

Projet de guide CNIG sur la gestion des identificateurs de ressource uniques (IRU) dans Inspire

Objectif du guide

La gestion des identificateurs de ressource uniques (IRU) est un aspect fondamental de l'architecture de services Inspire car ces identificateurs font le lien entre les métadonnées de données et les métadonnées de services, et également entre les capacités des services et les ressources servies par le service.

Le guide CNIG sur les métadonnées (Guide de saisie des éléments de métadonnées INSPIRE appliqué aux données, version 1.1 – Décembre 2013) prévoit 2 cas pour les IRU et recommande le cas A qui consiste à gérer des adresses Internet et à défaut propose le cas B d'une solution intermédiaire utilisant un identificateur non Internet.

A l'occasion de la rédaction des guides CNIG sur les services, il apparaît que l'infrastructure ne sera pleinement opérationnelle qu'à condition que les fournisseurs respectent la recommandation du cas A et, en outre, que cette recommandation n'est aujourd'hui quasiment pas suivie.

Le présent guide a pour objet, d'une part, de préciser la recommandation mentionnée et, d'autre part, de fournir des cas de bonnes pratiques de cette recommandation afin d'inciter les fournisseurs de données à la suivre.

Rappel du contexte réglementaire et des recommandations

Le règlement métadonnées

Le règlement métadonnées définit les 2 éléments de métadonnées suivants :

Identificateur de ressource unique

Une valeur identifiant la ressource de manière unique.

Le domaine de valeur de cet élément de métadonnées est un code obligatoire sous forme de chaîne de caractères, généralement attribué par le propriétaire des données, et un espace de noms sous forme de chaîne de caractères qui identifie de manière unique le contexte du code d'identification (par exemple le propriétaire des données).

1.6. Ressource couplée

Si la ressource est un service de données géographiques, cet élément de métadonnées identifie, le cas échéant, la série ou les séries de données géographiques cibles du service grâce à leurs identificateurs de ressource uniques (Unique Resource Identifiers, URI).

Le domaine de valeur de cet élément de métadonnées est un code obligatoire sous forme de chaîne de caractères, généralement attribué par le propriétaire des données, et un espace de noms sous forme de chaîne de caractères qui identifie de manière unique le contexte du code d'identification (par exemple le propriétaire des données).

Il prévoit de plus que :

- chaque MD de série de données doit comporter un ou plusieurs Identificateur de ressource unique.

- chaque MD de service peut comporter une ou plusieurs ressources couplées ; cette ressource est obligatoire si des liens avec les séries de données avec lesquelles le service opère sont disponibles.

Le règlement service

Le règlement service exige que la réponse d'un service de consultation à la demande « Accéder à des métadonnées du service de consultation » contienne des métadonnées des couches qui comportent pour chaque couche « L'identificateur unique de ressource utilisé pour créer la couche ».

Il exige qu'en réponse à la demande « Accéder à des métadonnées du service de téléchargement » adressée à un service de téléchargement « Les éléments de métadonnées Inspire des séries de données géographiques disponibles doivent être fournis »

Le guide technique sur les services de téléchargement propose 2 options : soit de fournir dans les capacités pour chaque série de données un lien vers la métadonnée de la série de données, soit de publier dans les capacités l'ensemble des éléments Inspire.

Le guide CNIG métadonnées

Le guide CNIG sur les métadonnées (Guide de saisie des éléments de métadonnées INSPIRE appliqué aux données, version 1.1 – Décembre 2013) prévoit 2 cas pour les IRU :

- A) Recommandation nationale : une adresse Internet
 - l'IRU doit être fourni comme une URI (Uniform Resource Identifier) au format http
 - cet URI soit également une URL (adresse internet) permettant d'accéder à la ressource (ou à ses métadonnées).
- B) Solution intermédiaire : une solution avec un identifiant non Internet

Mise en oeuvre d'un identifiant basé sur le numéro SIREN de l'INSEE, précédé par le code pays.

Le bloc identifiant l'organisme producteur sera constitué du code du pays et du code SIREN de l'organisme. Exemple : fr-243500139.

Le bloc identifiant la donnée pourrait être issu d'une nomenclature, porter une date ou une version. Exemples : orthophoto2004, ou 35206plu20100612.

Au final, par exemple : fr-243500139-35206plu20100612.

Il est reconnu qu'il faut éviter les noms de services dans le bloc identifiant la donnée. Plus généralement, toutes les chaînes de caractères signifiantes (nom d'organismes...) sont à éviter car trop susceptibles de changer. De plus, il est recommandé de fournir l'identifiant en minuscule, et d'utiliser le tiret (-) en guise de séparateur.

En cas d'absence de SIREN, une solution parallèle consiste à prendre le nom de domaine (celui de son site internet).

Recommandations

Recommandation n° 1

La recommandation du guide CNIG métadonnées d'utiliser comme IRU une URI (Uniform Resource Identifier) au format http qui soit en même temps une URL (adresse internet) laisse la possibilité que cette URL permette d'accéder à la ressource elle-même ou à ses métadonnées.

