

**Objet :** 1<sup>ère</sup> réunion du groupe de travail IRU du CNIG

**Ordre du jour :**

- Tour de table des membres du groupe
- Rappel du mandat
- Rappel des discussions ayant pu avoir lieu sur le sujet dans le groupe Services
- Passage en revue des documents à considérer dans le groupe
- Discussions sur les technologies à considérer et éléments à inclure dans le guide
- Suite des actions

**Résumé des discussions :**

La présentation est disponible [ici](#).

## 1 Tour de table des membres du groupe

Le tour de table est l'occasion de recenser à la fois les attentes du groupe et l'existant en ce qui concerne les IRU dans les différents organismes représentés.

La question des IRU se pose dans beaucoup d'organismes mais elle n'est pas tranchée dans tous. Les plateformes constatent également que les pratiques sont très diverses.

### IGN

Aucune politique d'identifiant résolvable n'a été définie pour l'instant. Les identifiants des fiches de métadonnées sont des chaînes de caractères sémantiquement parlantes.

### Museum d'histoire naturelle

Des travaux ont été initiés dans le cadre de la mise en conformité INSPIRE. Deux documents seront mis à la disposition du groupe. Le choix a été fait d'un identifiant composé d'une URL et un UUID, permettant de combiner une partie opaque et une partie lisible. La partie opaque permettant de ne pas avoir à gérer de versionnement des identifiants et de privilégier la permanence de l'identifiant.

### Etalab

L'absence d'identifiant unique est un frein dans le moissonnage des catalogues : différents catalogues contiennent ainsi la même donnée décrite n fois sans possibilité de traquer ces doublons. Ex : Scan25 IGN.

La stratégie actuelle est de créer un UUID sur la plate-forme data.gouv.fr par un hash de certaines infos. Une API permet de résoudre une URL contenant l'UUID.

L'UUID pointe sur des jeux de données, pas des métadonnées.

### MIG

Les recommandations du guide de saisie des métadonnées de données sont implémentées dans Geo-IDE et GéoSource.

Le fait qu'il y ait deux règles qui cohabitent (NDR : Un scénario à court terme composé de FR-SIREN et d'un identifiant unique et un scénario à long terme qui préconise une URL.) pose des problèmes de cohérence dans le Géocatalogue.

### BRGM

Aucune règle précise n'a été fixée. Un grand nombre de métadonnées sont gérées et une harmonisation nationale serait donc appréciée.

### GéoPicardie

Pour l'instant aucune règle commune n'a été fixée, dans l'attente de recommandations nationales qui puissent être appliquées par tous les partenaires.

### PIGMA

Le nombre de partenaires est très important et aucune règle n'a pu être imposée pour l'instant. Un grand nombre de partenaires semble avoir opté pour le scénario avec une URL contenant leur nom de domaine. A défaut, l'URL de PIGMA peut être utilisée.

### GéoBretagne

Le nombre de partenaire étant important, les règles de chacun s'accumulent sur la plateforme. Le problème de l'identifiant sera différent entre un producteur qui dispose de son propre catalogue pour créer ses fiches et un qui utilise le catalogue de la plate-forme. La solution préconisée devra tenir compte du cout de mise en place.

### Rennes Métropole

La règle est d'appliquer l'option 1 du guide : FR-SIREN + chaîne composée du codage de la nomenclature thématique GéoBretagne + jeu de données.

L'utilisation de la nomenclature permet d'avoir un identifiant sémantiquement parlant ce qui est important pour pouvoir hiérarchiser les données.

Une attente largement partagée est de comprendre l'intérêt que présentent des identificateurs resolvables dans une architecture.

**D1** : Un chapitre présentant les avantages de ce mécanisme sera ajouté au guide.

## **2 Rappel du mandat**

Le mandat est disponible [ici](#).

A la lecture du mandat ont lieu les discussions suivantes :

Deux éditeurs de logiciel sont mentionnés dans le mandat mais il serait bon de faire des recommandations génériques à destination de l'ensemble des éditeurs de logiciels.

**D2** : Le guide contiendra des recommandations spécifiques à destination des éditeurs de logiciel.

Une architecture basée sur des IRU n'a de sens que si un certain nombre de bonnes pratiques sont respectées : non-duplication des fiches de métadonnées (qui provoque souvent une non-unicité des identificateurs de ressources) et versionnement des métadonnées.

**D3** : Un lien vers le guide de gestion des catalogues sera fait pour rappeler ces bonnes pratiques.

La mise en place d'une telle architecture devra se faire à moindre coût.

**D4** : La description des solutions dans le guide prendra en compte le cout de mise en place et de gestion de la solution.

## **3 Rappel des discussions ayant pu avoir lieu sur le sujet dans le groupe Services**

Suite aux discussions qui avaient été menées dans le Groupe de travail Services un document avait été initié. Il est disponible [ici](#). Il pourra servir de base aux travaux du groupe IRU.

Le GT services a conclu qu'un utilisateur qui fait une recherche sur un catalogue ne veut pas accéder à la métadonnée mais directement au jeu de données.

Un identifiant resolvable facilite cette opération. Les arguments seront à détailler dans le guide. La question suivante devra également être posée : Vers quel(s) éléments pointe l'identifiant (donnée, métadonnée...) ?

## **4 Passage en revue des documents à considérer dans le groupe**

Le tour de table du début de la réunion a permis d'identifier des documents techniques ou de bonnes pratiques qui pourraient servir à alimenter les discussions.

