

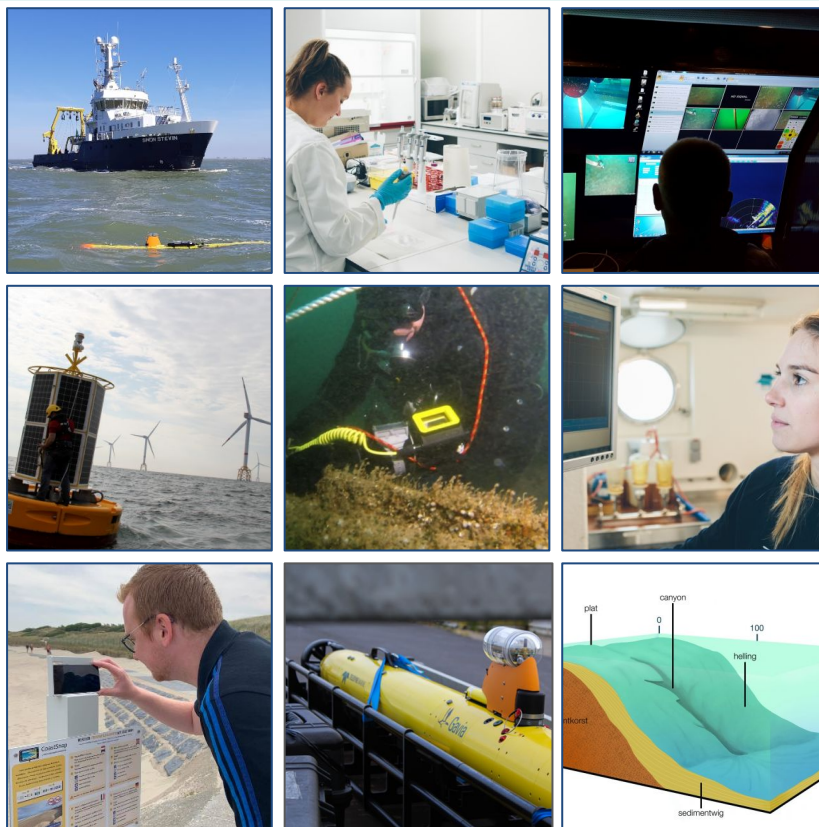


Bouwen aan de mariene informatiehub

Lennert Tyberghein
Datacentrum VLIZ

Het VLIZ versterkt de mariene kennisopbouw en excellentie van het marien onderzoek in Vlaanderen.

- Onderzoeksdomeinen: oceanen, zeeën, de kust en het intertidaal
- Doelgroepen:
 - Mariene onderzoeksgemeenschap
 - Educatieve instellingen
 - Breed publiek
 - Beleidsmakers
 - Blauwe economie



- Opgericht in 1999, operationeel sinds 2000
- Personeelsbestand: 154
- Leden: >1000 (filantropie)

Vlaams Marien Data Centrum

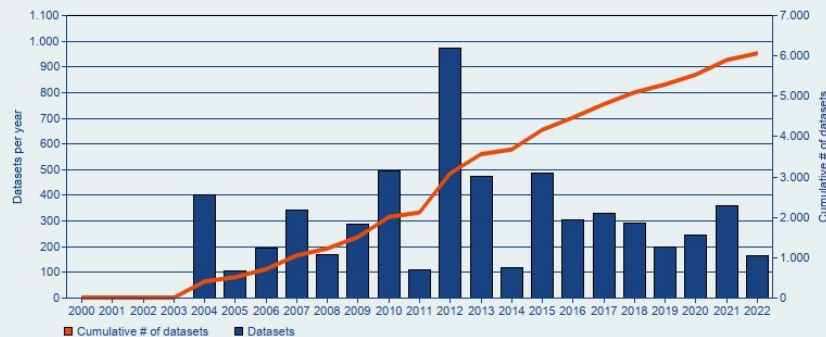
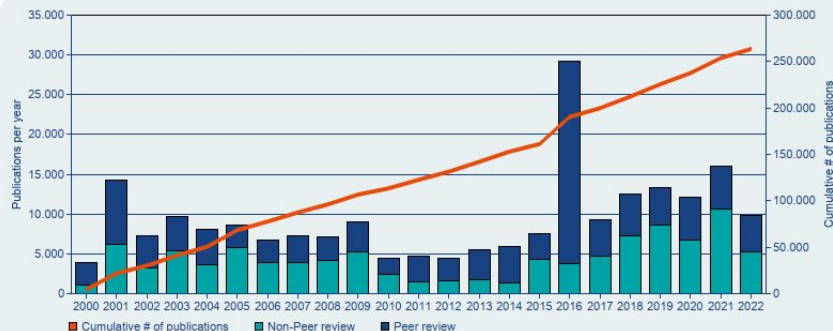
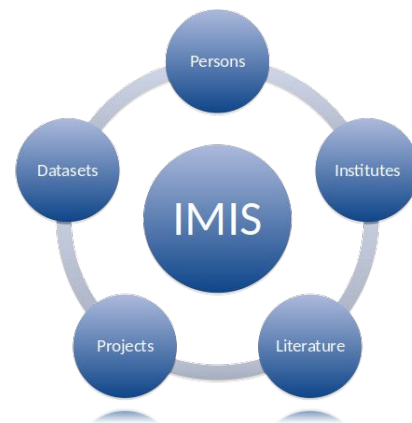
- Data management (biodiversiteit)
- Datasysteem design
- Data kwaliteitscontrole
- Data product ontwikkeling
- Kaart materiaal
- Real Time data monitoring
- Open Science
- Data beleid



