessous :

Colonne	Description	Type MySQL	Obligatoire
noticevcard_id_interne	Identifiant interne de la relation de la notice avec la vcard	INT(11)	Oui
noticevcard_id_notice	Identifiant de la notice associée	INT(11)	Oui
noticevcard_id_contact	Identifiant du contact extrait de la VCard, composé ainsi : [SIREN]_[ISNI]. Le SIREN est obligatoire et tronqué à 9 caractères. L'ISNI est remplacé par des 0 s'il est absent de la Vcard Ex : « 354786945_354122 »	CHAR(26)	Oui
noticevcard_vcard_contact	Vcard telle qu'extrait de la notice	VARCHAR(1024)	Oui
noticevcard_role_contact	Code du rôle correspondant à la Vcard extraite : « EDI » pour l'éditeur « DTR » pour le distributeur technique « DCO » pour le distributeur commercial « VAT » pour le valideur technique	CHAR(3)	Oui
noticevcard_courriel_contact	Adresse de courriel extraite de la Vcard	VARCHAR(255)	Non
noticevcard_fn_contact	Nom complet du contact extrait de la VCard (attribut « FN »)	VARCHAR(255)	Non
noticevcard_org_contact	Nom de l'organisation du contact extrait de la VCard (attribut « ORG »)	VARCHAR(255)	Non
noticevcard_plateforme_id	Nom de la plateforme extrait de la Vcard (attribut « X-PLATEFORME-ID »). Ce champ n'est fourni que dans la VCARD du DTR	VARCHAR(2)	Non



## 4.2.5 Table « notice\_liaison »

Cette table contient les informations relatives aux liens entre notices. Ces liaisons peuvent être de 4 types :

- membre de famille
- ressource appelante
- ressource appelée
- ressource appelée de type RTC

Colonne	Description	Type MySQL	Obligatoire
notice_liaison_id_interne	Identifiant interne d'enregistrement dans la base de données	INT(11)	Oui
notice_liaison_id_notice	Identifiant de la notice associée	INT(11)	Oui
notice_liaison_type	Type d Type de liaison : - « membre » pour les membres de la famille - « appelante » pour les ressources appelantes -« appelee » pour les ressources appelées - « rtc » pour les ressources appelées de type RTC	VARCHAR(9)	Oui
notice_liaison_id_lie	Identifiant de la ressource liée, ex : « ark://9999/w8zjdll1 »	VARCHAR(1024)	Oui

Note : Précision pour le cas des ressources techniques communes (RTC), il est attendu l'id interne de la ressource technique commune (RTC) dans le champ « notice\_liaison\_id\_notice » et l'id ARK pour la ressource appelante dans le champ « notice\_liaison\_id\_lie ».

## 5 Annexe : GAR – Type de présentation

Ce tableau présente les valeurs d'initialisation de la table « gar\_type\_presentation ». Celles-ci sont extraites du document de référence « SpecifDoc\_V14\_30092016.docx »

Code	Libellé
ACC	ressources d'entraînement et d'accompagnement scolaire
DIC	ressources de référence, dictionnaires et encyclopédies
DOC	ressources documentaires et de presse
MAN	manuels numériques
MUL	ressources d'enseignement multimédias
ORI	ressources d'orientation
PRO	ressources de production pédagogique

## 6 Annexe : GAR – Liste des attributs valides

Ce tableau présente les attributs valides stockés dans la table « gar\_dcp\_attribut ». Ces données sont extraites du document de référence « SpecifDoc\_V14\_30092016.docx »