Pour satisfaire de manière simple le règlement service et notamment, d'une part, que les capacités des services de consultation dans les métadonnées de couche contienne l'IRU de la ressource utilisée pour créer la couche et, d'autre part, que les capacités des services de téléchargement contient l'URL vers les métadonnées de la série de données, il est nettement préférable que chaque série de données ait au moins une IRU qui soit une URL permettant d'accéder aux métadonnées de la série de données.

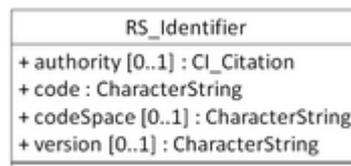
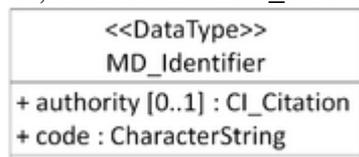
Ainsi il est nécessaire de compléter la recommandation CNIG par la recommandation suivante :

Pour chaque série de données, la fiche de métadonnées doit comporter en tant qu'IRU au moins une URL permettant d'accéder à la fiche de métadonnées de la série de données conforme à la norme ISO 19139.

Recommandation n° 2

Dans les documents XML des métadonnées, les normes ISO spécifient que les IRU sont encodées dans une structure MD_Identifier.

Afin de permettre la fourniture d'un espace de nom prévue par le règlement métadonnées, le guide technique européen sur les métadonnées autorise pour encoder les IRU d'utiliser soit une structure MD_Identifier, soit structure RS_Identifier.



Ainsi la recommandation est la suivante :

Les outils qui utilisent des documents XML de métadonnées doivent pouvoir reconnaître les IRU encodées soit sous la forme MD_Identifier soit sous la forme RS_Identifier.

Cette recommandation doit être complétée par les critères permettant aux outils de distinguer si un RS_Identifier décrit un système de projection ou un identificateur de ressource.

Bonnes pratiques de la première recommandation

Géocatalogue

Le nom de domaine du Géocatalogue peut être utilisé pour créer des URI.

Dans le cas B du guide CNIG d'un identificateur fondé sur le numéro SIREN du fournisseur de données, une URI peut être construite à partir de cet identificateur de la manière suivante :

`http://id.geocatalogue.fr/<identificateur_dorigine>`

Par exemple :

Si une fiche de métadonnées comportant l'IRU « fr-243500139-35206plu20100612 » est moissonnée par le Géocatalogue alors elle peut utiliser l'IRU suivante :

`http://id.geocatalogue.fr/fr-243500139-35206plu20100612`

Il faut dans ce cas que le fournisseur de données ajoute dans la fiche de MD l'IRU ainsi définie, et qu'il fasse moissonner sa fiche par le Géocatalogue.

Géosource

La version 2.11 de Géosource permet de définir automatiquement des IRU utilisables comme URL de la fiche de métadonnées.

Ces URI sont générés par défaut en s'appuyant sur l'URI non URL de la fiche de métadonnées (champ « fileIdentifier »), sous la forme : `http://id.<adresse du catalogue>/<fileIdentifier>`

Le logiciel Géosource peut modifier les fiches déjà saisies pour ajouter à chaque fiche ce nouvel IRU, et actualiser les liens entre fiches au sein du catalogue.

Ces IRU sont utilisables comme URL de la fiche de métadonnées en définissant sur le serveur web du catalogue des règles simples de réécriture d'URL pour la traduire en une URL utilisant le service GéoSource qui permet de résoudre cette IRU :

`http:// <adresse du catalogue>/srv/fre/xml.metadata.get?uuid=<fileIdentifier>`

Il est bien sûr possible de définir au sein de GéoSource d'autres IRU et sous d'autres formes utilisant un identifiant local autre que le « fileIdentifier ».

Utilisation de DOI

Pour les organismes de recherche les DOI (http://fr.wikipedia.org/wiki/Digital_Object_Identifier) constituent un bon IRU. Il est recommandé dans ce cas qu'une au moins des IRU de la série de données soit encodée sous la forme d'une URL. Une telle URL peut être constituée en utilisant un service résolveur des DOI, par exemple sous la forme : <http://dx.doi.org/<DOI>>, mais cette URL doit bien renvoyer la fiche de métadonnées suivant la norme ISO 19139.

Utilisation d'URN

Il est possible d'utiliser des URN (http://fr.wikipedia.org/wiki/Uniform_Resource_Name) comme URI, cependant il est pour cela nécessaire d'enregistrer l'espace de nom que l'on utilise auprès de l'IANA et ses formats d'URN. Il convient ensuite qu'une au moins des IRU de la série de données soit encodée sous la forme d'une URL qui utilisera un service résolveur d'URN à mettre en œuvre. Cette solution n'est pas recommandée dans une première approche car elle est plus complexe à mettre en œuvre que les précédentes.