

GN RDF Factory : python scripts to generate Geonetwork compliant RDF Files

Claudia Lavalley Guzman

UMR TETIS

Séminaire SIST 2019 – Toulouse 4-8 Nov 2019

Contexte

- Mise en place d'une IDG pour l'UMR TETIS (C. Lavalley, S. Fortuno, R. Decoupes):
 - Valorisation du patrimoine de données spatiales :
 - Organisation et sécurisation des données
 - Documentation et mise à disposition des données produites
 - Accompagnement du changement de pratiques (Science Ouverte PGD)
 - Accès centralisé aux données produites et externes (référentiels, services de l'état, OpenIG...)
- Nécessité absolue d'associer les chercheurs : mise à disposition de vocabulaires
 - « communautés » : GCMD (NASA), WWF Ecoregions, COVADIS (services de l'état), ...
 - « maison » : Axes thématiques TETIS, capteurs (LiDAR, hyperspectral), ...
- Solution adoptée pour le moment :
 - Geonetwork + Geoserver avec facettes ad-hoc
 - Automatisation ++ : balayage de systèmes de fichiers avec scripts geometa + geonapi

Thesauri dans Geonetwork

- Fichiers SKOS/RDF (Simple Knowledge Organization System / Resource Description Framework).
RDF : Langage de base du Web sémantique (W3C), description des ressources Web et leurs métadonnées.
- SKOS se sert du cadre RDF comme moyen standard pour représenter un système de connaissances transférable entre applications, de manière interopérable.
- Blocs skos/rdf utilisés par GN :
 - Espace de nommage XML (**xmlns** : URLs avec spécifications rdf, skos, dc)
 - **skos:ConceptScheme** : permet de définir le titre, les conditions d'utilisation, les informations du producteur du Thesaurus et... le(s) point(s) d'entrée du thésaurus (top Concept)
 - **skos:Concept** : labels, relations hiérarchiques (broader(s), narrower(s)), relations ontologiques(?)

Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml xml(version="1.0" encoding="UTF-8")
<rdf:RDF xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept rdf:about="Données Hétérogènes"
  <skos:Concept rdf:about="ModDynSpatiales"
  <skos:Concept rdf:about="Teledetection"
  <skos:Concept rdf:about="Textes"
  <skos:Concept rdf:about="Science des Données"
  <skos:Concept rdf:about="Corridors Rivulaires"
  <skos:Concept rdf:about="Hydrosystèmes"
  <skos:Concept rdf:about="Paysages"
  <skos:Concept rdf:about="Biodiversité"
  <skos:Concept rdf:about="Numeriques-apprenants"
  <skos:Concept rdf:about="Multifonctionnels"
  <skos:Concept rdf:about="Territoires"
  <skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:ConceptScheme rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
    <dc:title TETIS Research Axes, version 1.0 31102019
    <dc:description TETIS Research Axes thesaurus for GeoNetwork opensource
    <skos:hasTopConcept rdf:resource="tetis_researchAxes_root"
```

Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml  xml(version="1.0" encoding="UTF-8")
<rdf:RDF  xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept  rdf:about="Données Hétérogènes"
  >
  <skos:Concept  rdf:about="ModDynSpatiales"
  >
  <skos:Concept  rdf:about="Teledetection"
  >
  <skos:Concept  rdf:about="Textes"
  >
  <skos:Concept  rdf:about="Science des Données"
  >
  <skos:Concept  rdf:about="Corridors Rivulaires"
  >
  <skos:Concept  rdf:about="Hydrosystèmes"
  >
  <skos:Concept  rdf:about="Paysages"
  >
  <skos:Concept  rdf:about="Biodiversité"
  >
  <skos:Concept  rdf:about="Numeriques-apprenants"
  >
  <skos:Concept  rdf:about="Multifonctionnels"
  >
  <skos:Concept  rdf:about="Territoires"
  >
  <skos:Concept  rdf:about="tetis_researchAxes_root"
  >
  <skos:ConceptScheme  rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
  >
    <dc:title  TETIS Research Axes, version 1.0 31102019
    >
    <dc:description  TETIS Research Axes thesaurus for GeoNetwork opensource
    >
    <skos:hasTopConcept  rdf:resource="tetis_researchAxes_root"
    >
```

Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml xml(version="1.0" encoding="UTF-8")
<rdf:RDF xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept rdf:about="Données Hétérogènes"
  <skos:Concept rdf:about="ModDynSpatiales"
  <skos:Concept rdf:about="Teledetection"
  <skos:Concept rdf:about="Textes"
  <skos:Concept rdf:about="Science des Données"
  <skos:Concept rdf:about="Corridors Rivulaires"
  <skos:Concept rdf:about="Hydrosystèmes"
  <skos:Concept rdf:about="Paysages"
  <skos:Concept rdf:about="Biodiversité"
  <skos:Concept rdf:about="Numeriques-apprenants"
  <skos:Concept rdf:about="Multifonctionnels"
  <skos:Concept rdf:about="Territoires"
  <skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:ConceptScheme rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
    <dc:title TETIS Research Axes, version 1.0 31102019
    <dc:description TETIS Research Axes thesaurus for GeoNetwork opensource
    <skos:hasTopConcept rdf:resource="tetis_researchAxes_root"
```

Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml xml(version="1.0" encoding="UTF-8")
<rdf:RDF xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept rdf:about="Données Hétérogènes"
  <skos:Concept rdf:about="ModDynSpatiales"
  <skos:Concept rdf:about="Teledetection"
  <skos:Concept rdf:about="Textes"
  <skos:Concept rdf:about="Science des Données"
  <skos:Concept rdf:about="Corridors Rivulaires"
  <skos:Concept rdf:about="Hydrosystèmes"
  <skos:Concept rdf:about="Paysages"
  <skos:Concept rdf:about="Biodiversité"
  <skos:Concept rdf:about="Numeriques-apprenants"
  <skos:Concept rdf:about="Multifonctionnels"
  <skos:Concept rdf:about="Territoires"
  <skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:ConceptScheme rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
    <dc:title TETIS Research Axes, version 1.0 31102019
    <dc:description TETIS Research Axes thesaurus for GeoNetwork opensource
    <skos:hasTopConcept rdf:resource="tetis_researchAxes_root"
```

```
<skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:prefLabel xml:lang="fr" tetis_researchAxes_root
  <skos:prefLabel xml:lang="en" tetis_researchAxes_root
  <skos:narrower rdf:resource="Territoires"
  <skos:narrower rdf:resource="Biodiversité"
  <skos:narrower rdf:resource="Paysages"
  <skos:narrower rdf:resource="Science des Données"
  <skos:narrower rdf:resource="ModDynSpatiales"
  <skos:narrower rdf:resource="Données Hétérogènes"
```

Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml  xml(version="1.0" encoding="UTF-8")
<rdf:RDF  xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept  rdf:about="Données Hétérogènes"
  <skos:Concept  rdf:about="ModDynSpatiales"
  <skos:Concept  rdf:about="Teledetection"
  <skos:Concept  rdf:about="Textes"
  <skos:Concept  rdf:about="Science des Données"
  <skos:Concept  rdf:about="Corridors Rivulaires"
  <skos:Concept  rdf:about="Hydrosystèmes"
  <skos:Concept  rdf:about="Paysages"
  <skos:Concept  rdf:about="Biodiversité"
  <skos:Concept  rdf:about="Numeriques-apprenants"
  <skos:Concept  rdf:about="Multifonctionnels"
  <skos:Concept  rdf:about="Territoires"
  <skos:Concept  rdf:about="tetis_researchAxes_root"
  <skos:ConceptScheme  rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
    <dc:title  TETIS Research Axes, version 1.0 31102019
    <dc:description  TETIS Research Axes thesaurus for GeoNetwork opensource
    <skos:hasTopConcept  rdf:resource="tetis_researchAxes_root"
```

```
<skos:Concept  rdf:about="tetis_researchAxes_root"
  <skos:prefLabel  xml:lang="fr" tetis_researchAxes_root
  <skos:prefLabel  xml:lang="en" tetis_researchAxes_root
  <skos:narrower  rdf:resource="Territoires"
  <skos:narrower  rdf:resource="Biodiversité"
  <skos:narrower  rdf:resource="Paysages"
  <skos:narrower  rdf:resource="Science des Données"
  <skos:narrower  rdf:resource="ModDynSpatiales"
  <skos:narrower  rdf:resource="Données Hétérogènes"
```

```
<skos:Concept  rdf:about="Territoires"
  <skos:prefLabel  xml:lang="fr" Territoires
  <skos:prefLabel  xml:lang="en" Territoires
  <skos:broader  rdf:resource="tetis_researchAxes_root"
  <skos:narrower  rdf:resource="Multifonctionnels"
  <skos:narrower  rdf:resource="Numeriques-apprenants"
```


Exemple : Axes recherche TETIS