Code	Libellé	Catégorie attributs GAR	Type DCP	Description
UAI	Code établissement	1	3	UAI (code d'Unité Administrative Identifiée)
IDO	Id opaque	1	3	Identifiant de l'utilisateur spécifique à la ressource, garantissant l'absence de retour vers les informations personnelles
PRO	Profil	2	3	Identifiant(s) du ou des profil(s) de l'accédant (PROFIL_NATIONAL : National_elv, National_ens, National_doc,...)
DIV	Division(s)	3	4	Code de la ou des divisions, libellé de la ou des divisions
GRO	Groupe(s)	3	4	Code du ou des groupes, libellé du ou des groupes, divisions d'appartenance des groupes
E_MS1	Degré d'enseignement (élève)	3	4	Code du (des) degrés d'enseignement (N_MEF_STAT_1)
E_MS2	Cycle de scolarité (élève)	3	4	Code du ou des cycle(s) de scolarité (N_MEF_STAT_2)
E_MS3	Dispositif de formation (élève)	3	4	Code du ou des dispositifs de formation (N_MEF_STAT_3)
E_MS4	Niveau de formation (élève)	3	4	Code du ou des niveaux de formation (N_MEF_STAT_4)
E_MS5	Filière (élève)	3	4	Code de la ou des filière(s) (N_MEF_STAT_5)
E_MAT	Enseignements suivis (élève)	3	4	Code des enseignements suivis tels qu'ils sont dans les référentiels sources + libellé
P_MAT	Matières enseignées (enseignant)	3	4	Codes des matières enseignées tels qu'ils sont dans les référentiels source + libellés
P_MEL	Adresse mel (enseignant)	4	4	Adresse courriel

Code	Libellé	Catégorie attributs GAR	Type DCP	Description
P_MS1	Degré(s) d'enseignement (enseignant)	3	4	Code(s) du (des) degrés d'enseignement (N_MEF_STAT_1)
P_MS2	Cycle(s) de scolarité (enseignant)	3	4	Code(s) du ou des cycle(s) de scolarité (N_MEF_STAT_2)
P_MS3	Dispositif(s) de formation (enseignant)	3	4	Code(s) du ou des dispositifs de formation (N_MEF_STAT_3)
P_MS4	Niveau(x) de formation (enseignant)	3	4	Code(s) du ou des niveaux de formation (N_MEF_STAT_4)
P_MS5	Filière(s) (enseignant)	3	4	Code(s) du ou des filière(s) (N_MEF_STAT_5)
CIV	Titre de civilité	4	4	M., Mme
NOM	Nom d'usage	4	4	Patronyme
PRE	Prénom usuel	4	4	Prénom
idENT	Code projet ENT	1	3	Code projet ENT encodé en base 64

Ce tableau présente les attributs Famille valides stockés dans la table « gar\_dcp\_attribut\_famille ». Ces données sont issues de la concertation pour la mise en place de l'évolution « Id unique, id complémentaire ».

Code	Libellé	Catégorie attributs GAR	Type DCP	Description
IDC	Identifiant complémentaire	4	4	
LRA	Liste des ressources appelées	2	2	Liste des autres ressources de la famille appelées par la ressource et qui sont affectées à l'accédant.

## 7 Annexe : GAR – Liste des types d'identifiants autorisés

Ce tableau présente les types d'identifiant autorisés stockés dans la table «type\_id\_scolomfr ». Ces données sont extraites du document de référence « Moissonnage-identifiants-V2\_29092016.docx ».

Rang	Nom	Description
1	ARK	Archival Resource Key; URL designed to support long - term access to information objects. In general, ARK syntax is of the form (brackets indicate [optional] elements: [ http: // NMA/ ] ark: /NAAN/Name
2	ISBN	International Standard Book Number; a unique numeric book identifier. There are 2 formats: a 10 - digit ISBN format and a 13 - digit ISBN.
3	DOI	Digital Object Identifier; a character string used to uniquely identify an object. A DOI name is divided into two parts, a prefix and a suffix, separated by a slash.
4	PURL	Persistent Uniform Resource Locator. A PURL has three parts: (1) a protocol , (2) a resolver address , and (3) a name.

Note : Dans les exploitations actuelles, seul le type d'identifiant ARK est autorisé. Toutes les autres formes d'ID sont à déprécier.

## 8 Annexe : GAR – Codes SIREN des éditeurs ayant adhéré au cadre de confiance

La table « editeur » contient les colonnes suivantes :

- editeur\_siren : numéro SIREN de l'éditeur
- editeur\_libelle : Libellé éditeur
- dtr\_libelle : Nom du distributeur technique de la ressource

Le module de moissonnage n'interroge que la première colonne (n° SIREN).

Il n'y a pas de données de référence pour cette table.

