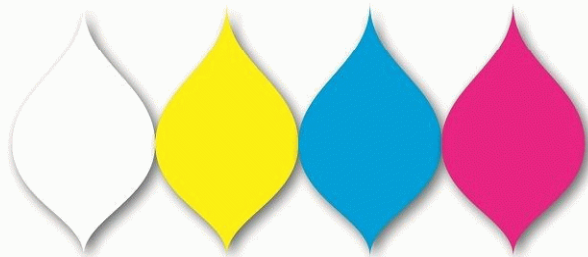


Ark-et-type d'identifiant : un nouveau mode d'identification pour les données archivistiques

Journée d'étude et d'échanges sur l'identifiant ARK (Archival Resource Key), sommet international ARK organisé par la BNF le 21 mars 2018

Delphine Jamet – Archives départementales de la Gironde
@Delphine_Jamet



Sommaire

- Archives et identifiants pérennes : quel usage ?
- Le choix de ARK
- Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin
- La question de la gouvernance des identifiants

Archives et identifiants pérennes : quel usage ?

- **Diffusion des contenus archivistiques sur le WEB**

permettre l'interopérabilité : référencement des contenus réutilisables par la communauté archivistique → Sources complémentaires, notices d'autorité...

- ***Archivage électronique :***

OAIS : constituer un paquet d'information pérenne à archiver → l'identifiant de la ressource numérique est un des éléments constitutifs de l'information de pérennisation

Le choix de Ark

Ark versus DOI

ARK et DOI fournissent tous deux un mécanisme d'identification pérenne des objets.

Historiquement DOI est associé au monde des éditeurs (notamment de publications techniques)

ARK est plutôt associé au monde des administrations publiques culturelles (bibliothèques, archives, etc.).

Les arguments qui nous ont fait choisir ARK :

- Gratuité
- Normalisation
- Pérennité
- Citabilité facilitée par une chaîne de caractères assez courte

Choix confirmé par la nouvelle version du [Référentiel général d'interopérabilité \(RGI\)](#) qui pose les règles d'échange entre les administrations. ARK est retenu comme standard pour identifier les objets numériques de manière pérenne.

Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin

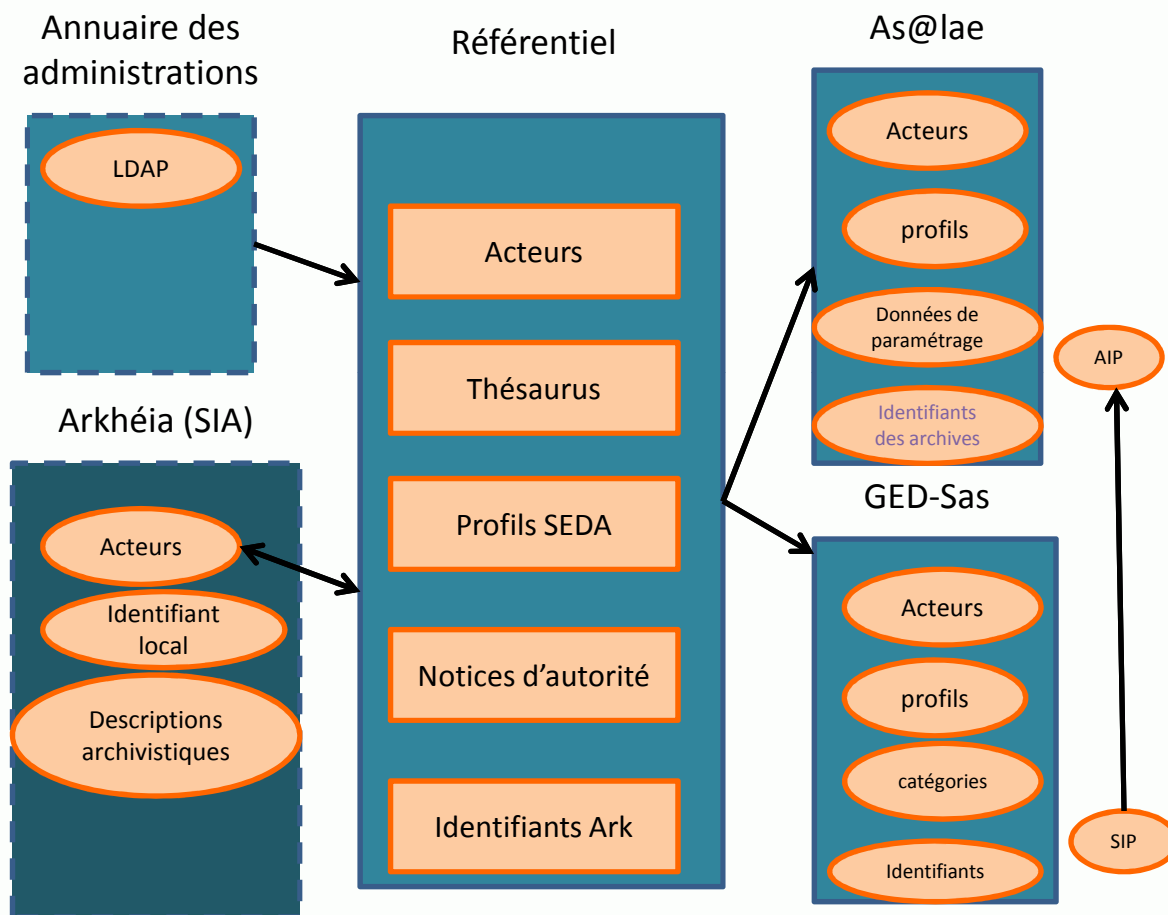
- **La brique logiciel « Le référentiel girondin » : clef de voute du système**