```
*Processing Instruction* <?xml xml(version="1.0" encoding="UTF-8")
<rdf:RDF xmlns:dc="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept rdf:about="Données Hétérogènes"
  <skos:Concept rdf:about="ModDynSpatiales"
  <skos:Concept rdf:about="Teledetection"
  <skos:Concept rdf:about="Textes"
  <skos:Concept rdf:about="Science des Données"
  <skos:Concept rdf:about="Corridors Rivulaires"
  <skos:Concept rdf:about="Hydrosystèmes"
  <skos:Concept rdf:about="Paysages"
  <skos:Concept rdf:about="Biodiversité"
  <skos:Concept rdf:about="Numeriques-apprenants"
  <skos:Concept rdf:about="Multifonctionnels"
  <skos:Concept rdf:about="Territoires"
  <skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:ConceptScheme rdf:about="http://geonetwork-opensource.org/tetisAxesThematiques"
    <dc:title TETIS Research Axes, version 1.0 31102019
    <dc:description TETIS Research Axes thesaurus for GeoNetwork opensource
    <skos:hasTopConcept rdf:resource="tetis_researchAxes_root"/>
```

```
<skos:Concept rdf:about="tetis_researchAxes_root"
  <skos:prefLabel xml:lang="fr" tetis_researchAxes_root
  <skos:prefLabel xml:lang="en" tetis_researchAxes_root
  <skos:narrower rdf:resource="Territoires"
  <skos:narrower rdf:resource="Biodiversité"
  <skos:narrower rdf:resource="Paysages"
  <skos:narrower rdf:resource="Science des Données"
  <skos:narrower rdf:resource="ModDynSpatiales"
  <skos:narrower rdf:resource="Données Hétérogènes"/>
```

