

Delen van milieudata met nieuwe technologie

Wies Vullings, Jandirk Bulens

2 oktober 2019



Agenda

- Introductie, Environmental Information systems (EIS)
 - Wat vinden we hiervan voor NL?
- 'State of play' in Europa, de evaluatie van EIS-sen per land en data delen
 - Moeten we data wel of niet zomaar delen?
- Nieuwe technologie, met de 'roadmap'
 - Hype of werkelijkheid?

Wat is een 'Environmental Information system'?

DEFINITION

An Environmental Information System (EIS) is defined as a coordinated assemblage of people, devices or other resources designed to exchange data or knowledge concerning any aspect of the ecosystem, the natural resources within or, more generally, the external factors surrounding and affecting human life.

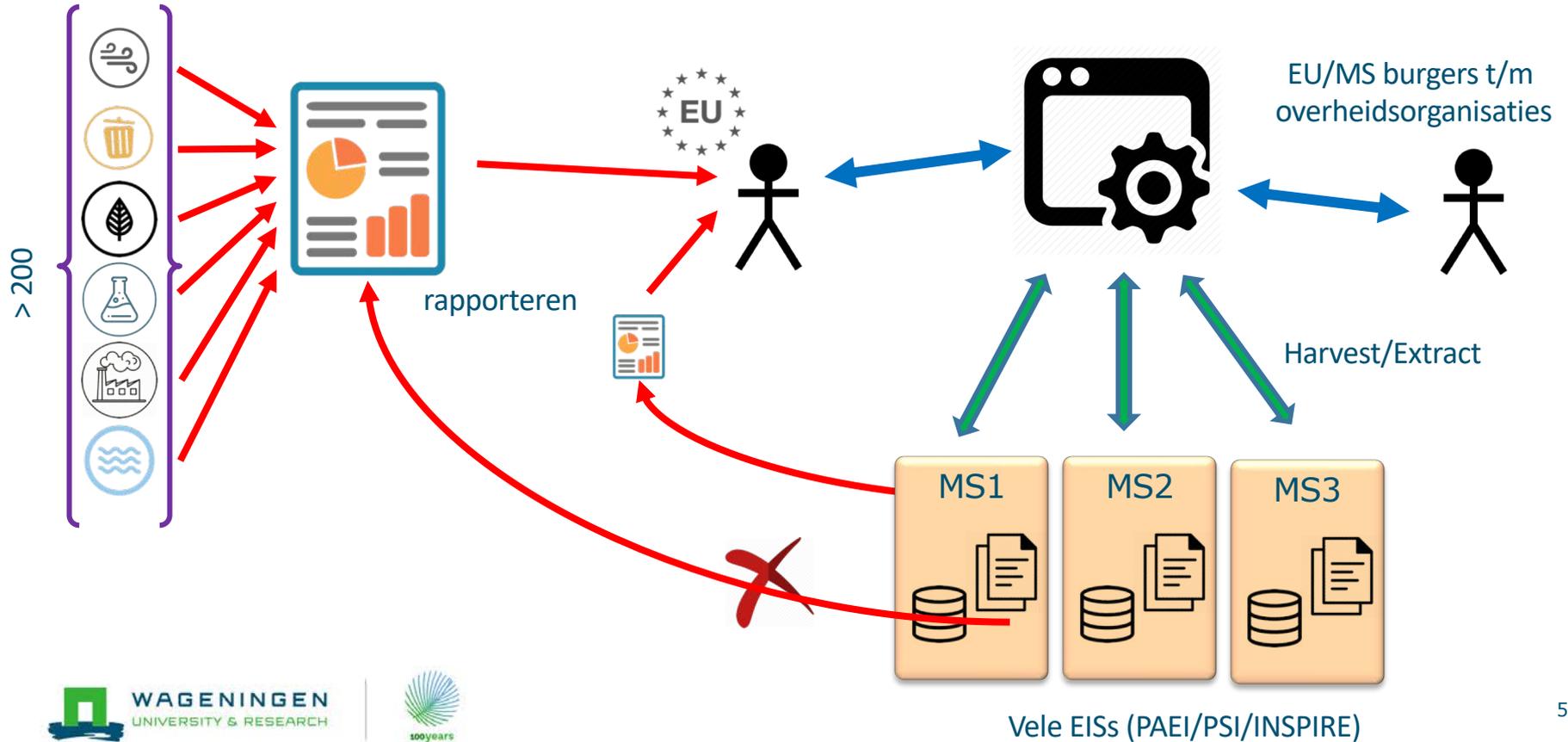
[source: GEMET, General Multilingual Environmental Thesaurus]



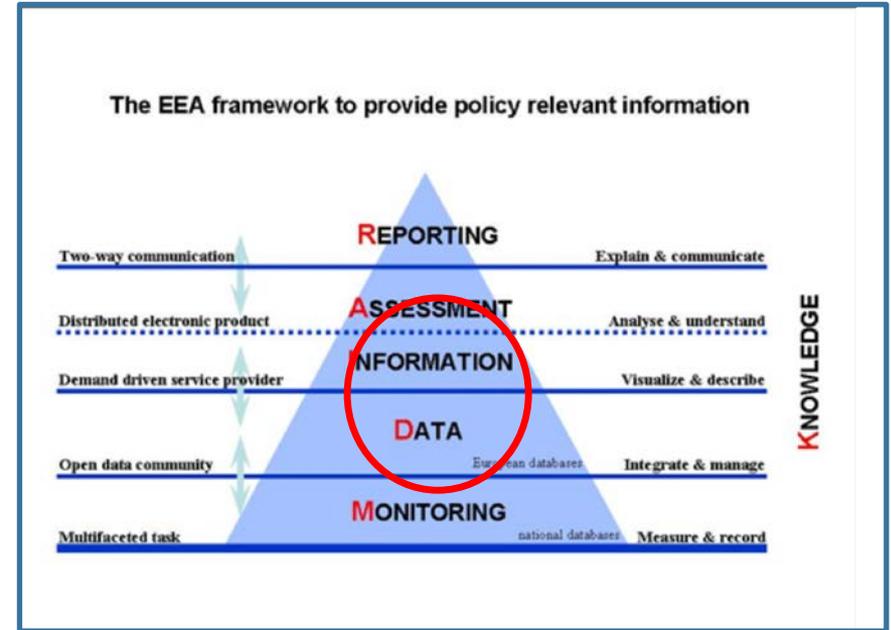
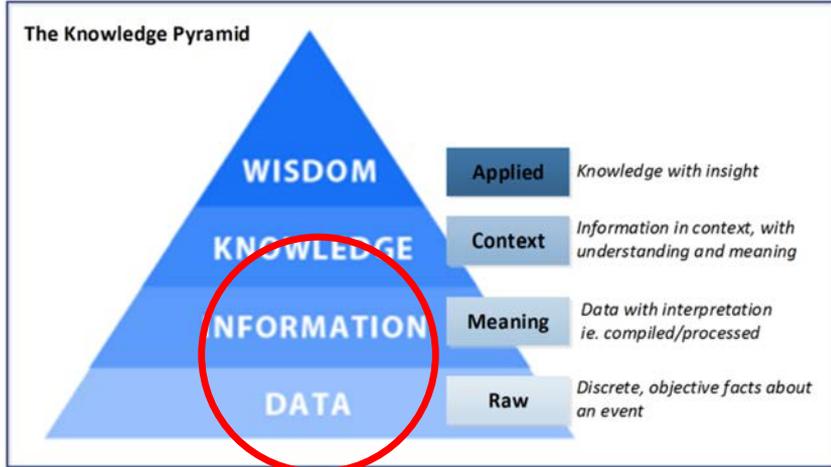
Doel:

De algemene doelstelling van de studie is het definiëren, identificeren en presenteren van 'Good Practice' uit EU- en nationale 'milieu-informatiesystemen' en portals die bijdragen aan een actieve verspreiding in de EU en de lidstaten door het verkennen, ontwikkelen en testen van tools om openbaar beschikbare gegevens te gebruiken om gegevens, geautomatiseerd, te verzamelen en te mijnen.

Project rationale



Governance van informatie

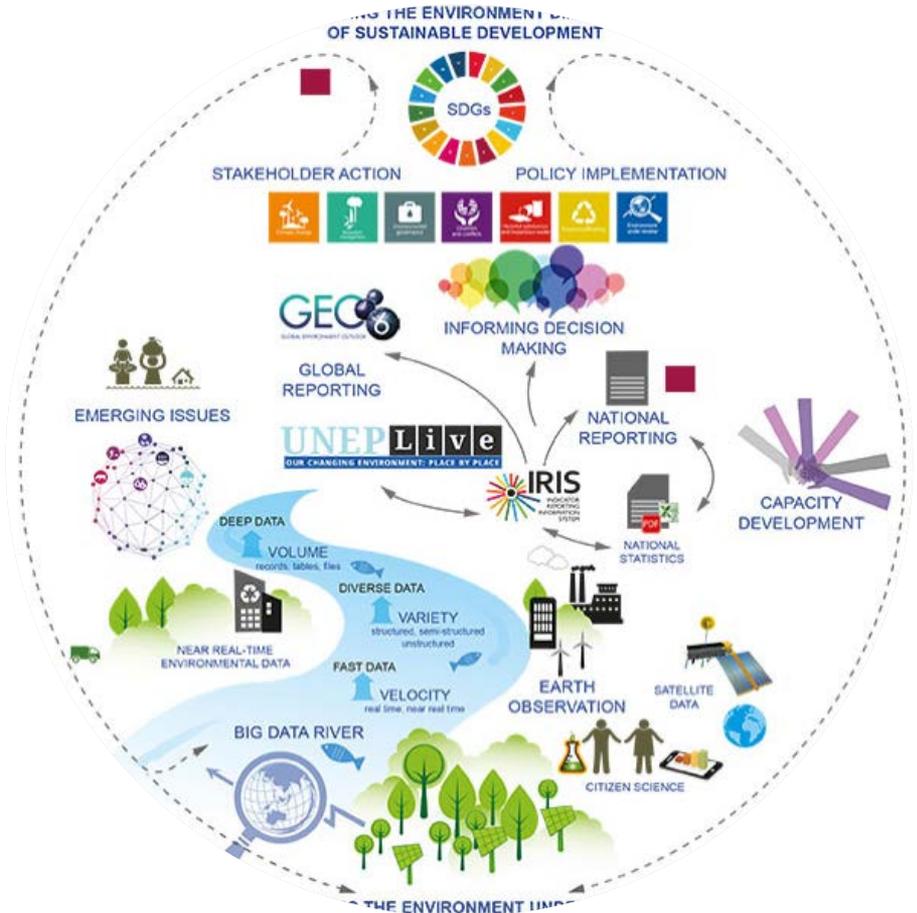


Discussie stelling 1

Zijn EIS-sen, milieuportalen
samenhangend georganiseerd
in Nederland?

A) Ja

B) Nee



Inventarisatie

EIS Germany

EIS Data - EIS Germany



| Latvia | | Map | GIS | Web | Mobile | API | Other |
|---------------------------------------|--|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| NATIONAL GEOPORTAL(S) | Ministry of Environmental Protection http://www.varam.gov.lv | <input checked="" type="checkbox"/> |
| INSPIRE GEOPORTAL | https://geolnija.lv | <input type="checkbox"/> |
| OTHER NATIONAL / REGIONAL EIS: | | | | | | | |
| | Nature Conservation Agency https://www.daba.gov.lv | | | <input checked="" type="checkbox"/> | | | |
| | Geospatial Information Agency http://map.lgia.gov.lv | | | | | | |

| Serbia | | Map | GIS | Web | Mobile | API | Other |
|---------------------------------------|--|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| NATIONAL GEOPORTAL(S) | Geosrbija http://www.geosrbija.rs National Open Data Portal http://data.gov.rs | <input checked="" type="checkbox"/> |
| INSPIRE GEOPORTAL | http://www.geosrbija.rs | <input type="checkbox"/> |
| OTHER NATIONAL / REGIONAL EIS: | | | | | | | |
| | Statistical Office Serbia http://www.stat.gov.rs | <input checked="" type="checkbox"/> | | | | | |

| Luxembourg | | Map | GIS | Web | Mobile | API | Other |
|---------------------------------------|--|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| NATIONAL GEOPORTAL(S) | Public.lu https://data.public.lu | <input checked="" type="checkbox"/> |
| INSPIRE GEOPORTAL | Geoport.lu https://www.geoport.lu | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| OTHER NATIONAL / REGIONAL EIS: | | | | | | | |