Cette dernière sera alimentée par l'administrateur du GAR au fur et à mesure de l'adhésion des éditeurs au cadre de confiance du GAR.

Les numéros SIREN stockés dans la colonne "editeur\_siren" ne doivent pas contenir d'espaces.

## 9 Annexe : Script de standardisation des vignettes

### 9.1 Préambule

Le script de normalisation est lancé avec un code d'entrepôt en paramètre, déterminant ainsi l'ensemble des images à traiter.

Il accède au fichier de configuration du module de moissonnage pour déterminer l'emplacement du dossier des images originales à traiter et les dossiers destinations contenant les images standardisées (GAR et MÈRE DE FAMILLE), prêtes à être utilisées par les modules du GAR.

Le traitement est appliqué selon les règles suivantes :

- Parcours du dossier des images originales, pour chaque image :
  - si l'image standardisée correspondante n'existe pas, celle-ci est créée.
  - si l'image originale est plus récente que l'image standardisée, l'image standardisée est recrée.
- Parcours du dossier des images standardisées, si une image standardisée ne correspond à aucune image originale, l'image standardisée est supprimée.

Chaque traitement de dossier lié à un entrepôt est consigné dans un fichier de logs nommé comme suit « normalisation-[codeentrepot]-[date-heure-debut].log » dans le dossier de log défini pour le module de moissonnage.

Ce fichier de log référencera l'ensemble des vignettes traitées en indiquant pour chacune si le traitement a été effectué avec succès ou si une erreur s'est produite. En cas d'erreur sur le traitement d'une vignette, le script de standardisation envoie un courriel de notification au DTR de la notice (l'adresse de courriel du DTR est récupérée de la base de données).

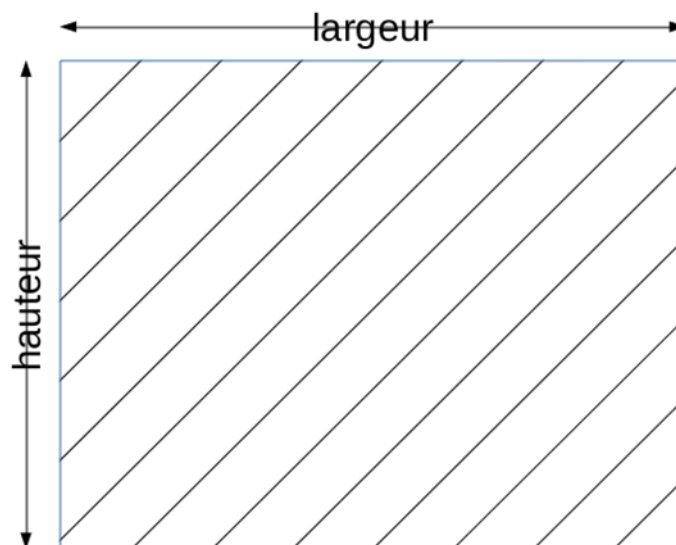
### 9.2 Traitement à apporter à chaque image

La transformation de la vignette éditeur en vignette pour usage par le GAR est constituée de plusieurs étapes :

- Création du fond de la vignette pour usage par le GAR
- Mise à l'échelle de la vignette éditeur
- Intégration de la vignette éditeur retaillée
- Intégration de la pastille GAR et si nécessaire de la pastille MÈRE DE FAMILLE
- Génération du fichier PNG

#### 9.2.1 Création du fond de la vignette pour usage par le GAR

Les vignettes à générer sont une composition de la vignette éditeur du titre, de la pastille GAR et si nécessaire de la pastille MÈRE DE FAMILLE. La première étape est la création de l'image transparente qui va accueillir ces deux éléments.



