

# Test Publication Géocatalogue / linked data -> réflexions sur l'architecture des MD

- Objectif : être découvrable via les outils classiques du Web
- Moyen : publication du contenu du GeoCatalogue en JSON-LD
  - Utilisation du vocabulaire Schema.org pour optimiser l'indexation
    - (expérience très récente au Canada : ça fonctionne)
  - Définir une architecture: structure des données, politique des URIs, liens vers les catalogues des fournisseurs, etc.

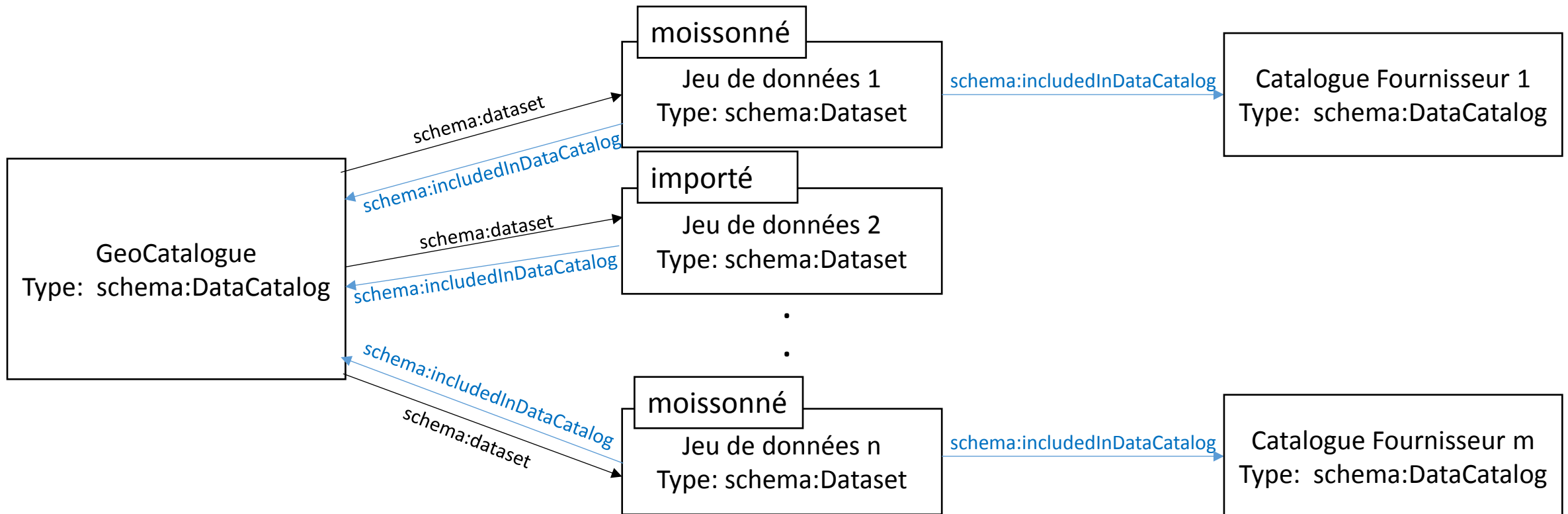
# Publier le catalogue en JSON-LD

- JSON-LD + Schema.org sont recommandés par Google, Bing, Yandex ... pour la structuration de données d'indexation.
- Génération statique vs dynamique:  
Google semble capable de lire du contenu structuré généré dynamiquement → Des tests sont nécessaires pour vérifier la qualité de l'indexation.
- Google propose un ensemble de guidelines à respecter lors du rajout de contenu structuré afin de garantir l'indexation:  
<https://developers.google.com/search/docs/guides/sd-policies>
- Le JSON-LD peut
  - être injecté dans une balise <script> dans les fiches HTML des entrées de catalogue (idem que les exemples <http://geocat.fr/>),
  - être déclaré dans des fichiers statiques ou au-dessus de flux.

```
<script type='application/ld+json'>
{
  "@context": "http://schema.org/",
  "@type": "Dataset",
  "@id": "http://data.brgm.fr/dataset/jeudedonneesXX",
  "includedInDataCatalog": "http://data.brgm.fr/catalogue/catalogueXX",
  "name": "Nom du jeu de données ou service",
  "description": "description du jeu de données ou service",
  "inLanguage": "fr",
  "distribution": {
    "encodingFormat": "XML",
    "contentURL": "URL_TELECHARGEMENT", ...
  },
  "about": [
    "http://inspire.ec.europa.eu/theme/XX"
  ]
}
</script>
```