Objectif : réalisation d'un référentiel de structuration et de gestion des métadonnées, interopérable, développé en open source, prenant en compte le contexte normatif et basé sur une ontologie pour créer, importer, centraliser, gérer et exposer :

- des autorités et unités administratives
- des vocabulaires contrôlés externes ou internes
- des notices d'autorité des acteurs de l'archivage
- des profils d'archivage conformes au SEDA (toutes versions)
- des identifiants pérennes : numéroteur centralisé du système d'archivage électronique

Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin

Le référentiel girondin : la clé de voute du système

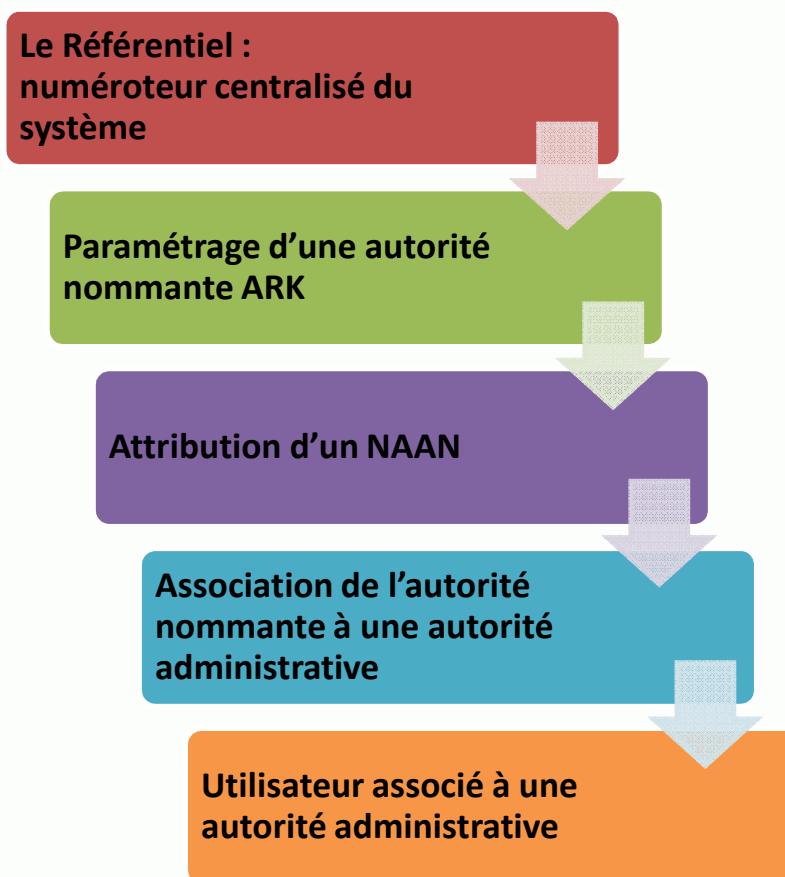


Fonctionnalités :