*Le fond transparent.*

Paramètre: largeur x hauteur en pixels. (cf. directive « `img.thn.size` » dans la section « Paramètres techniques »)

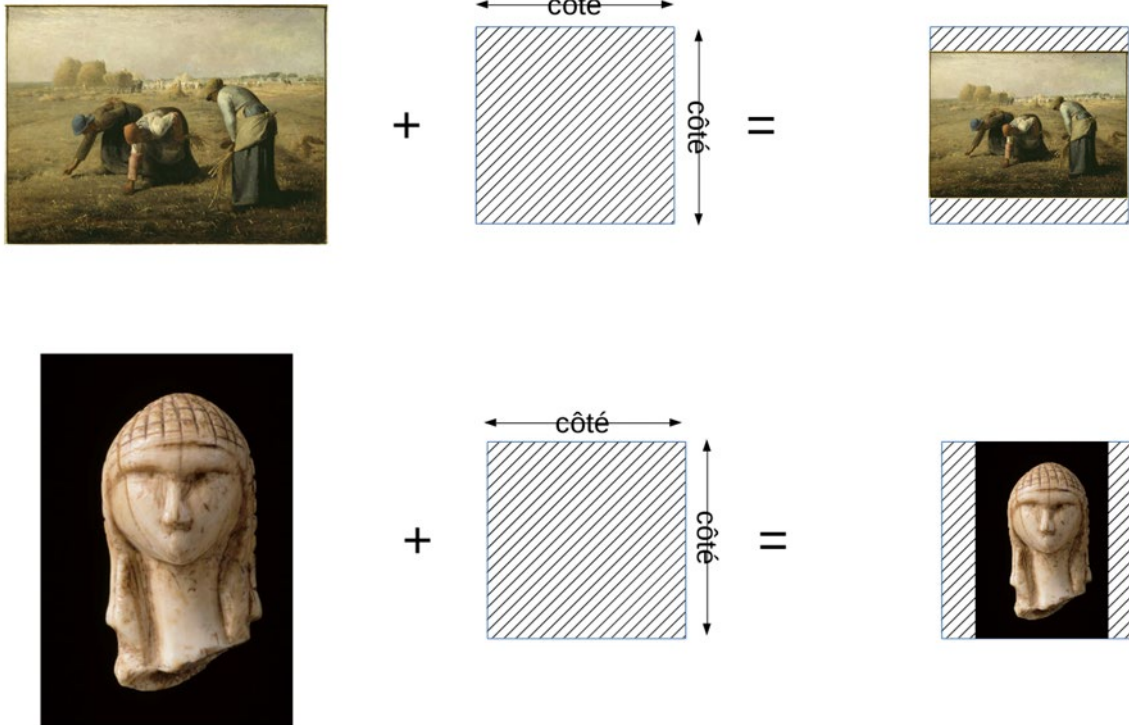
Paramètre par défaut : 310x250

### 9.2.2 Mise à l'échelle de la vignette éditeur

La taille de la vignette éditeur décrite dans la notice ScoLOMFR peut varier. Il est nécessaire de mettre à l'échelle cette vignette pour permettre son insertion harmonieuse au sein des vignettes pour usage par le GAR. Cette mise à l'échelle s'effectue en respectant les règles suivantes :

- l'opération de mise à l'échelle conserve le ratio largeur x hauteur de l'image originale
- l'image mise à l'échelle doit tenir dans un carré de n pixels de côté
  - si l'image originale est en mode portrait, l'image retaillée verra sa hauteur redimensionnée à n pixels et sera centrée horizontalement
  - si l'image originale est en mode paysage, l'image retaillée verra sa largeur redimensionnée à n pixels et sera centrée verticalement
  - si l'image originale est carrée, elle remplira complètement le gabarit fixé.





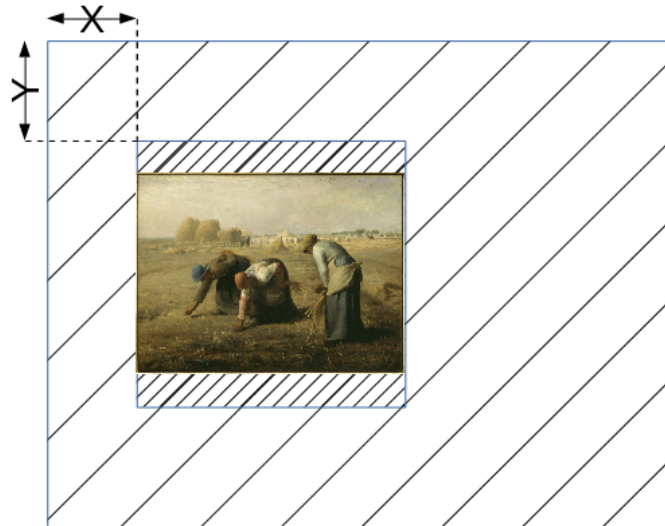
*Mode de redimensionnement des vignettes éditeurs dans la boîte englobante (grisée). Images provenant du site « Images d'art » (<http://art.rmngp.fr/fr>).*