Germany

| Germany | | Map | GIS | Web | Mobile | API | Other |
|---------------------------------|--|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| GEOPORTAL(S) | Umweltbundesamt https://www.umweltbundesamt.de | <input checked="" type="checkbox"/> |
| INSPIRE GEOPORTAL | http://www.geoport.de | <input checked="" type="checkbox"/> |
| NATIONAL / REGIONAL EIS: | | | | | | | |
| | Geoportal Baden-Württemberg https://www.geoportal-bw.de | <input checked="" type="checkbox"/> |
| | GISU https://gis.uba.de | <input checked="" type="checkbox"/> |
| | Geoportal SH http://www.gis-sh.de | <input checked="" type="checkbox"/> |
| | Digitale Atlas Nord http://danord.gis-sh.de | <input checked="" type="checkbox"/> |
| | Geodata Portal Niedersachsen http://www.geodaten.niedersachsen.de | <input checked="" type="checkbox"/> |
| | GeoMIS Thüringen http://www.geoportal-th.de | <input checked="" type="checkbox"/> |
| | GeoPortal.MV http://www.geoportal-mv.de | <input checked="" type="checkbox"/> |
| | Geoportal NRW https://www.geoportal.nrw | <input checked="" type="checkbox"/> |
| | Germany Viewer http://www.geodatenzentrum.de | <input checked="" type="checkbox"/> |
| | GEF https://gef.bkg.bund.de | <input checked="" type="checkbox"/> |
| | Environmental Atlas of Bavaria http://www.umweltatlas.bayern.de | <input checked="" type="checkbox"/> |
| | Env-IT https://www.env-it.de | <input checked="" type="checkbox"/> |
| | Geoportal Wasser http://www.geoportal-wasser.rlp.de | <input checked="" type="checkbox"/> |
| | GEOTIS https://www.geotis.de | <input checked="" type="checkbox"/> |
| | GeoSeaPortal https://www.geoseaportal.de | <input checked="" type="checkbox"/> |
| | GGINA https://geoport.bafg.de | <input checked="" type="checkbox"/> |
| | GIS Umwelt im Saarland https://www.saarland.de | <input checked="" type="checkbox"/> |
| | Umweltportal des Landes NRW https://www.umweltportal.nrw.de | <input checked="" type="checkbox"/> |
| | Schadstoffemissionsportal https://www.thru.de | <input checked="" type="checkbox"/> |
| | Altbaufinder NRW http://www.altbaufinder-nrw.de | <input checked="" type="checkbox"/> |
| | ELWAS-WEB http://www.elwasweb.nrw.de | <input checked="" type="checkbox"/> |
| | Datenbank/-sammlung Naturschutz https://www.bfn.de | <input checked="" type="checkbox"/> |
| | Umgebungsärm in NRW http://www.umgebungslaerm.nrw.de/ | <input checked="" type="checkbox"/> |
| | emission of ionizing radiation https://www.imis.bfs.de | <input checked="" type="checkbox"/> |
| | Raumbeobachtung.de http://www.raumbeobachtung.de | <input checked="" type="checkbox"/> |
| | GeoSeaPortal https://www.geoseaportal.de | <input checked="" type="checkbox"/> |
| | Marine Geodateninfrastruktur https://www.mdi-de.org | <input checked="" type="checkbox"/> |
| | Statistisches Informationssystem https://www.destatis.de | <input checked="" type="checkbox"/> |



(s)

| Description | URL | Owner | Survey | Workshop |
|---------------|---|-----------------|-------------------------------------|--------------------------|
| Umwelt (GISU) | https://gis.uba.de/GIScatalogStart.de | Umweltbundesamt | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

(l)(s)

| Description | URL | Survey | Workshop |
|----------------------------|---|-------------------------------------|--------------------------|
| Landes Nordrhein-Westfalen | https://www.umweltportal.nrw.de | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

| URL | Owner | Survey | Workshop |
|---|-------------------|-------------------------------------|--------------------------|
| http://www.geoport.de/DE/GDI-DE/gdi-de.htm?lang=de | German Government | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