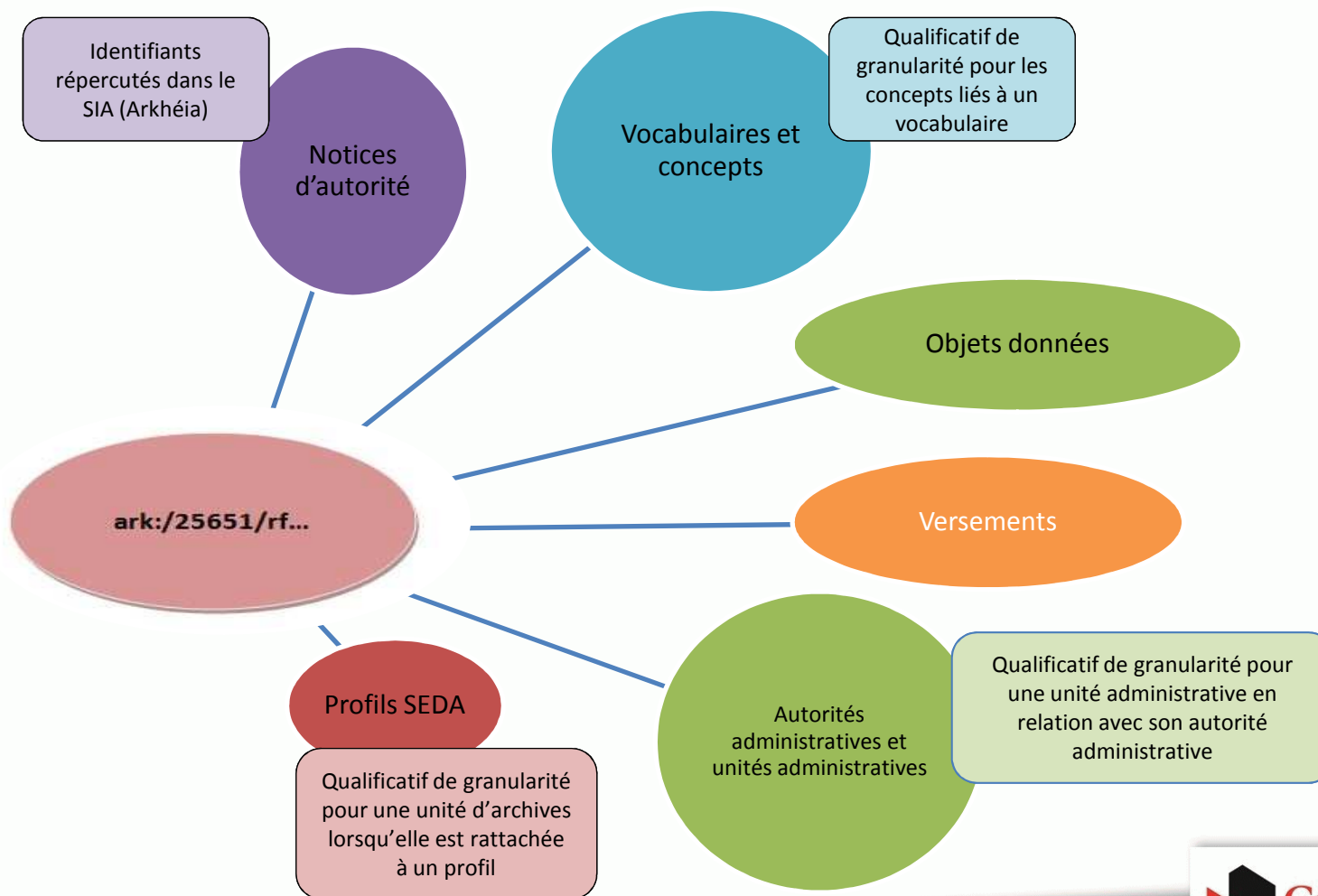
Simplification des processus d'archivage par réutilisation des données sans re-saisie

Moissonnage et mise à disposition des données de référence et des métadonnées utiles à la gestion du cycle de vie documentaire : dialogue des outils en utilisant le protocole **OAI-PMH**

Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin



Ark en Gironde : Le modèle d'attribution des identifiants au sein du SAE girondin



Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin

Structure et intégrité des identifiants ARK dans le SAE girondin

La société Logilab développe le référentiel avec l'outil cubicweb qui permet :

- de relier les données entre elles via des liens et **des identifiants** en supprimant les ambiguïtés.
- de fournir à des briques logicielles externes des identifiants pérennes via le protocole **OAI-PMH**

— Structure des identifiants

Les parties **Name et qualifier** sont construites sous la forme d'une chaîne de caractères aléatoires avec les contraintes suivantes :

- une chaîne de caractères alphanumériques avec **une séquence en base 29** : uniquement des consonnes (sauf l) et des chiffres 0 à 9
- la partie Name commence par un préfixe « *rf* » **fixant le contexte d'attribution des identifiants**
- une longueur fixe de **10 caractères**

Exemples

Identifiant de l'autorité administrative Département de la Gironde :

ark:/25651/rfk54146jg

Identifiant de du vocabulaire d'indexation géographique girondin :

ark:/25651/rf6v3254hg

Identifiant d'un concept du vocabulaire d'indexation géographique :

ark:/25651/rf6v3254hg/5g4z671vxz

Ark en Gironde : le modèle d'attribution des identifiants au sein du SAE girondin

