Service Public
Fédéral
FINANCES

PORTAIL DE TÉLÉCHARGEMENT

14/12/2021

Service Public
Fédéral
FINANCES

PLAN DE LA PRÉSENTATION

- Objectif
- Présentation de la solution développée
- Avantages et perspectives

OBJECTIFS

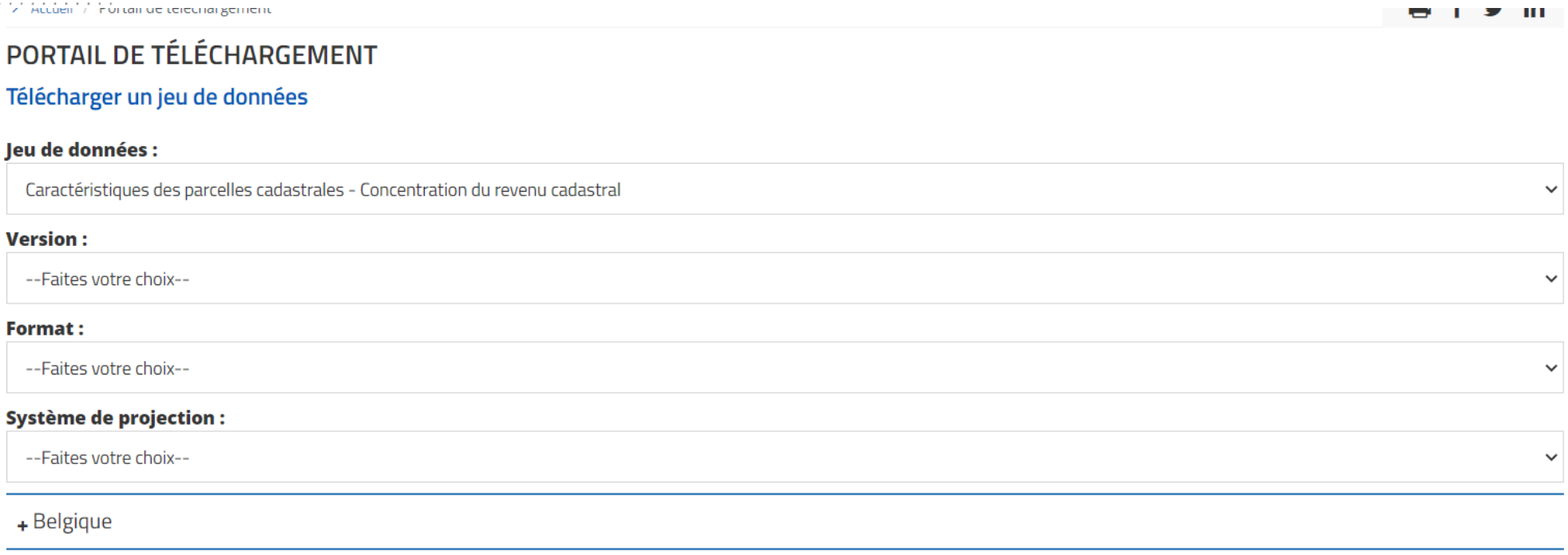
- Infrastructure de bulkdownload qualitative pour l'AGDP

- Infrastrcture qualitative

  - Publication automatique et quotidienne des fichiers téléchargeables
  - Pérennité des URL d'accès fichiers téléchargeables
  - URL de téléchargement dans fichiers *machine readable* standardisés (XML/JSON)
  - URL de téléchargement dans des pages web *human readable* (HTML)
  - Solution applicable à tous les jeux de données actuels et futurs

Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION



Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

Fichiers zippés

- Zips correspondant au jeu de données A
- Zips correspondants au jeu de données B

XML/JSON

- Liens vers jeu de données A
- Liens vers jeu de données B

HTML

- Liens vers jeu de données A
- Liens vers jeu de données

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

- Étape 1 : fichiers à publier placés dans le NAS interne
  - Fichiers zippés correspondant à un jeu de données (process métiers propres à chaque jeu de données)
  - Pdf de spécifications techniques
  - Thumbnails
  - Publication vers opendata.fin.belfgium (clone R)