Print Exit

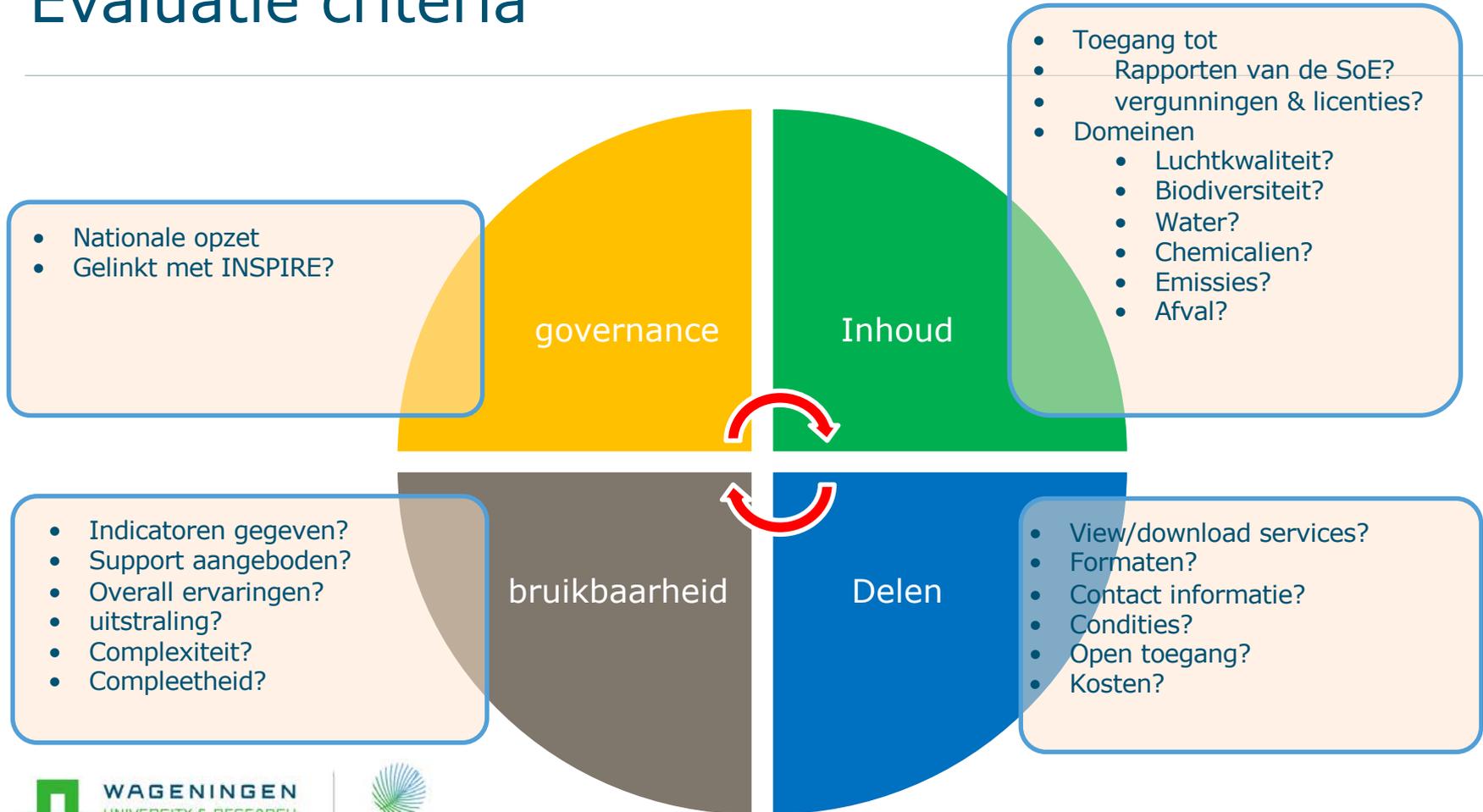
| Name | Description |
|------------------------------|--|
| DigitaleAtlasNord | Geodata portal |
| Geoportal Saarland | Geoportal |
| Altbaufinder NRW | |
| ELWAS-WEB | |
| Geoportal SH | Geodateninfrastruktur Schleswig-Holstein |
| Geoportal Sachsen | Geoportal |
| Geodata Portal Niedersachsen | Integrated use of distributed geodatabases of the Land |

Evaluatie criteria

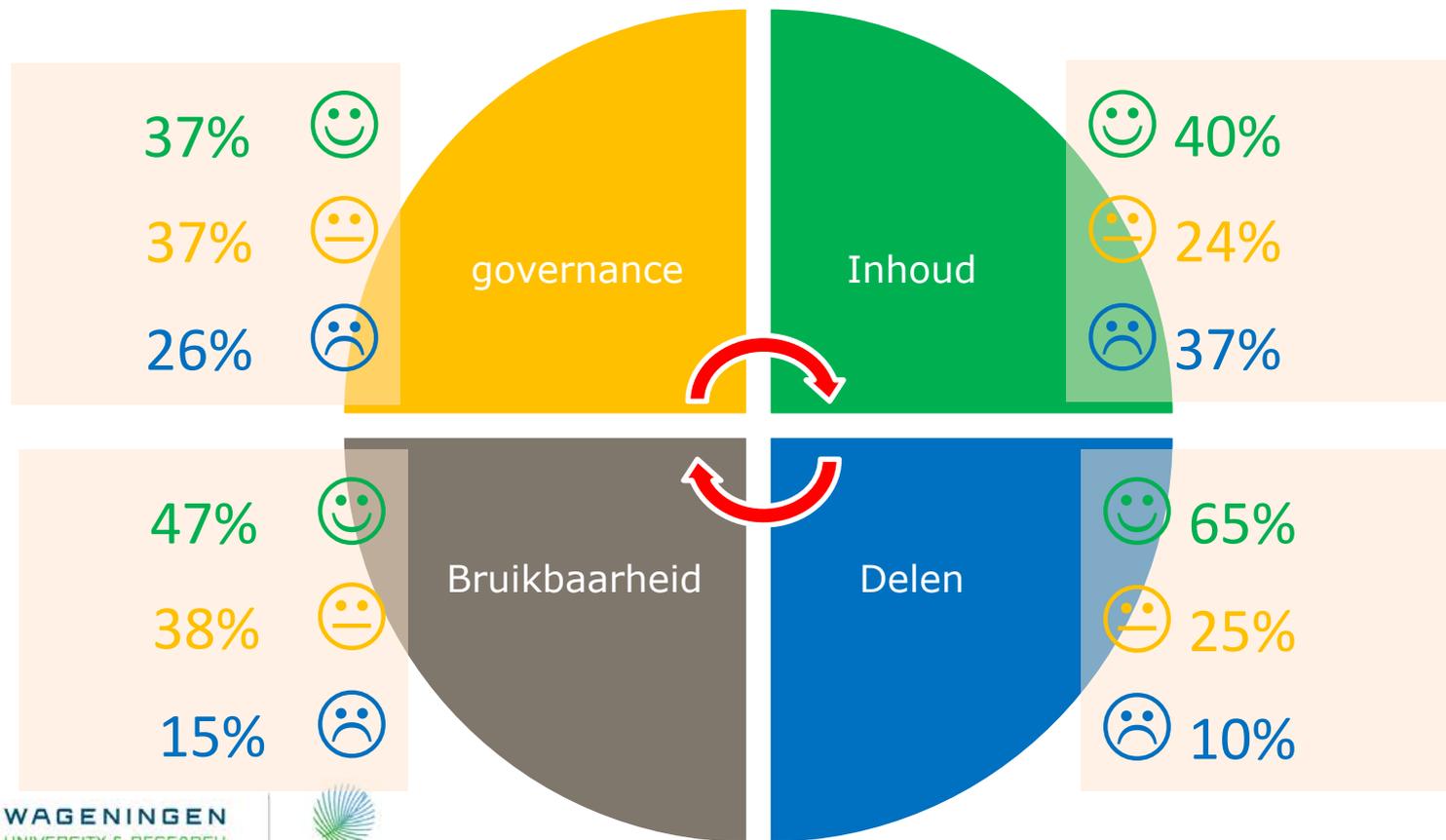
- 4 categorieën
 - Governance (2 criteria)
 - Inhoud: 6 milieu domeinen (content, 50 criteria)
 - Delen (sharing, 6 criteria)
 - Bruikbaarheid (usability, 10 criteria)

Er zijn 68 criteria gedefinieerd op basis van eerdere rapporten zoals de Fitness Check, SEIS-kookboek enz.