# Architecture proposée

- Utiliser des URIs pérennes pour identifier les catalogues et les jeux de données.
- Utiliser `schema:dataset` et `schema:includedInDataCatalog` pour lier les catalogues et les jeux de données
- Politique d'URIs: deux cas de figures
  - Les jeux de données importés directement dans le GéoCatalogue
  - Les données moissonnées par le GéoCatalogue



# Architecture proposée

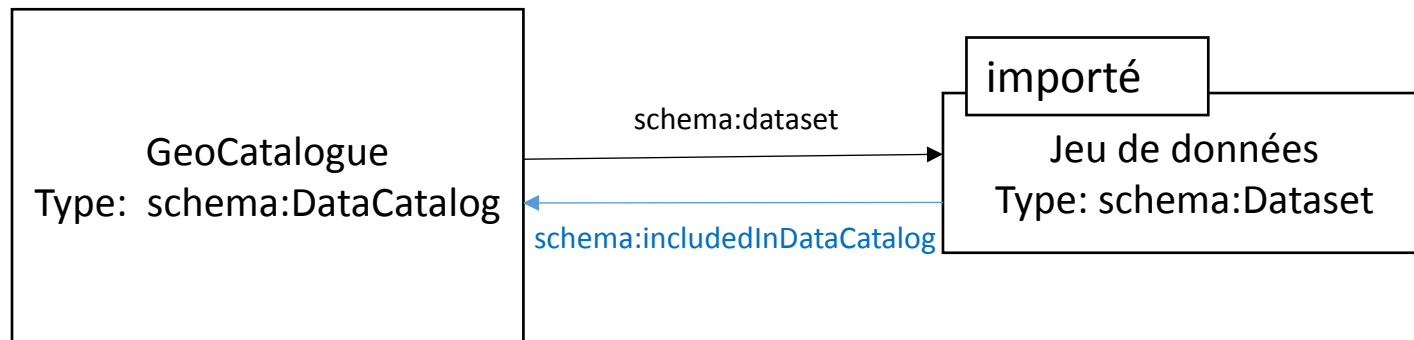
Politique d'URIs:

- Les jeux de données importés directement dans le GéoCatalogue :

→ Utiliser l'URI de base du geocat. Utiliser le pattern

[http://www.geocatalogue.fr/id/dataset/identifiant\\_jeu\\_de\\_donnees](http://www.geocatalogue.fr/id/dataset/identifiant_jeu_de_donnees)

→ lier au GéoCatalogue



# Architecture proposée

Politique d'URIs:

- Les données moissonnées par le GéoCatalogue :

→ les fournisseurs de données n'ayant pas de politique d'URI: proposer des URIs au niveau du GeoCatalogue, les remplacer si le fournisseur finit par définir sa politique d'URI avec ses propres JSON-LD qui décrivent les jeux de données / services (animation communauté)

→ les fournisseurs ayant déjà une politique d'URI: quels URIs dans les JSON-LD ?

- à court terme : Pas de JSON-LD chez les fournisseurs, définir des JSON-LD dans les fiches locales avec URIs geocat. Possibilité d'utiliser les équivalences d'URI avec schema:sameAs

- Solution à long terme: Proposer une structure JSON-LD que les fournisseurs doivent remplir en utilisant leur propre politique d'URI