Paramètre : taille en pixels du côté du carré englobant la vignette éditeur. (cf. directive « `img.thn.bb.size` » dans la section « Paramètres techniques »)

Paramètre par défaut : 250x250

### 9.2.3 Intégration de la vignette éditeur retournée

Le carré contenant la vignette éditeur mise à l'échelle s'intègre ensuite dans le fond de la vignette pour usage par le GAR à un emplacement déterminé par un point x, y (en pixels) à partir du point haut gauche du fond

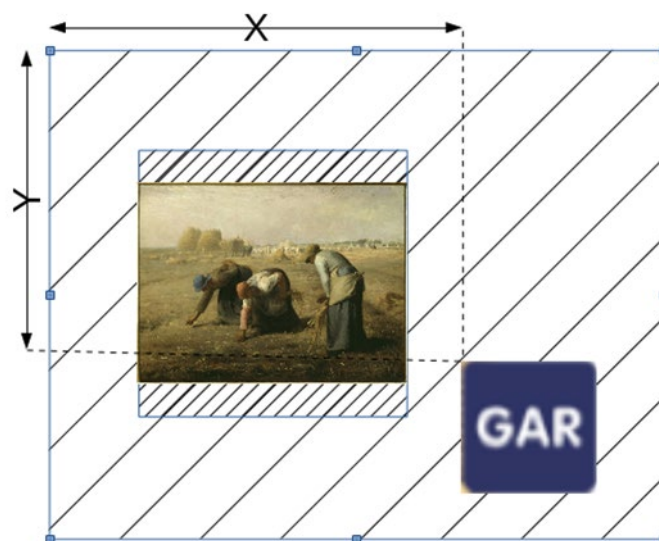


Paramètre : positionnement x, y en pixels du coin haut gauche du carré contenant la vignette éditeur mise à l'échelle. (cf. directive « `img.thn.bb.pos` » dans la section « Paramètres techniques »)

Paramètre par défaut : +0x0

#### 9.2.4 Intégration de la pastille GAR et de la pastille MÈRE DE FAMILLE

La pastille GAR est une image externe fournie par le GAR. Elle est utilisée telle quelle (sans remise à l'échelle) et positionnée dans le fond contenant la vignette éditeur retaillée. Elle est apposée pour toutes les vignettes utilisées par le GAR. En fonction de la position choisie pour la pastille GAR, celle-ci peut se trouver en superposition avec la vignette éditeur retaillée.



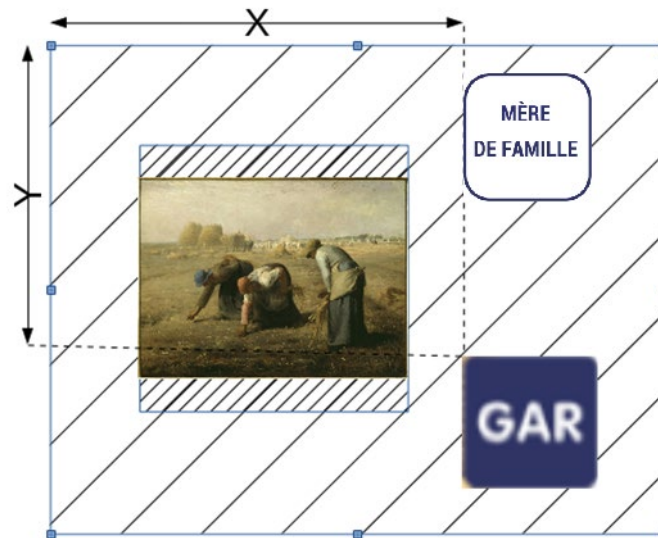
Paramètre : positionnement x,y en pixels du coin haut gauche de la pastille GAR. (cf. directive « `img.label.pos` » dans la section « Paramètres techniques »)

Paramètre par défaut : +250x160

La pastille MÈRE DE FAMILLE est une image externe fournie par le GAR. Elle est apposée en haut à droite (voir paramètre pour les coordonnées de positionnement exact) lorsque la notice est une notice « mère de famille »

Elle est utilisée telle quelle (sans remise à l'échelle) et positionnée dans le fond contenant la vignette éditeur retaillée.