Evaluatie criteria

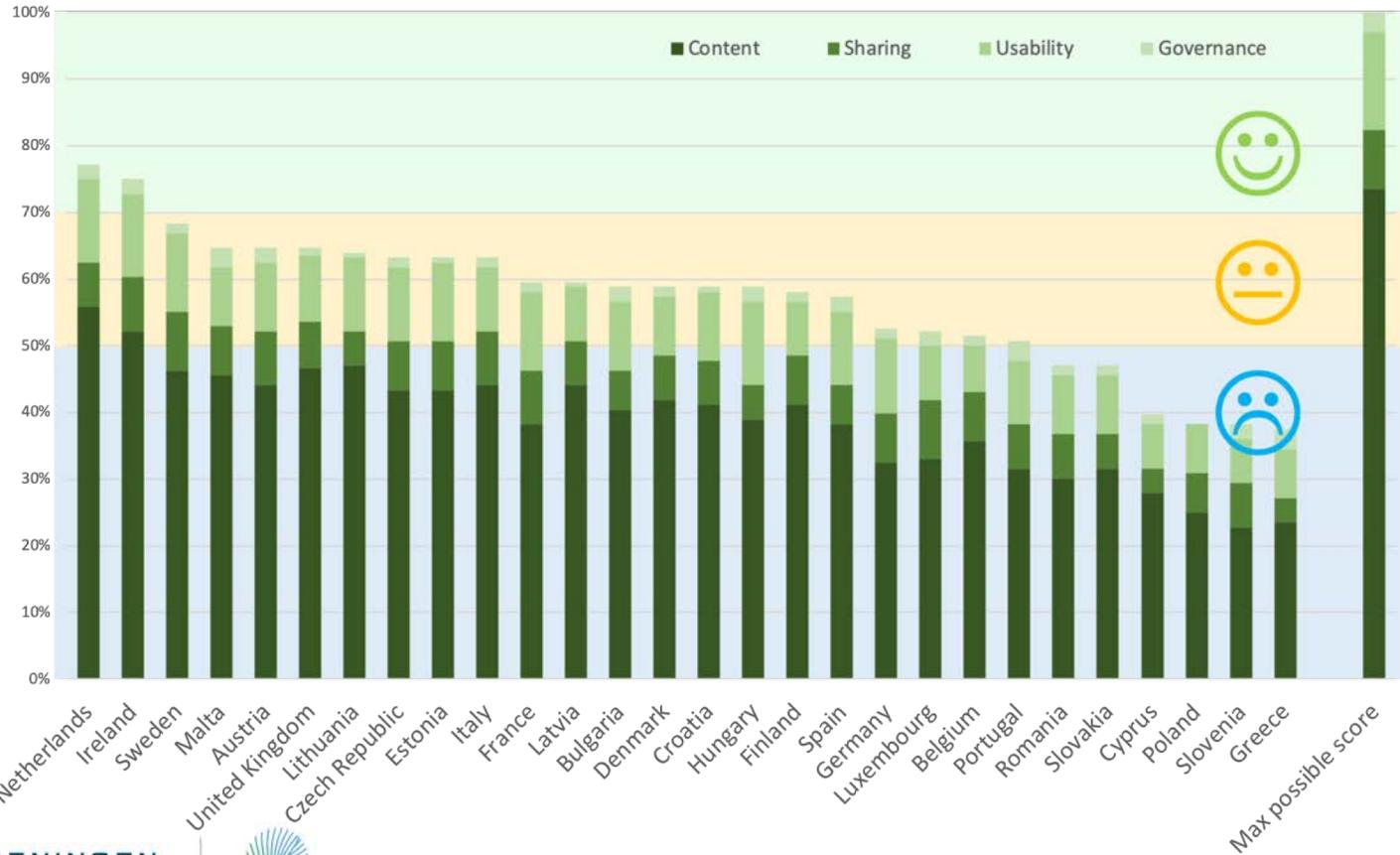


Algemeen overzicht score Europa



Scores per category and country

(all 68 criteria)



Richtlijnen 'good practice'