Integrated Marine Information System (IMIS)

Beschrijven en linken van informatie

- Bewaren, beheren en publiek maken van metadata
- 8 gelinkte modules
- > 100 implementaties



Integrated Marine Information System (IMIS)

Historiek

Y 2000: oorsprong en Publicatie module

Y 2001: Personen module

Y 2003: Projecten en Instituut module

Y 2004: Dataset module

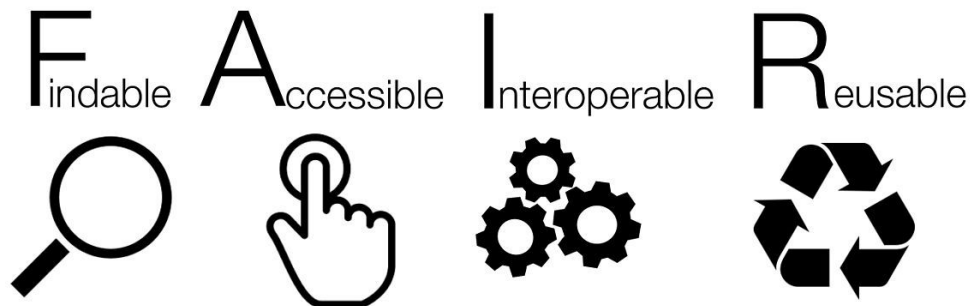
...

Y 2016: FAIR principles (Wilkinson paper)

Missie VLIZ 1999: coördinatie- en informatieplatform voor zeewetenschappelijk onderzoek in Vlaanderen.

- Thematische aanpak centraal
- Opzetten van [Belgische Mariene Bibliografie](#) en [Open Marien Archief](#)
- Verdere groei in kader van tal van (inter)nationale projecten en initiatieven

F.A.I.R principes zijn het middel, *maar wat is het doel?*



The Open Science challenge

general observations about achieving the goals hidden behind the F.A.I.R acronym

who is the actress playing the wife of stephen hawking



Alle Afbeeldingen Nieuws Video's Shopping Meer Tools

Ongeveer 7.810.000 resultaten (0,79 seconden)

Felicity Jones

Her performance as **Jane Hawking** in the 2014 biographical film *The Theory of Everything* garnered critical acclaim, earning her nominations for the Golden Globe Award, Screen Actors Guild Award, BAFTA Award and Academy Award for Best Actress.

...