- Étape 2 : génération des XML/JSON
  - Fichiers XML/JSON générés sur base des fichiers zippés présents dans le répertoire convenu (script python)
  - XML/JSON contient des informations relatives
    - Au jeu de données (titre, résumé, ...)
    - À la distribution précise (format, projection, version, ...)
  - Publication vers opendata.fin.belfgium

Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

gispadoc-public > OPENDATA > datasets >

Nom	Modifié le
8894f840-51ca-11eb-bfa6-3448ed25ad7c_20210101_csv_NA_01000.zip	01-10-21 09:20
AU-FiscSit_20200101_gml+xml_3812_01000.zip	27-05-21 15:38
AU-FiscSit_20200101_gml+xml_3812_1000.zip	27-05-21 15:38
AU-FiscSit_20200101_shp_3812_01000.zip	27-05-21 15:38
AU-FiscSit_20200101_shp_3812_1000.zip	27-05-21 15:38
AU-FiscSit_20200101_shp_31370_01000.zip	27-05-21 15:38
AU-FiscSit_20200101_shp_31370_1000.zip	27-05-21 15:38
AU-FiscSit_20210101_gml+xml_3812_1000.zip	27-05-21 15:38
AU-FiscSit_20210101_shp_3812_01000.zip	27-05-21 15:38
AU-FiscSit_20210101_shp_31370_01000.zip	27-05-21 15:38
PP-FiscSit_20200101_gml+xml_3812_01000.zip	27-05-21 15:38
PP-FiscSit_20200101_gml+xml_3812_11001.zip	27-05-21 15:38
PP-FiscSit_20200101_gml+xml_3812_11002.zip	27-05-21 15:38
PP-FiscSit_20200101_gml+xml_3812_11004.zip	27-05-21 15:38

https://opendata.fin.belgium.be/download/datasets/

Index of /download/datasets

Name	Last modified	Size	Description
Parent Directory		-	
8894f840-51ca-11eb-b..>	2021-10-01 14:24	240	
AU-FiscSit_20200101_..>	2021-10-01 14:24	0	
AU-FiscSit_20200101_..>	2021-10-01 14:24	0	
AU-FiscSit_20200101_..>	2021-10-01 14:24	0	
AU-FiscSit_20210101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	
PP-FiscSit_20200101_..>	2021-10-01 14:24	0	

Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

https://opendata.fin.belgium.be/download/.test/JSON/config_FR.json

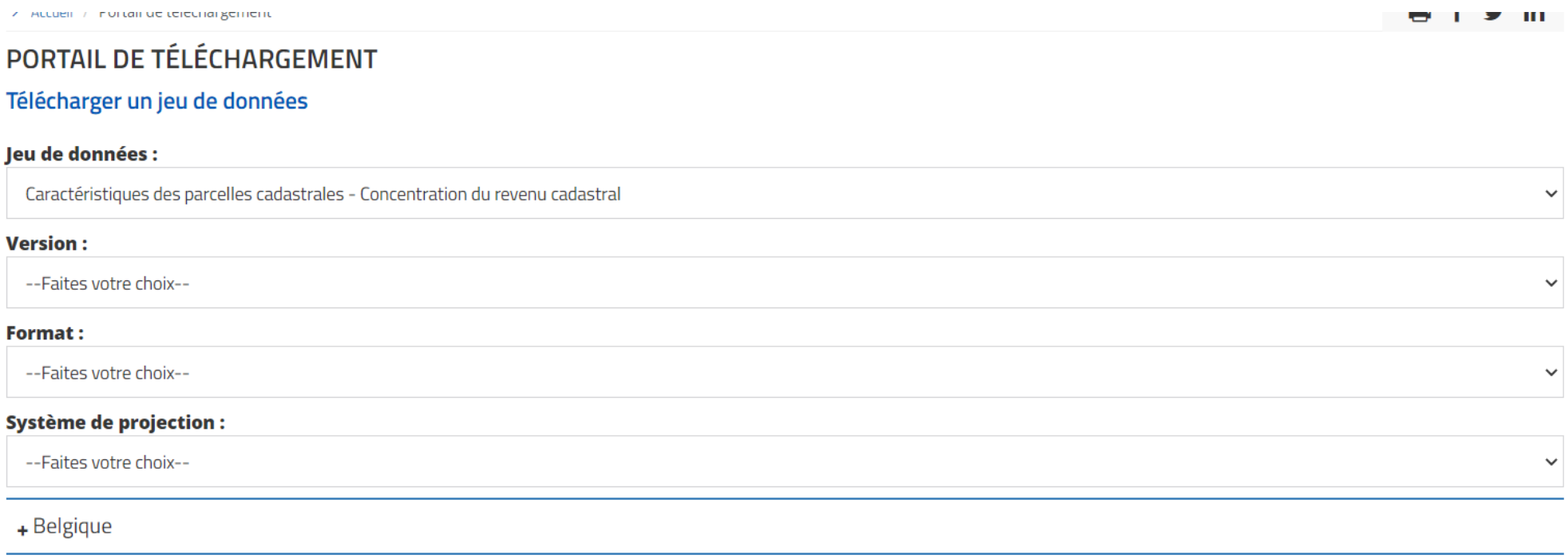

```
{
   "MetadataURL": "https://www.geo.be/catalog/details/8894f840-51ca-11eb-bfa6-3448ed25ad7c",
   "versionID": "20210101",
   "ProjectionName": "NA",
   "UUID": "8894f840-51ca-11eb-bfa6-3448ed25ad7c",
   "versionName": "01 janvier 2021",
   "Format": "csv",
   "DistributionLinks": [
      {
         "Province": {
            "PlaceName": "Belgique",
            "Links": [
               {
                  "Municipality": {
                     "URL": "https://opendata.fin.belgium.be/download/8894f840-51ca-11eb-bfa6-3448ed25ad7c_20210101_csv_NA_1000.zip",
                     "PlaceName": "Belgique",
                     "NISCode": "1000"
                  }
               }
            ],
            "NISCode": "01000"
         }
      }
   ],
   "EPSG": "NA",
   "NameFormat": "CSV",
   "Name": "Caract\u00e9ristiques des parcelles cadastrales - Concentration du revenu cadastral"
},
{
   "MetadataURL": "https://www.geo.be/catalog/details/629ad470-71dc-11eb-af47-3448ed25ad7c",
   "versionID": "20200101",
   "ProjectionName": "Lambert belge 2008 (EPSG 3812)",
   "UUID": "629ad470-71dc-11eb-af47-3448ed25ad7c",
   "versionName": "01 janvier 2020",
   "Format": "gml+xml",
   "DistributionLinks": [
```

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

- Étape 3 : génération des pages HTML
  - Création des pages HTML par les serveurs des sites concernés à partir des XML/JSON
    - Portail de téléchargement du SPF Finances généré à partir des JSON à partir d'un scrit JS
    - [Pages de téléchargement du portail INSPIRE générées à partir des XML
    - (pas de portions HMTL similaires sur geo.be/data.gov)]

Service Public Fédéral FINANCES

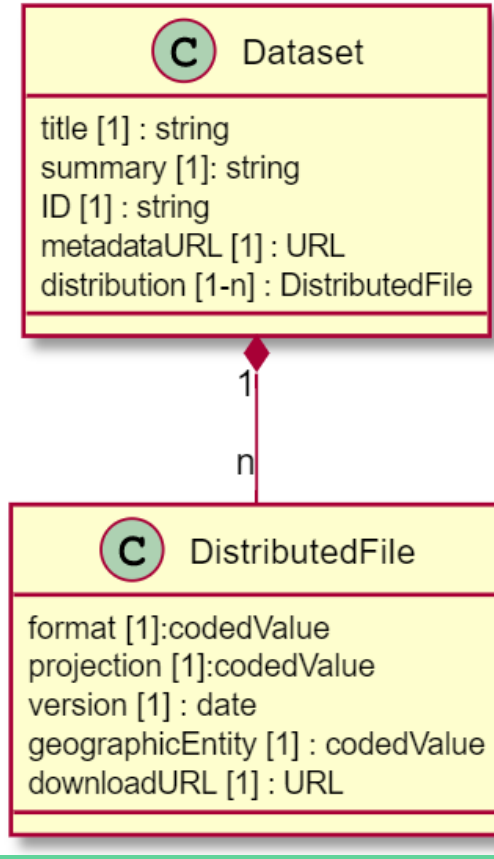
TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION



Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

- Génération des JSON/XML
  - JSON *ad hoc*
  - ATOM INSPIRE
- Deux types d'objets
  - Jeux de données → informations extraites de ISO 19139
    - Titre
    - Résumé
    - Fichiers téléchargeables
  - Fichiers téléchargeables → informations propres à chaque zip
    - Projection
    - Format
    - Version
    - Entité géographique



Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

```
class listDatasets():

    def __init__(self):
        self.folderdatasets =localisation_script+"\\"+"datasets"
        self.listdataset=os.listdir(self.folderdatasets)
        self.AdministrativeUnitsName=NISCodes('EN')

        for i in self.listdataset:
            if i.split('.')[1] !="zip":
                self.listdataset.remove(i)

        self.columns=['id', 'version', 'format', 'projection', 'section', 'uuid', 'zip', "NodeID",]   #prévoir ID et UUID car MEOW n'utilise pas les UUID
        self.rows=[]

    def _getuuid(self, id):
        if ID.has_key(id):
            self.uuid=ID[id]

    def listDistributions(self):
        for self.i in self.listdataset:
            self.zip=self.i
            self.i=self.i.split('.')[0].split(' ')
            self.id=self.i[0]
            self.version=self.i[1]
            self.format=self.i[2]
            self.projection=self.i[3]
            self.section=self.i[4] #NISCODE de la section géoémtrique
            if self.id in cswconnexion.list_records():
              self.uuid=self.id
            else :
                self. getuuid(self.id)
            self.AdministrativeUnitsName.GetUnitSup(self.section)
            self.nodeID=self.AdministrativeUnitsName.unitsup
            self.rows.append([self.id, self.version, self.format, self.projection, self.section, self.uuid, self.zip, self.nodeID])
        self.distributions=pandas.DataFrame(self.rows,columns = self.columns)
        self.versionsdataset=self.distributions.drop_duplicates(subset=['version', 'format', 'projection', 'uuid'])[['version', 'format', 'projection', 'uuid']]
        print self.versionsdataset
```

Service Public Fédéral FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

```
class ReadISO19139():
    def __init__(self, uuid, langue):
        self.uuid=uuid
        self.langue=langue
        self.iso19139=cswconnexion.get_record(self.uuid)
        self.MDRecord=MDRecord(self.iso19139)

    def readcoreMDRecord(self):
        self.name=self.MDRecord.select_tag_by_key("title", path_suffix_key='maintag')
        self.orgaMain=self.MDRecord.select_tag_by_key("title", path_suffix_key='maintag')


        self.title=self.MDRecord.select_tag_by_key("title", path_suffix_key='PT_FreeText', predicates=[self.langue])
        self.abstract=self.MDRecord.select_tag_by_key("abstract", path_suffix_key='PT_FreeText', predicates=[self.langue])
        self.orga=self.MDRecord.select_tag_by_key("RS_organisationName", path_suffix_key='PT_FreeText', predicates=[self.langue])
        self.mail=self.MDRecord.select_tag_by_key("RS_electronicMailAddress", path_suffix_key='maintag')
        self.rights=self.MDRecord.select_tag_by_key("Use_resourceConstraints",path_suffix_key='PT_FreeText', predicates=[self.langue])

    def readDatasetMDRecord(self):
        self.readcoreMDRecord()
        self.projection=self.MDRecord.select_tag_by_key("referenceSystemInfo", path_suffix_key='maintag')
        self.resourceCode=self.MDRecord.select_tag_by_key("resourceCode", path_suffix_key='maintag')
        self.codeSpace=self.MDRecord.select_tag_by_key("codeSpace", path_suffix_key='maintag')
        self.geobox=self.MDRecord.select_tag_by_key("geobox", path_suffix_key='maintag')
        self.graphicOverview=self.MDRecord.select_tag_by_key("graphicOverview", path_suffix_key='maintag')
        self.INSPIREConcepts=self.MDRecord.select_tag_by_key("thesaurus_descriptiveKeywords", path_suffix_key='maintag', predicates=["'INSPIRE feature concept dic
        self.INSPIRETheme=self.MDRecord.select_tag_by_key("thesaurus_descriptiveKeywords", path_suffix_key='PT_FreeText', predicates=["'GEMET - INSPIRE themes, ve
        self.FederalTheme=self.MDRecord.select_tag_by_key("thesaurus_descriptiveKeywords", path_suffix_key='PT_FreeText', predicates=["'Belgian Federal Thesaurus'
        self.INSPIREThemeMain=self.MDRecord.select_tag_by_key("title", path_suffix_key='maintag')
        self.FederalThemeMain=self.MDRecord.select_tag_by_key("title", path_suffix_key='maintag')
```