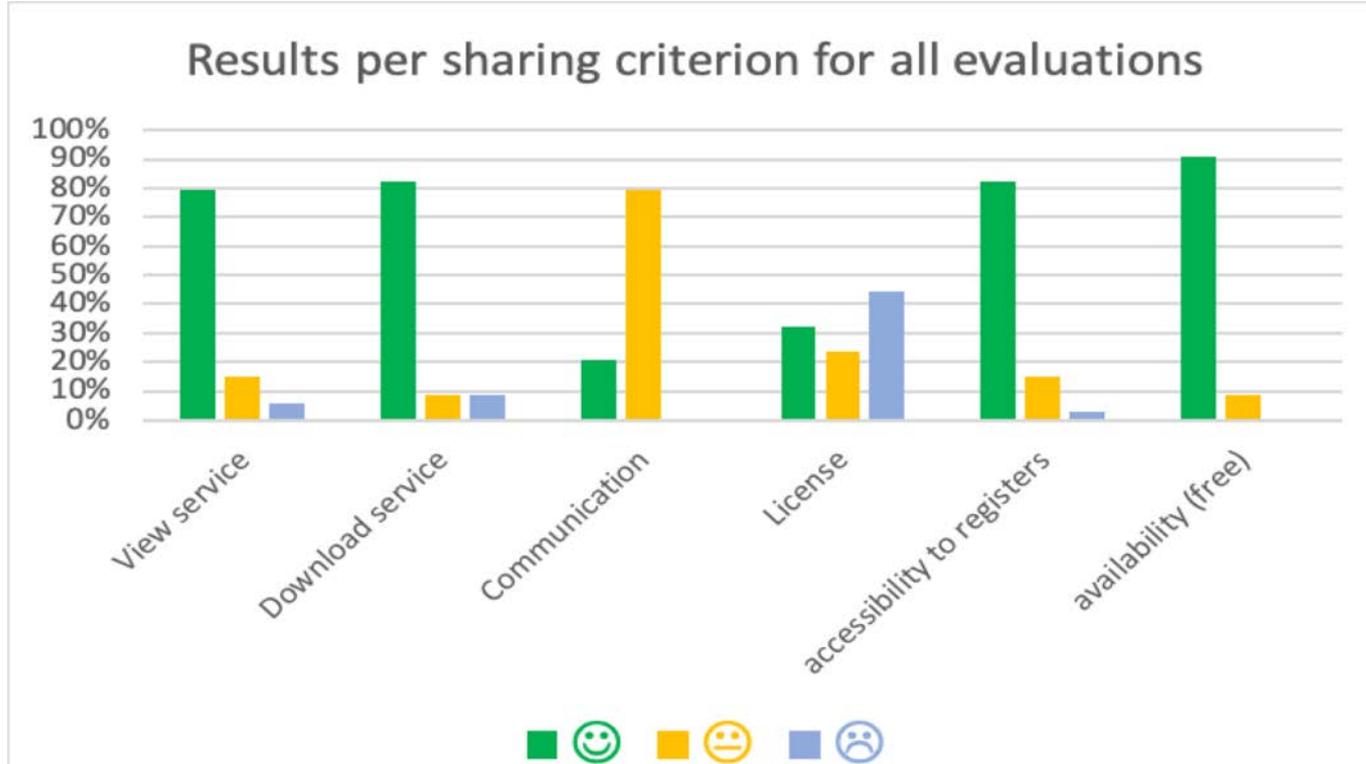
- Aanbevelingen hoe een goed portaal opgezet kan worden
- Methode voor evaluatie, matrix met criteria aanpak



Data delen

| Criterion | |  |  |  |
|------------|---|--|---|---|
| 3.1 | Is it possible to view the data? | Presence of view services | presence of links to view services | Not found |
| 3.2 | Is it possible to download the data in most commonly used file formats? | Downloads in common file formats possible | - | Not found |
| 3.3 | What communication possibilities are available? | Possibility to communicate with the owner/administrator or with other users through interactive form | Possibility to communicate with the owner/administrator (e-mail address or contact form) | Not found |
| 3.4 | Is the license to re-use data published? | Clear licensing for data re-use per dataset | Clear licensing policy for data re-use not indicated for each dataset, but on an organisational level | Not found |
| 3.5 | Is access to registers or lists open? (catalogues, thesaurus, vocabularies) | Access to registers is open | Access to registers only to registered users | Not open |
| 3.6 | Is environmental data freely available (free of charge)? | Data is open free of charge | Data is reasonably charged: add INSPIRE definition | Data is available on a commercial basis |