– Intégrité des identifiants

Les identifiants ARK sont stockés dans une table SQL dont la structure est la suivante :

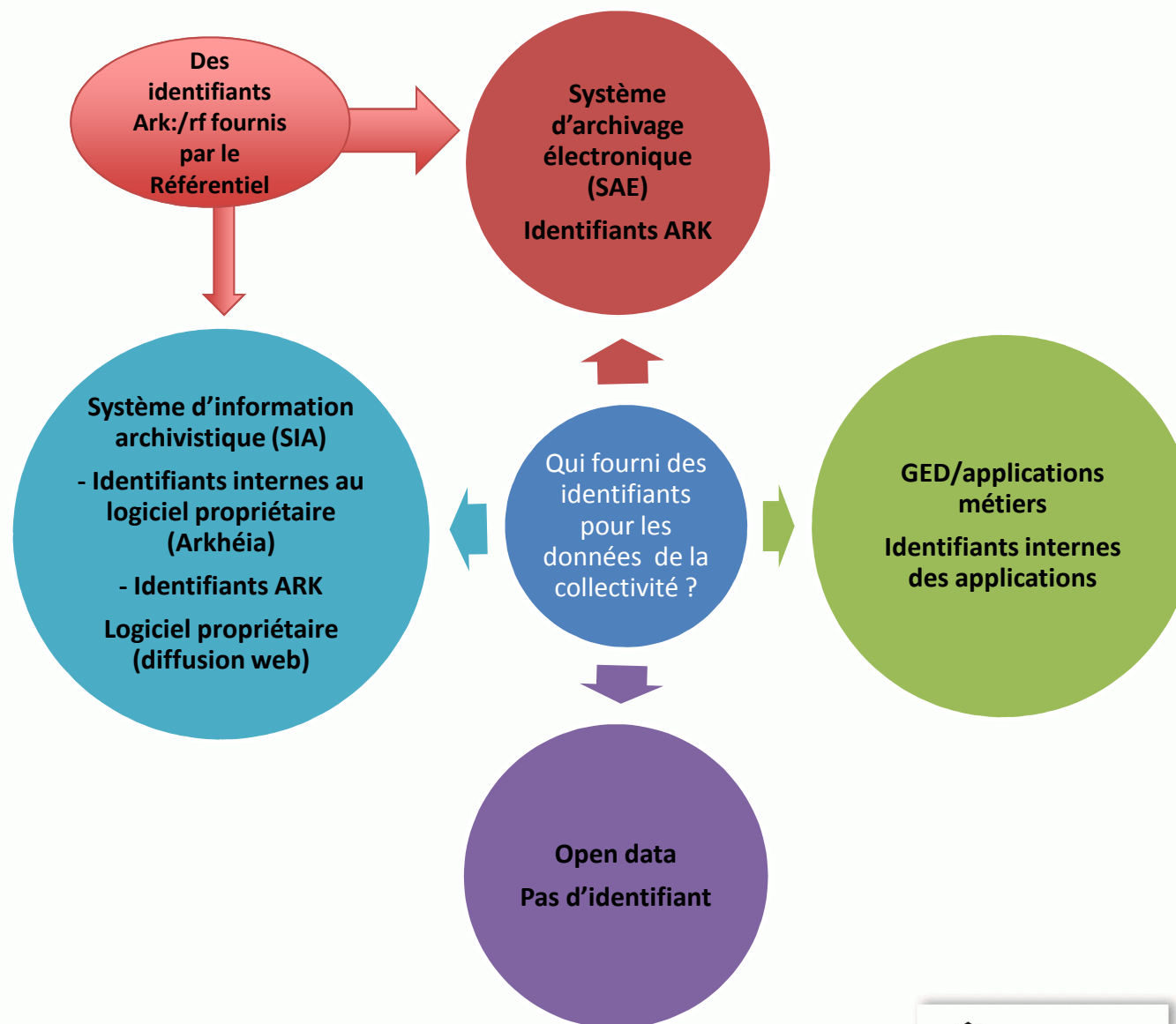
Colonne	Type	Modificateurs
naan	integer	non NULL
name	text	non NULL
qualifier	text	non NULL

Contrainte d'unicité sur le tuple (naan, name, qualifier)

- Pour générer un nouvel identifiant, on produit les chaînes de caractères aléatoires selon les règles évoquées ci-dessus en s'assurant (via le gestionnaire de base de données) de l'unicité de l'identifiant.