La dimension préconisée est 60 sur 60.



Paramètre : positionnement x,y en pixels du coin haut gauche de la pastille MÈRE DE FAMILLE. (cf. directive « img.mere.pos » dans la section « Paramètres techniques »)

Paramètre par défaut : +250x30

Note : si le positionnement choisi pour la pastille GAR la place hors du cadre défini pour le fond, celle-ci ne sera pas visible. De même, si la pastille GAR dépasse du cadre, celle-ci sera automatiquement tronquée. Idem pour la pastille MÈRE DE FAMILLE.

### 9.2.5 Génération du fichier PNG

L'image composée est alors à générer au format Portable Network Graphics (PNG) dans les dimensions précisées à la première étape.

Note : l'ensemble des traitements nécessaires sera réalisé avec l'outil libre en ligne de commande « GraphicsMagick » piloté par un script Python pour le parcours des dossiers et le paramétrage des opérations graphiques.

## 9.3 Gestion des erreurs

La normalisation d'une vignette éditeur peut échouer pour différentes raisons :

- format de la vignette éditeur non reconnu
- fichier original corrompu

- problème d'entrée/sortie (disque plein, dossier de génération interdit en écriture...)
- ...

Dans ce cas, le script de normalisation émet une erreur dans les logs, envoie un mail au DTR de la notice ou à défaut au gestionnaire de l'entrepôt et passe à la vignette suivante. Ainsi, les vignettes GAR ou MÈRE DE FAMILLE générées par un traitement précédent ne sont pas altérées.

En cas de problème constaté sur les vignettes générées, il est toujours possible de consulter les logs pour diagnostiquer le problème et le résoudre.

## 9.4 Lancement du script

Le script de standardisation des vignettes se présente sous la forme d'un script Python nommé « standardisation.py ». Il peut être lancé en ligne de commande, via un script shell ou par une tâche cron. Il nécessite l'installation de Python 3 et la définition d'une variable d'environnement (cf. §3.1.1).

Le lancement du module s'effectue comme suit :

```
standardisation.py ( --all | codeentrepot ) [ --config fichierdeconfiguration ]
```

Exécuté sans paramètre, le script affiche la notice d'usage du script.

Pour traiter les vignettes de l'ensemble des entrepôts, on utilisera la commande :

```
standardisation.py --all
```

Pour traiter les vignettes d'un entrepôt en particulier, on précise son code, ex :

```
standardisation.py valenciennes
```

Par défaut, le script de standardisation utilise le fichier de configuration dont l'emplacement est indiqué par la variable d'environnement « GARHARVEST\_CONF ». Il est possible de spécifier un autre fichier de configuration en utilisant le paramètre « --config fichier\_de\_configuration ».

## 9.5 Prérequis techniques

Le script de standardisation des vignettes est développé en langage Python version 3 et exploite la bibliothèque logicielle GraphicsMagick spécialisée dans le traitement d'images en ligne de commande. Les formats acceptés en entrée (vignettes éditeur) sont les formats d'images Web, à savoir : PNG, JPEG et GIF.

Les fonctionnalités de l'outil ImageMagick utilisées principalement ici seront la mise à l'échelle (commande « resize ») et la composition de plusieurs (commandes « composite »).

Il est donc nécessaire que les paquets logiciels correspondants soient installés sur la plate-forme cible.