Data delen

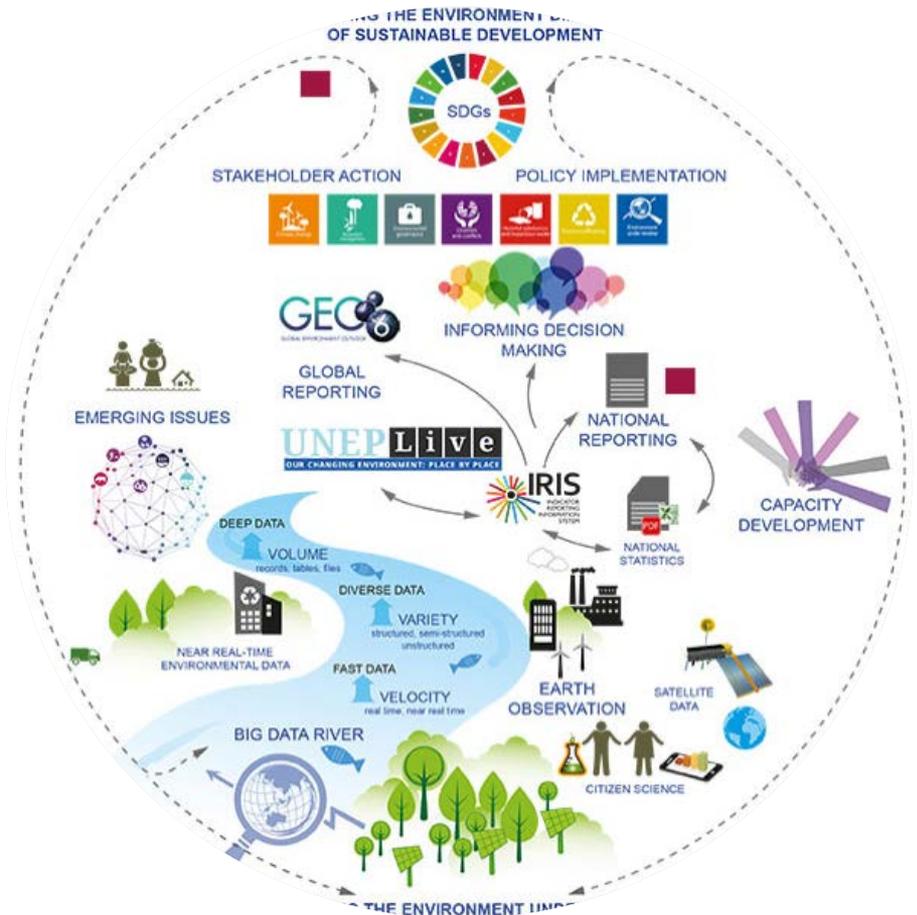


Discussie stelling 2

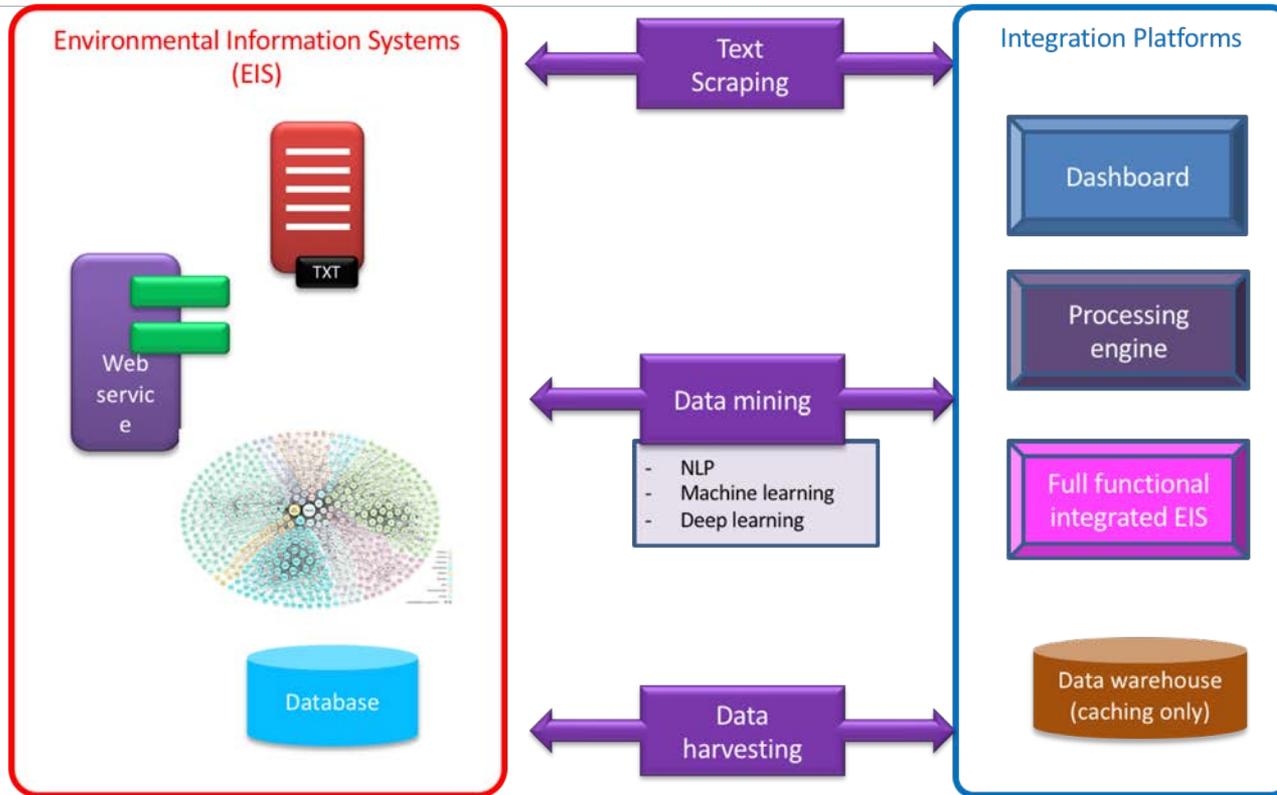
Data gerealiseerd met publiek geld (in beide gevallen)

A) Er zijn geen knelpunten, data delen moet altijd, tenzij dat wettelijk niet mag

B) Er zijn wel degelijk knelpunten, data delen kan niet zomaar.



Nieuwe technologieën



Nieuwe technologieën

- Data harvesting met
 - OGC webservices
 - RestFul web service
- Web crawling/web scraping
- Semantisch web/linked data
 - SPARQL endpoints
 - Microformats and RDFa
- Artificial Intelligence (AI)
 - Computational linguistics,
 - natural language processing (NLP)
 - language engineering
 - text mining
 - machine learning
 - Deep learning

SWOT analysis (AI)

- Computational linguistics,
- natural language processing (NLP)
- language engineering
- text mining
- machine learning

STRENGTHS

- Data does not have to be structured
- Can process large/huge amounts of data
- If there is a lot of data from various sources, useful information can be extracted in an automated way
- Possible correlations which have not been identified beforehand, can be discovered using these techniques

WEAKNESSES

- Needs a lot of computing power
- The relevance of the results is not always clear, the results might not always be meaningful

OPPORTUNITIES

- Most data available on the web is still unstructured, the techniques mentioned can make this data and information usable/reusable in ways which have never been thought possible
- Large datasets can be analysed without an initial hypothesis on possible correlations.
- This can lead to new (scientific) insights which could be used for the purpose of e-reporting

THREATS

- N/a

Technologiën – bevindingen 1

- Hergebruik van gegevens via services en gedistribueerde databases is vrij algemeen (IPCHEM)
- Hergebruik van ruimtelijke gegevens via OGC-webservices komt veel voor, maar meestal WMS (alleen voor weergave, niet opvraagbaar)
- Vooraf gedefinieerde formaten (\approx standaarden) voor gegevensbeschrijving en vergelijkbaarheid erg belangrijk (E-PRTR)
- De meeste tools zijn toegewijd aan een enkel doel, geen hergebruik van tool mogelijk, maar hergebruik van concepten wel

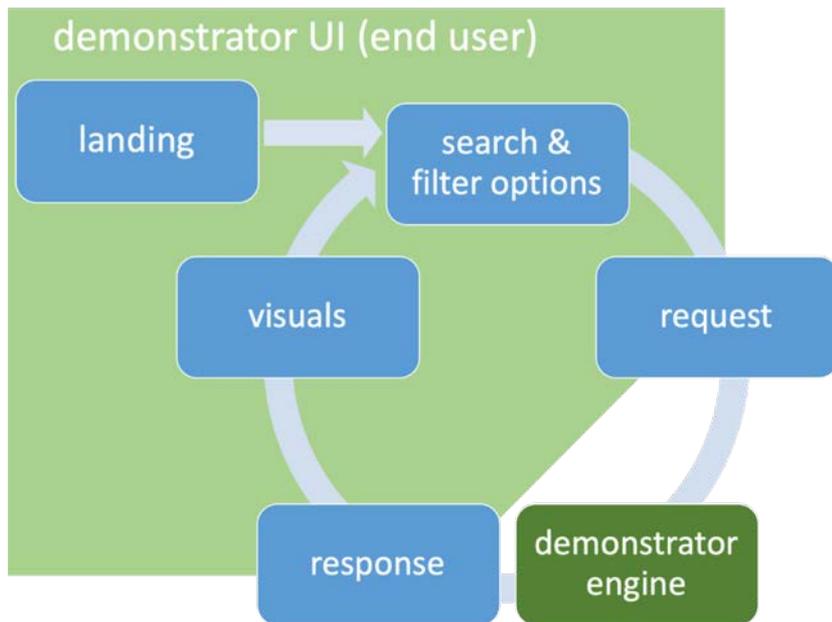
Technologiën – bevindingen 2

- Slechts 1 state-of-the-art (NLP / text scraping) geïdentificeerd 'harvestings' mechanisme (EMM)
- Gebruik van semantiek bij het beschrijven van data 'evoluerende' techniek. Gebruik van RDF- en SPARQL-eindpunten groeit
- Het leveren van inline metadata via annotaties in micro-formaat (JSON-LD, RDF-a) lijkt veelbelovend