Service Public
Fédéral
FINANCES

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION


```
▼<gmd:identificationInfo xmlns:geonet="http://www.fao.org/geonetwork">
 ▼<gmd:MD_DataIdentification>
  ▼<gmd:citation>
   ▼<gmd:CI_Citation>
    ▼<gmd:title xsi:type="gmd:PT_FreeText_PropertyType">
       <gco:CharacterString>Land register plan - situation on January 1st</gco:CharacterString>
     ▼<gmd:PT_FreeText>
      ▼<gmd:textGroup>
         <gmd:LocalisedCharacterString locale="#EN">Land register plan - situation on January 1st</gmd:LocalisedCharacterString>
       </gmd:textGroup>
      ▼<gmd:textGroup>
         <gmd:LocalisedCharacterString locale="#FR">Plan parcellaire cadastral - situation au 1e janvier</gmd:LocalisedCharacterString>
       </gmd:textGroup>
      ▼<gmd:textGroup>
         <gmd:LocalisedCharacterString locale="#NL">Kadastraal percelenplan - toestand op 1 januari</gmd:LocalisedCharacterString>
       </gmd:textGroup>
      ▼<gmd:textGroup>
         <gmd:LocalisedCharacterString locale="#DE">Katasterplan - Stand 1. Januar</gmd:LocalisedCharacterString>
       </gmd:textGroup>
     </gmd:PT_FreeText>
    </gmd:title>
   ▼<gmd:date>
    ▼<gmd:CI_Date>
     ▼<gmd:date>
        <gco:Date>2020-12-15</gco:Date>
      </gmd:date>
     ▼<gmd:dateType>
        <gmd:CI_DateTypeCode codeList="http://standards.iso.org/iso/19139/resources/gmxCodelists.xml#CI_DateTypeCode" codeListValue="creation"/>
      </gmd:dateType>
    </gmd:CI_Date>
   </gmd:date>
   ▼<gmd:date>
    ▼<gmd:CI_Date>
     ▼<gmd:date>
        <gco:Date>2021-01-01</gco:Date>
      </gmd:date>
     ▼<gmd:dateType>
        <gmd:CI_DateTypeCode codeList="http://standards.iso.org/iso/19139/resources/gmxCodelists.xml#CI_DateTypeCode" codeListValue="revision"/>
      </gmd:dateType>
    </gmd:CI_Date>
   </gmd:date>
   ▼<gmd:identifier>
    ▼<gmd:RS_Identifier>
     ▼<gmd:code>
        <gco:CharacterString>64624e00-71dc-11eb-9a37-3448ed25ad7c</gco:CharacterString>
      </gmd:code>
     ▼<gmd:codeSpace>
        <gco:CharacterString>GAPD</gco:CharacterString>
      </gmd:codeSpace>
    </gmd:RS_Identifier>
   </gmd:identifier>
   ▼<gmd:identifier>
    ▼<gmd:RS_Identifier>
     ▼<gmd:code>
        <gco:CharacterString>PP-FiscSit</gco:CharacterString>
      </gmd:code>
     ▼<gmd:codeSpace>
        <gco:CharacterString>GAPD</gco:CharacterString>
      </gmd:codeSpace>
    </gmd:RS_Identifier>
   </gmd:identifier>
   ▼<gmd:series>
     <gmd:CI_Series/>
    </gmd:series>
  </gmd:CI_Citation>
 </gmd:citation>
 ▼<gmd:abstract xsi:type="gmd:PT_FreeText_PropertyType">
   <gco:CharacterString>Land register plan - situation on January 1st corresponds to the dataset of land register from the reference database for the land register data CadGIS. The land register plan is one of the elements that make up the patrimonial documentation. It is defined as “the graphic representation and assembling on a plan of all the cadastral
   territory”. A plan parcel, such as defined in Article 2 of the Royal Decree of 30 July 2018, is a part of Belgian territory, geographically delimited and identified by the General Administration of Patrimonial Documentation on the land register plan, which corresponds to the ground surface area of one or more patrimonial cadastral parcel(s) (real estate
   Article 472 of the 1992 Income Tax Code in respect of which the cadastral income is fixed). This dataset corresponds to the legal situation as defined on 1 January of the reference year. It is composed of thirteen geometric classes and one class without geometry. The first class includes the cadastral plan parcels present in the corresponding administr
   includes the buildings managed by the General Administration of Patrimonial Documentation, the third class includes the buildings managed by the regions, the fourth class includes equipment and tools, the fifth class includes water bodies, the sixth class includes cadastral blocks, the seventh, easements and paths, the eighth, toponymic lines, the ninth
   property boundaries, the eleventh, parcel numbers, the twelfth, cadastral divisions, and finally, the thirteenth, cadastral limits. The class without geometry corresponds to the table of tax situations. The dataset can be freely downloaded as a zipped shapefiles.</gco:CharacterString>
 ▼<gmd:PT_FreeText>
  ▼<gmd:textGroup>
     <gmd:LocalisedCharacterString locale="#EN">Land register plan - situation on January 1st corresponds to the dataset of land register from the reference database for the land register data CadGIS. The land register plan is one of the elements that make up the patrimonial documentation. It is defined as “the graphic representation and assembling on a
     parcels of the Belgian territory”. A plan parcel, such as defined in Article 2 of the Royal Decree of 30 July 2018, is a part of Belgian territory, geographically delimited and identified by the General Administration of Patrimonial Documentation on the land register plan, which corresponds to the ground surface area of one or more patrimonial cadas
     property as mentioned in Article 472 of the 1992 Income Tax Code in respect of which the cadastral income is fixed). This dataset corresponds to the legal situation as defined on 1 January of the reference year. It is composed of thirteen geometric classes and one class without geometry. The first class includes the cadastral plan parcels present in
     administrative unit. The second class includes the buildings managed by the General Administration of Patrimonial Documentation, the third class includes the buildings managed by the regions, the fourth class includes equipment and tools, the fifth class includes water bodies, the sixth class includes cadastral blocks, the seventh, easements and pat
     lines, the ninth, toponymic points, the tenth, property boundaries, the eleventh, parcel numbers, the twelfth, cadastral divisions, and finally, the thirteenth, cadastral limits. The class without geometry corresponds to the table of tax situations. The dataset can be freely downloaded as a zipped shapefiles.</gmd:LocalisedCharacterString>
```