## 10 Annexe : Résultat de l'étude du passage de ScoLOMFR 8.0 à ScoLOMFR 9.0

### 10.1 Résultat de l'étude

Le GAR utilise actuellement la version 8.0 du format ScoLOMFR pour le traitement des notices éditeurs. Il est demandé d'effectuer la migration vers le format ScoLOMFR v9.0. Cette migration prendra en charge les changements de :

- ScoLOMFR 8.1 (publiée le 05 janvier 2023), il faut noter que cette version est une livraison partielle des fichiers et elle prend en compte uniquement les vocabulaires impactés dans les modifications :
  - vocabulaires 015, 022, 028, 047, 0151, 0152.
  - Cette version comprend les ajouts et les modifications suivantes :
    - Complément aux programmes de mathématiques sur l'entrée « enseignement scientifique (enseignement commun) » en vocabulaires 0151 et 0152 incluant les relations inter-vocabulaires 015 et 022 ;
    - Ajout du vocabulaire 047 IA – Intelligence Artificielle (et de sa mention dans le vocabulaire 028) ;
    - Fichiers XSD & documentation d'accompagnement de la V8.1.
- ScoLOMFR 9.0 (publiée le 19 juin 2023), cette version comprend les ajouts et les modifications suivantes (<https://www.reseau-canope.fr/scolomfr/telecharger-tous-les-fichiers>) :
  - Mises à jour sur les vocabulaires associés aux programmes (vocs 015GTPX, 015A, 015, 0151, 0152, 0153) ;
  - Associations inter-vocabulaires 15-22 / 15-151 / 22-151 / 15-152 / 22-152 (associations visibles en sorties SKOS) ;
  - Mise à jour du vocabulaire Lexique de l'intelligence artificielle (voc 047) ;
  - Mise à jour du vocabulaire des Diplômes (voc 029) ;
  - Ajout du vocabulaire 039 Référentiel de compétences numériques des métiers de l'enseignement, de l'éducation et de la formation (CRCNédu) ;
  - Mise à jour du vocabulaire Objectif de classification (voc 028) avec l'ajout du vocabulaire 039 ;
  - Mise à jour du modèle de données, de l'ontologie ScoLOMFR et des règles SHACL ;
  - Mise à jour des XSD et de la documentation utilisateur.

L'évolution principale « Migration ScoLOMFR de la version 8.0 à la version 9.0 » consiste en :

- La vérification de la compatibilité ascendante du schéma 9.0
- L'intégration de la xsd ScoLOMFR 9.0 et la mise à jour des vocabulaires utilisés par le batch de moissonnage
- La modification des tables de correspondance de vocabulaires issus de skos - les matières enseignées (voc-015) et les niveaux éducatifs (voc-022) - avec les tables BCN par l'utilisation du vocabulaire d'alignement ScoLOMFR avec les tables BCN.

## 10.2 Présentation de la solution

L'intégration de la version de ScoLOMFR utilisée par le GAR est une opération délicate. Afin d'opérer une transition douce, il est nécessaire de s'assurer de :

- la vérification de la compatibilité ascendante du nouveau schéma : les notices valides ScoLOMFR V8.0 doivent être valides après la prise en compte de la V9.0 ainsi que la mise à jour des vocabulaires ;
- l'intégration de la xsd ScoLOMFR 9.0 utilisée par le module de moissonnage pour profiter des nouveaux vocabulaires mis à disposition ;
- l'adaptation des tests de validité si nécessaire du module de moissonnage pour autoriser les nouveaux concepts tout en supportant encore les précédents ;

Ces 3 points sont détaillés dans les sections suivantes.

### 10.2.1 Vérification de la compatibilité ascendante des schémas 9.0

Le schéma XSD ScoLOMFR 9.0 n'apporte aucune différence structurelle par rapport au schéma ScoLOMFR 8.0. Les différences constatées le sont sur le contenu des vocabulaires :

- la liste des valeurs de « source » a été complétée avec « SCOLOMFRv9.0 »
- La mise à jour des vocabulaires

Note : Certains concepts sont classés « Archivé », « Obsolète », « Ancien », « Fin de vie » ou « Déprécié » dans le SKOS, mais ils doivent rester présents dans les fichiers voc des schémas, pour être toujours valides.