Technologiën – aanvullend

- Het landschap is zeer divers, vooral de governance varieert sterk in de lidstaten
- Weinig ervaring met AI (EU-niveau)
- Het draait meestal om afstemming en harmonisatie
 - Juridisch, inhoud, governance, normen en technologie
- Meer ondersteuning is nodig !!

Road map



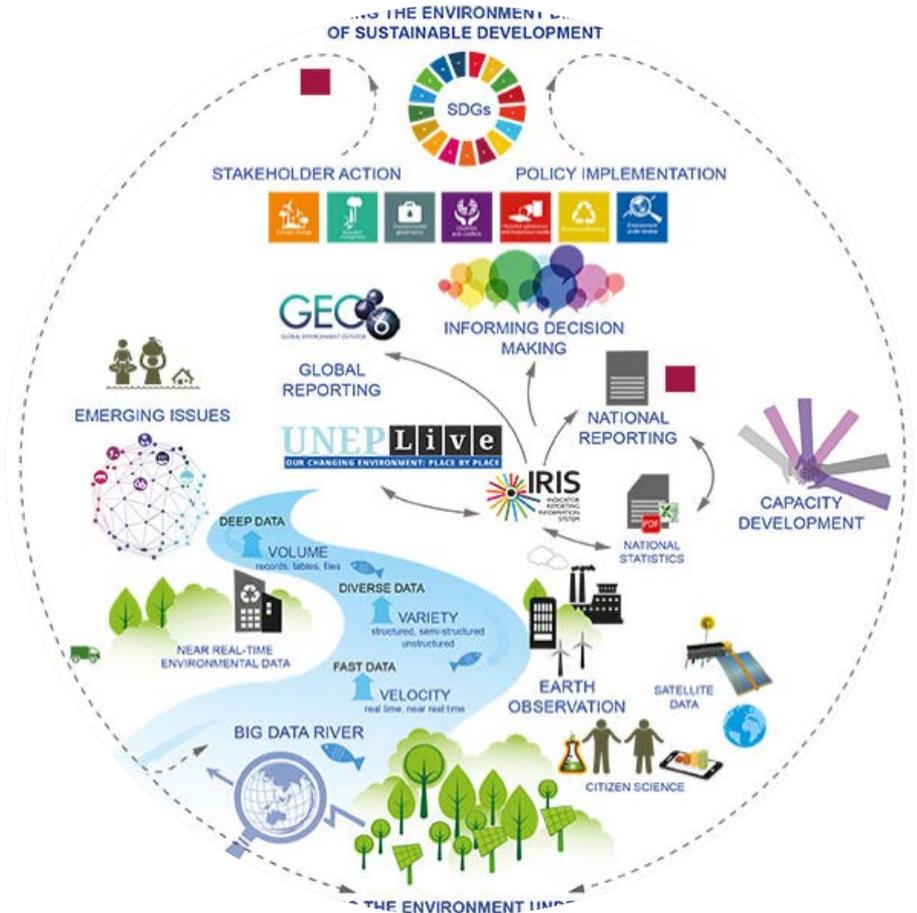
| | Starting | Learning | Ready to use | Experienced | Matured | Cost level EU | Cost level MS | Difficulty level | Time to mature |
|---|----------|----------|--------------|-------------|---------|---------------|---------------|------------------|----------------|
| Harvesting Toolbox | | | | | | | | | |
| <i>baseline measurement</i> | | | | | | | | | |
| H.1) End-user applications | █ | | | | | €€ | €€ | ★★ | P |
| H.2) Harvesting services | | | | | | €€€ | €€ | ★★ | PP |
| H.3) Catalogues | █ | █ | | | | € | €€€ | ★★★ | PP |
| H.4) Maintenance | █ | | | | | € | € | ★ | PPP |
| Technical | | | | | | | | | |
| T.1) Apply agreed technical standards | | | █ | | | € | € | ★★ | P |
| T.2) Data & information all machine readable | | █ | | | | €€ | €€€ | ★★★ | PP |
| T.3) Access sources Machine2machine | | █ | | | | €€ | €€ | ★★ | PP |
| T.4) Machine2machine authorisation | | | █ | | | €€€ | | ★★ | P |
| T.5) Useful metadata | | █ | | | | € | €€ | ★★★ | P |
| T.6) Validation tooling for QA | | █ | | | | €€ | € | ★ | P |
| Organisational | | | | | | | | | |
| O.1) Assign roles and responsibilities/ data stewardship | | █ | | | | € | €€ | ★★ | P |
| O.2) Provide collecting platforms | | | █ | | | €€€ | | ★★ | P |
| O.3) Create common vocabularies/ etc. (semantic standards) | | █ | | | | €€€ | €€ | ★★★ | PPP |
| Legal | | | | | | | | | |
| L.1) Align reporting obligations with all the (INSPIRE) data | | | █ | | | €€€ | €€ | ★★★ | PP |
| L.2) Accepting digital signatures to authenticate | | | | | | €€ | €€ | ★★ | PP |
| L.3) Provide provisions to accept also third party/citizen science/sensor/etc. data & information | | | | | | € | € | ★★★ | PP |
| Capacity building | | | | | | | | | |
| C.1) Provide guidance | | █ | | | | €€ | € | ★ | P |
| C.2) Introduce accreditation | | | | | | €€ | €€ | ★★★ | PPP |

Discussie item 3

We kijken 5 jaar vooruit:

Is nieuwe technologie:

- A) Een hype die overwaait
- B) Nee, wordt werkelijkheid



Bedankt!

Project resultaten op:

<https://www.eis-data.eu>

