

Open Data Portals overleg

Datum : 10 september 2024, 14:00 – 15:10

Locatie: MS-Teams

Aanwezig:

- Agence du Numérique – ODWB
- Belgisch Platform Biodiversiteit
- BEP Namur
- DigitaalVlaanderen – Datavindplaats
- FOD Binnenlandse Zaken
- FOD BOSA – DG SD, BB
- FOD Financiën – AAPD
- FOD Mobiliteit – SD ICT
- FOD Sociale Zekerheid – DG Analyse & Monitoring
- FuturoCité
- Infrabel – SCPA
- IWEPS
- Kon. Belgisch Instituut Ruimte-Aeronomie
- Kon. Museum Natuurwetenschappen
- Kruispuntbank Sociale Zekerheid / Smals
- Nationaal Geografisch Instituut
- Paradigm Brussels – Datastore
- Sciensano
- Stad Antwerpen – Studiedienst
- Stad Gent – DI, District09
- Stad Namen
- Vlaamse Provincies – Data & Analyse

Bespreking

Project VLOED

De steden Gent en Brugge hebben samen het project “Voorspellen Lokale Ondernemers en Economie door Drukke” (VLOED) gerealiseerd. Met dit project wil men vooral de handelaars en horeca informatie verschaffen over de te verwachte drukte, door op verschillende plaatsen metingen te doen en AI-modellen te trainen op allerlei datasets.

Dit is niet zo eenvoudig te voorspellen: de drukte is seizoensgebonden, en hangt ook af van evenementen (vb Gentse Feesten), maar metingen kunnen ook verstoord raken door allerlei storingen (vb de stekker van meetapparatuur die per ongeluk uitgetrokken wordt). In praktijk komt het erop neer dat elke datasets wel hiaten vertoont.

Ook is het mogelijk dat de AI-modellen niet goed presteren omdat ze getraind werden op afwijkende datasets, zodat bijvoorbeeld bepaalde types van bussen worden herkend maar andere niet.

Het resultaat van dit alles is dat men de drukte enkel voor de komende 2 à 3 uren correct kan inschatten. Door het toevoegen van historische gegevens en andere real-time bronnen (gegevens over het weer, anonieme data van mobiele operatoren...) kan dit uitgebreid worden tot 7 à 8 uur, maar dat zou dan continu onderhoud van de data vergen. De resultaten voor beide steden lopen ook uiteen: Gent heeft meer studenten, terwijl Brugge meer toeristisch is (minder voorspelbaar)

Zie ook: <https://stad.gent/nl/over-gent-stadsbestuur/stadsbestuur/wat-doet-het-bestuur/gent-internationaal/gesubsidieerde-projecten/vloed-kunnen-we-drukke-de-stad-voorspellen>

Portaal Brussels-Hoofdstedelijk Gewest

Het Brusselse data portaal werd vernieuwd. De lay-out werd verbeterd en er zijn meer details zichtbaar bij onder andere geo-data. De zoekmotor werd aangepast en linken naar technische specificaties werden toegevoegd. Om datasets te downloaden is het niet meer verplicht om aan te melden.

Zie ook: <https://datastore.brussels/web>

GBIF portaal

Het internationale portaal rond biodiversiteit werd vernieuwd. Het bevat ruim 100.000 datasets met in totaal meer dan 3 miljard waarnemingen van planten, dieren en schimmels. Het Belgische portaal bevat bijna 400.000 waarnemingen.

Het gaat hierbij vooral over waarnemingen door mensen, al worden de meldingen meer en meer geautomatiseerd (vb met camera's of microfoons die het uitzicht en/of geluid van dieren kunnen herkennen). Bij voorkeur wordt ook de GPS-locatie genoteerd, want de plaatsnaam is niet altijd bekend of niet precies genoeg (vb in een groot bos).

Onderzoekers kunnen deze waarnemingen koppelen met andere gegevens, maar dan buiten het portaal. Uitwisseling van gegevens tussen landen is niet altijd eenvoudig: de bestandsformaten zijn wel gekend, maar inhoudelijk kunnen termen en IDs verschillen.

Zie ook: <https://www.gbif.org>

GeoNetwork

Het aantal portalen dat CKAN als basis gebruikt lijkt af te nemen, in plaats daarvan lijken er meer gebruikers te zijn van de open source oplossing GeoNetwork of de cloud-based dienst van OpenDataSoft. Daarnaast zijn er ook nog Piveau en (Harvard) Dataverse.

GeoNetwork is vooral bedoeld om INSPIRE metadata te gebruiken, maar er is een (Geo)DCAT plugin beschikbaar. De Vlaamse Datavindplaats gebruikt GeoNetwork, er zijn gesprekken met het team van GeoNetwork om de DCAT(-AP) plugin te integreren in de "core" van GeoNetwork. Daarnaast zal er gekeken worden hoe in dit systeem de High-Value Datasets aangeduid kunnen worden, het Nederlandse GeoNovum werkt hier ook aan.

Zie ook: <https://geonetwork-opensource.org>

Varia

Gezien steeds meer bedrijven gegevens willen hergebruiken of (moeten) publiceren, is het nuttig dat de beheerders van de verschillende data portalen elkaar hiervan op de hoogte houden, om ervaringen uit te wisselen en dubbele aanvragen te vermijden.

Open data evenementen

November : OpenBelgium conferentie, georganiseerd door OpenKnowledge Belgium (nog onder voorbehoud)