Pour le passage de la V8.0 à la V8.1 :

Nouveaux concepts qui ont été ajoutés :

- Dans le vocabulaire 028 : ajout d'un concept lié à l'ajout du vocabulaire 047 ;
- Dans le vocabulaire 0151 : 19 concepts liés au complément aux programmes de mathématiques sur l'entrée « enseignement scientifique (enseignement commun) » ;
- Mise à jour des relations inter-vocabulaire dans les autres vocabulaires (0151, 0152, 22 et 15).

Il n'existe pas de tableau différentiel pour ce passage partiel, cette pré-étude a été réalisée pour répondre au complément d'information demandées.

Pour le passage de la V8.1 à la V9.0 :

- 2 concepts ont fusionné, 20 sont archivés, 20 sont en fin de vie, 13 sont dépréciés et le reste (59) des modifications sont des signalements d'obsolescence de concepts.
- Un tableau est désormais fourni identifiant les changements apparus lors de la migration de la version 8.0 vers 9.0 identifiant les 113 modifications :
- [https://www.reseau-canope.fr/scolomfr/sites/default/files/2023-06/tableau\\_diff%C3%A9rentiel\\_V8.1\\_V9.pdf](https://www.reseau-canope.fr/scolomfr/sites/default/files/2023-06/tableau_diff%C3%A9rentiel_V8.1_V9.pdf)

Pour compléter cette information, les actions suivantes ont été réalisées :

- Analyse des documents publiés sur la V8.1 et identification du différentiel entre la V8.0 à la V8.1 ;

- Consolidation des changements entre la V8.0, V8.1 et V9.0 ;
- Etude des impacts d'une seule publication de la V8.0 à la V9.0 sans passer par une publication intermédiaire vers la V8.1.

Suite à cette étude, il est ainsi préconisé de réaliser une seule montée de version de V8.0 à la V9.0 incluant les concepts ajoutés dans V8.1.

- Il reste un impact à la marge, s'il faut revenir à la version précédente (retour à la V8.0), nous perdrons les concepts du vocabulaire 047 ajoutés en V8.1 à cause de l'introduction d'un nouveau concept dans le vocabulaire 028 dû à l'ajout du vocabulaire 047.

NB : La montée de version vers V8.1 n'ayant jamais été réalisée au niveau du GAR, ce cas ne peut pas se produire.

### 10.2.2 Intégration du format ScoLOMFR 9.0

L'intégration du schéma ScoLOMFR 9.0 au sein du module de moissonnage consiste en 2 opérations distinctes :

- Intégration du schéma XSD pour la validation systématique des notices moissonnées ;
- Intégration du contenu des vocabulaires ScoLOMFR 9.0.

L'intégration du nouveau schéma XSD au sein du module de moissonnage tient de l'opération de remplacement de fichiers. L'intégration des concepts de vocabulaire nécessite l'ajout de vocabulaires et l'ajout ou la suppression de données non contraintes.

En effet, depuis la sortie des premières versions de ScoLOMFR, la génération d'un format pivot\* stable est nécessaire pour s'affranchir des variations de formats de vocabulaires. Ce format est généré à partir d'un script shell d'extraction des données dans les fichiers ScoLOMFR en entrée.

Notes :

Le format SKOS-XL est le format utilisé pour les données de vocabulaire. C'est donc à partir de ce format que nous extrayons les concepts des vocabulaires pour l'exploitation au sein du GAR.

### 10.2.3 Mise à jour des contrôles effectués sur les notices

Il n'y a pas d'impact sur les contrôles de validité des notices au niveau du code.

Seule une écriture des tests prenant en compte la nouvelle version de ScoLOMFR 9.0 est fournie, afin de vérifier les compatibilités ascendantes.

Remarque : comme évoqué en préambule, des modifications des vocabulaires sont apparues et sont testées au niveau de la brique logicielle de génération des formats pivots\* utilisés lors de l'indexation du module de moissonnage, ils n'ont aucun impact sur le module de moissonnage du GAR.

\* Un format pivot est un format interne indépendant des variations qui peuvent avoir dans les vocabulaires et permet de maintenir un code source stable. Les outils mis en place sont mis à jour pour le mapping vers ce format pivot à chaque version.