La question de la gouvernance des identifiants

Les identifiants des données nativement numériques produites aujourd'hui au sein du Département de la Gironde



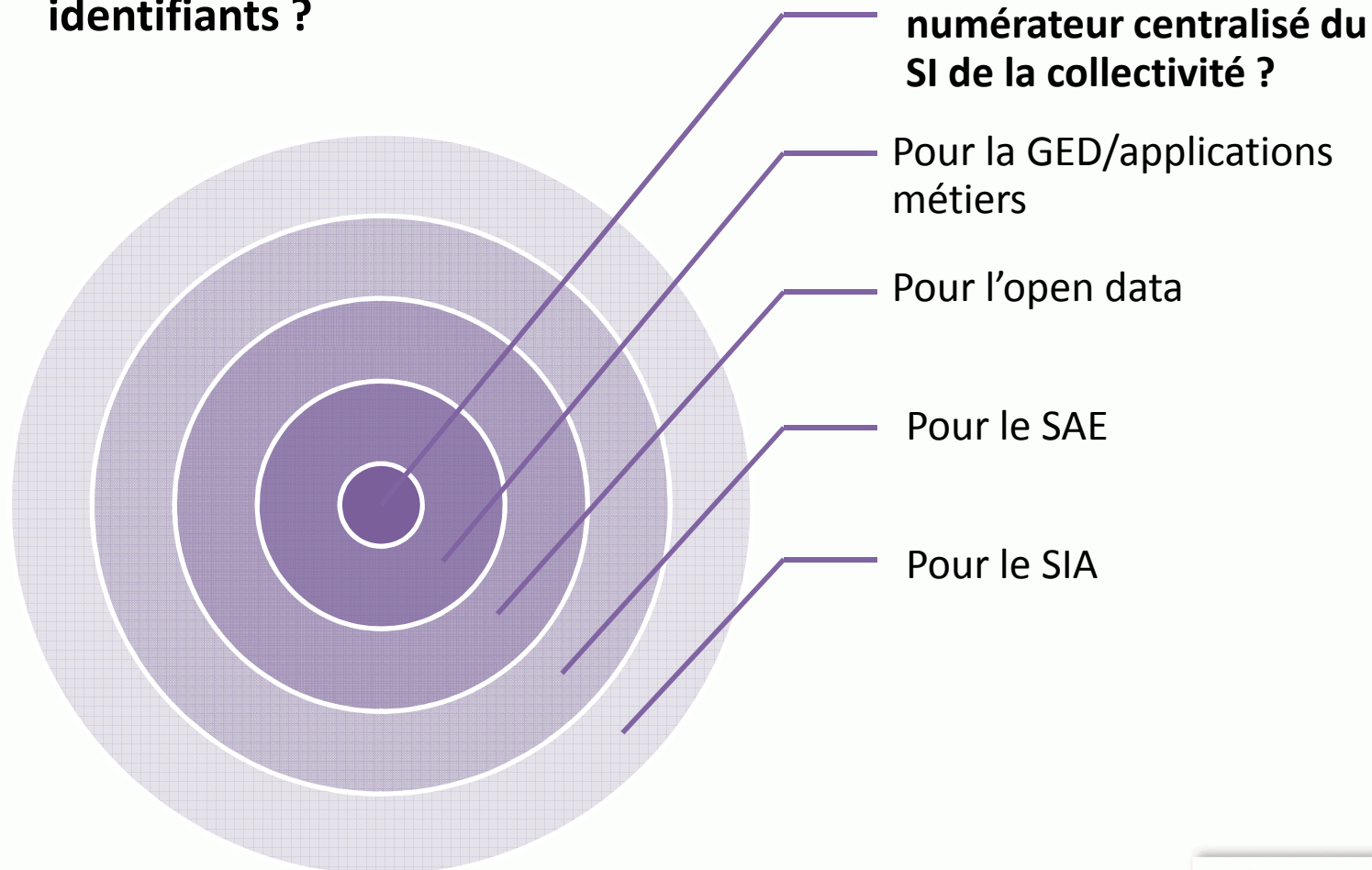
La question de la gouvernance des identifiants

Les limites du système actuel d'attribution des identifiants

- l'absence d'identifiants
- des identifiants qui ne prennent pas en compte le cycle de vie de l'information
- des identifiants non homogènes au sein du SI
- des identifiants non pérennes au sein du SI
- des identifiants dépendant de logiciels propriétaires

La question de la gouvernance des identifiants

La vision cible pour l'attribution des identifiants ?



Merci de votre attention



Delphine Jamet

d.jamet@gironde.fr

@Delphine_Jamet

<http://saem.e-bordeaux.org/>