TÉLÉCHARGEMENT DIRECT : PRÉSENTATION DE LA SOLUTION

- Méthode robuste
  - Indépendance opérationnelle vis-à-vis de l'IT
  - Mise à jour quotidienne
  - Développements relativement basiques
  - Aucune ressource business pour l'exécution quotidienne
  - Corrections faciles (corrections en cascade depuis la source)

- Méthode basée sur les standards
  - INSPIRE (fiches ISO 19139 +ATOM)
  - JSON
  - OOP friendly

- Web dynamique

Service Public
Fédéral
FINANCES

AVANTAGES ET PERSPECTIVES

- Valorisation du contenu des fiches ISO 19139
- Cohérence du contenu garantie
- Méthode bottom up : description suit la réalité >< méthode top down
  - JSON/XML produits en fonction des zip vraiment publiés
  - Garantie d'exactitude des URL
  - ATOM générés sur base des zip publiés (INSPIRE)
  - Instances de Distribution générée sur base des zip publiés (DCAT)
- Solution basée sur des standards : modèle de fonctionnement réplicable
  - Lecture de JSON/XML semblables pour générer du HTML
  - Geo.be / autres portails basés sur INSPIRE (ATOM)
  - Data.gov (Distribution)

Service Public
Fédéral
FINANCES

FAQ

- Questions/remarques

Benoit.fricheteau@minfin.fed.be