```
<skos:Concept rdf:about="Territoires"
  <skos:prefLabel xml:lang="fr" Territoires
  <skos:prefLabel xml:lang="en" Territoires
  <skos:broader rdf:resource="tetis_researchAxes_root"
  <skos:narrower rdf:resource="Multifonctionnels"
  <skos:narrower rdf:resource="Numeriques-apprenants"/>
```

« TETIS Research Axes » sur GN

▼ Mots clés - TETIS Research Axes, version 1.0 16102019

Biodiversité x

Rechercher dans le thésaurus ...

Concept en cours [tetis_researchAxes_root](#) ⬇

Concepts plus spécifiques

- + Biodiversité
- + Données Hétérogènes
- + Modélisation et Dynamiques Spatiales
- + Paysages
- + Science des Données
- + Territoires

Current concept [Territoires](#) ⬇

Parent concept

- + [tetis_researchAxes_root](#)

More specific concepts

- + Multifonctionnels
- + Numeriques-apprenants

« SENSORS TETIS » sur GN

SENSOR TETIS, version 1.0

Search ...

Search ...

Current concept [tetis_s](#)

More specific concepts

- + [HYPERSPETRAL](#)
- + [LiDAR](#)

Current concept [HYPERSPETRAL](#) +

Parent concept

- + [tetis_sensors_root](#)

More specific concepts

- + [PHYSICAL](#)
- + [SIMULATED](#)

Current concept [SIMULATED](#) +

Parent concept

- + [HYPERSPETRAL](#)

More specific concepts

- + [Hypxim 4m](#)
- + [Hypxim 8m](#)
- + [Hypspex 8m](#)
- + [Pleiades 1.6m](#)
- + [Sentinel2 1.6m](#)
- + [Sentinel2 9.6m](#)

Catalogue IDG TETIS

Search Map Contribute Admin console

TOPICS

INSPIRE THEMES

GEMET CONCEPTS KEYWORDS

- Agriculture (12)
- Environment (8)
- Forest conservation (11)
- Forest ecosystem (11)
- Remote sensing (26)

5 more

GCMD KEYWORDS

- + ATMOSPHERE (2)
- + HUMAN DIMENSIONS (10)
- + LAND SURFACE (27)
- + MODELS (11)
- + SPECTRAL/ENGINEERING (18)

1 more

REFERENTIELS NATIONAUX

- BD_PARCELAIRE (2)
- BD_TOPO (1)
- BDCARTO (3)

TETIS SENSORS KEYWORDS

- Tetis_sensors_root (17)
 - + HYPERSPETRAL (6)
 - + PHYSICAL (1)
 - SIMULATED (5)
 - Hypxim 8m (2)
 - Pleiades 1.6m (1)
 - Sentinel2 1.6m (1)
 - Sentinel2 9.6m (1)
 - + LIDAR (1)

WWF ECOREGIONS KEYWORDS

- WWF EcoRegionsRoot (6)
 - Mediterranean Forests, woodla... (2)
 - Temperate broadleaf and mixed... (4)

GN Facettes

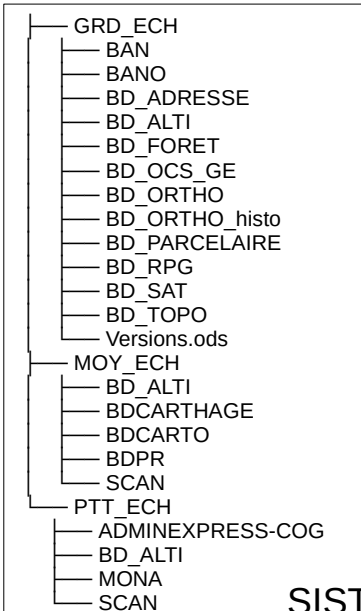
Implementation disponible pour:

CVS File
(2 niveaux)

Theme	Sous-theme
AGRICULTURE	AGRI_ENVIRONNEMENT
	AGRO_ALIMENTAIRE
	EXPLOITATION_ELEVAGE
	PARCELLAIRE_AGRICOLE
	SANTE_ANIMALE
	SANTE_VEGETALE
	ZONAGES_AGRICOLES
AIR_CLIMAT	CHANGEMENT_CLIMATIQUE
	METEO
	QUALITE_POLLUTION
AMENAGEMENT_URBANISME	ASSIETTE_SERVITUDE
	POLITIQUE_EUROPEENNE
	ZONAGES_AMENAGEMENT
	ZONAGES_ETUDE
	ZONAGES_PLANIFICATION
CULTURE_SOCIETE_SERVICE	ENSEIGNEMENT
	EQUIPEMENT_SPORTIF_CULTUREL
	ERP_AUTRE
	PATRIMOINE_CULTUREL

06/11/19

UNIX tree
output
(2 niveaux)



SIST 2019 - Toulouse 5-8 Nov 2019

Structured
TXT File

axesThematiquesTETIS.txt

<http://geonetwork-opensource.org/tetisAxesThematiques>
TETIS Research Axes, version 1.0
TETIS Research Axes thesaurus for GeoNetwork opensource
tetis_researchAxes_root
Territoires
Territoires Multifonctionnels
Territoires Numeriques-apprenants
Biodiversité
Paysages
Paysages Hydrosystèmes
Paysages Corridors Rivulaires
Science des Données
Science des Données Textes
Science des Données Teledetection
ModDynSpatiales Label:Modélisation
et Dynamiques Spatiales
Données Hétérogènes

Customisation of WEB
RDF files
(GCMD + GEMET By Groups)

Inputs :

- https://gcmdservices.gsfc.nasa.gov/kms/concepts/concept_scheme/sciencekeywords
- <https://www.eionet.europa.eu/gemet/latest/gemet.rdf.gz>

Outputs :

- GCMD_ForGN.rdf
- GEMETByGroups_ForGN.rdf

12

Codes sur dépôt gitlab IRSTEA

umr-tetis > IDG TETIS > GN RDF Factory > Details

GN RDF Factory

Project ID: 1500

8 Commits 1 Branch 0 Tags 3.4 MB Files

Python codes to generate Geonetwork compliant RDF Files

master gn-rdf-factory / +

History Find file Web IDE

fix gcmd script namespace bug
Claudia Lavalley authored 3 hours ago

README Add CHANGELOG Add CONTRIBUTING Add Kubernetes cluster Set up CI/CD

Name	Last commit
RDFPack	GCMD script finished
use_cases	fix gcmd script namespace bug
README.md	update README
init.py	initial Commit

GN_RDFFactory: python scripts to generate Geonetwork compliant RDF Files

<https://gitlab.irstea.fr/umr-tetis/idg-tetis/gn-rdf-factory>

Name	Last commit	Last update
..		
input_files	fix gcmd script namespace bug	3 hours ago
output_files	fix gcmd script namespace bug	3 hours ago
createCovadisRDF.py	initial Commit	1 month ago
createGCMDForGNFromNASARDFFile.py	fix gcmd script namespace bug	3 hours ago
createGEMETConceptsByGroups.py	GCMD script finished	3 days ago
createRDFFromStructureTXT.py	GEMETConceptsSchemeName fixed	2 weeks ago
createReferentielsRDF.py	GCMD script finished	3 days ago

Merci !