Un document a été créé [ici](#) pour rassembler ces documents.

**D5** : Chacun est invité à renseigner les documents dont il a connaissance dans cette bibliographie.

## 5 Discussions sur les technologies à considérer et éléments à inclure dans le guide

### Liste des technologies

La liste des technologies pré-identifiée est la suivante :

- HTTP URIs
- Persistent URIs ( PURL + ARK)
- URN
- DOI
- LSID

Le tour de table a permis de mettre en évidence une technologie à ajouter : UUID. Il est en effet rappelé que l'utilisation de chaînes de caractères combinées avec un service qui permet de les résoudre est une pratique très courante.

**D6** : Cette technologie sera ajoutée à la liste.

**D7** : Les technologies pourront ne pas faire être traitées de manières exhaustives si elles s'avèrent représenter peu d'intérêt.

### Critères de sélection

La liste des critères qui permettent d'évaluer une technologie par rapport à une autre est la suivante :

- Unicité
- Persistance
- Résolvabilité
- Fiabilité
- Traçable
- Peu coûteuse
- Conforme à la stratégie des différents organismes

**D8** : Les critères suivants seront ajoutés

- L'opacité de l'identifiant : un utilisateur est-il capable de comprendre ou non comment est structuré l'identifiant ? En effet certains organismes privilégient un identifiant opaque, qui permet de ne pas avoir à modifier l'identifiant en cas de modifications sur la donnée alors que d'autres privilégient une lisibilité qui permet de mieux organiser les données hiérarchiquement.
- Le degré de persistance : Cette persistance peut être dépendante de l'organisme (probabilité que son nom de domaine change, possibilité de mettre en place des redirections ou non).
- Les possibilités d'évolution de l'identifiant
- La facilité de génération de l'identifiant (via une API centralisée par exemple)
- L'indépendance de la technologie : La charge de création peut-elle être répartie entre différents partenaires d'une plateforme ?
- Le cout de la solution (cf. D4)
- La possibilité de résoudre les problématiques de résolution internes (LAN) / externes (WAN)

### Autres éléments à inclure

Le guide pourra également inclure d'autres éléments qui rappellent les bonnes pratiques, en lien avec les guides existants.

**D9** : Au guide sera ajouté une partie sur :

- la possibilité d'avoir un jeu de données identifié par plusieurs identificateurs de ressources. Cette possibilité peut permettre de gérer à la fois identifiants internes et externes qui peuvent répondre à des exigences différentes.
- des recommandations à destination des éditeurs (cf.D2)
- l'implémentation des identificateurs (MD\_Identifier vs RS\_Identifier) : Ces recommandations seront faites en suivant les travaux menés au niveau européen.
- le rôle des plateformes régionales dans le filtrage des doublons et le relais des bonnes pratiques.
- Un état des lieux des resolver existants
  - Note : Un resolver va être développé au BRGM.

## 6 Suite des actions

Les documents suivants seront partagés pour permettre à l'ensemble du groupe de les compléter :

Bibliographie : [ici](#)  
 Plan du guide : [ici](#)  
 Pratiques actuelles : [ici](#)

**Prochaine réunion** : le 17 novembre 2014 à 10h00.

**Liste de diffusion**

Participants	Organisme
CHARTIER Benjamin CHATAIGNER Julie CHAUVEAU Chloé DAVID Benoît DESBOEUFS Jérôme JAQUEMET Clément LAMBOIS Marie PHUNG Fabrice	<b>GeoPicardie</b> <b>MNHN</b> <b>GIPATGERI</b> <b>CGDD/DRI/MIG</b> <b>SGMAP / Etalab</b> <b>CGDD/DRI/MIG</b> <b>IGN</b> <b>DREAL</b> <b>Bretagne/COPR</b> <b>EV</b> <b>Ville de Rennes</b> <b>/ Rennes</b> <b>Métropole</b> <b>BRGM</b>
REBOUX Maël TAFFOUREAU Etienne	

Personnes à informer	Organisme
RYCKELYNCK Guillaume CHARTIER Benjamin SATRA LE BRIS Catherine QUIMBERT Erwan MEILLON Julien HOUBIE Frédéric CARTOIXIA Mathieu TAFFOURAU Etienne DAVID Benoît JACQUEMET Clément	<b>CIGAL</b> <b>GeoPicardie</b> <b>Ifremer</b> <b>Ifremer</b> <b>Ifremer</b> <b>Geomatys</b> <b>Isogéo</b> <b>BRGM</b> <b>CGDD/DRI/MIG</b> <b>CGDD/DRI/MIG</b> <b>DREAL</b> <b>Bretagne/COPREV</b>
PHUNG Fabrice COUDERCY Laurent	<b>Onema</b> <b>Ville de Rennes /</b> <b>Rennes Métropole</b>
REBOUX Maël CHAUVEAU Chloé WATIER Loïc CHATAIGNER Julie DESBOEUFS Jérôme PECRIAUX Gilles LORY Pascal LEOBET Marc VERGEZ Pierre	<b>GIPATGERI</b> <b>GIPATGERI</b> <b>MNHN</b> <b>SGMAP / Etalab</b> <b>IUT de Troyes (AUBE)</b> <b>IGN</b> <b>CGDD/DRI/MIG</b> <b>IGN</b>

Date	Visa	Nom	Service
06/10/2014		Marie Lambois	IGN/SAI