Felicity Jones

Children

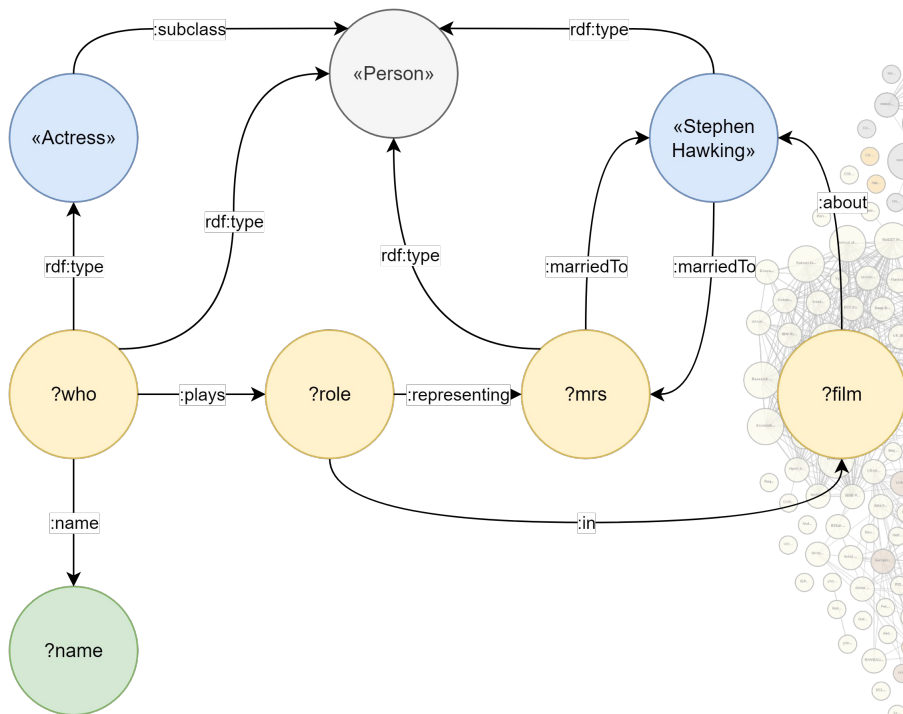
1

Historisch @y2000

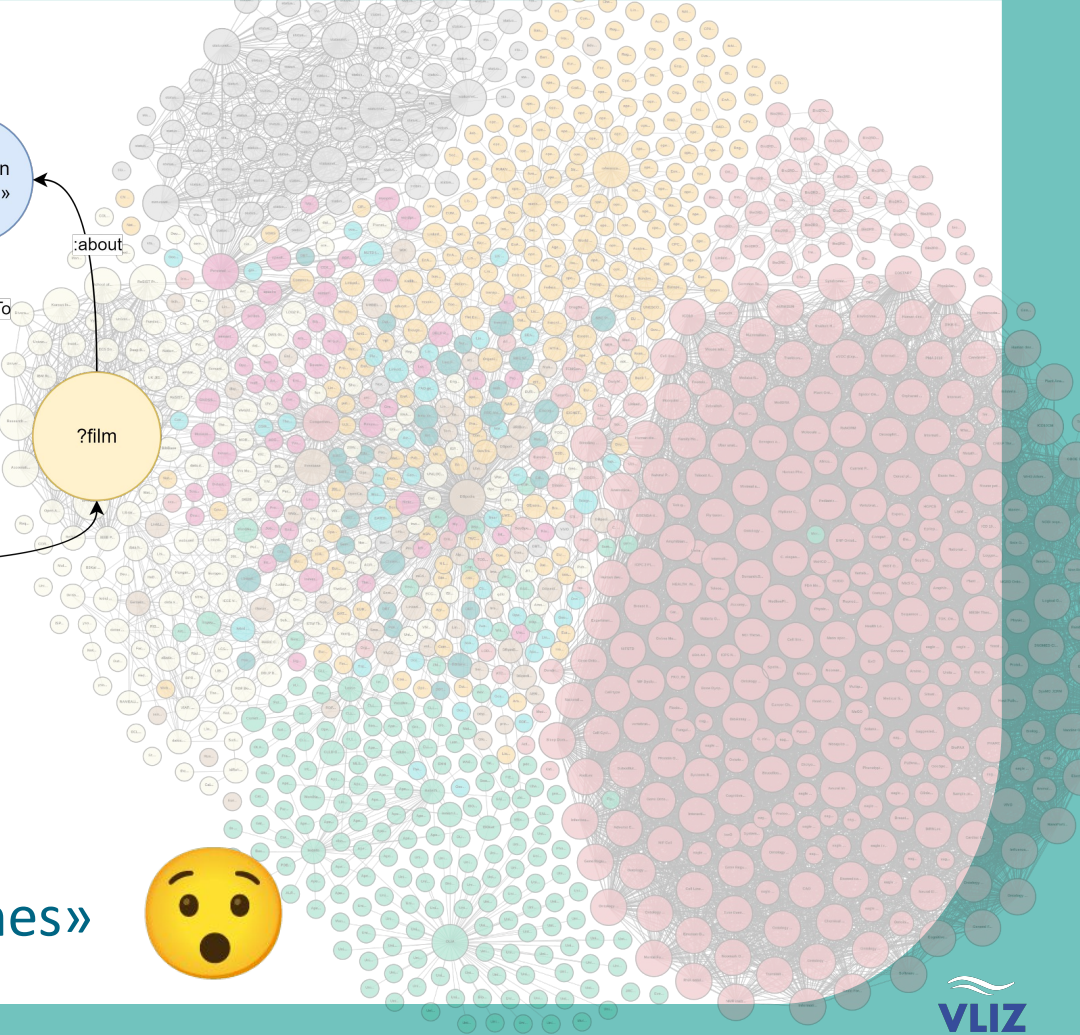
Aantal gerangschikte
textuele matches

Vandaag @2022

het «**answer**» gebaseerd op
de geaggregeerde
«**knowledge graph**»



«Felicity Jones»



Dus. Wat ALS?



Onderzoeksinformatie en -data ...

- ... gepubliceerd zouden worden als webpagina's
- ... zelf de semantiek zou bevatten
- ... bij zouden dragen aan de «global knowledge graph»

Zodat het vinden van antwoorden gebaseerd daarop ...

- ... zo gemakkelijk is als een web-search
- ... zouden spelen in dezelfde arena als (social) media

- ❑ Sea water temperatures, measured since 2020-01-01 in the whole 'North Sea'
- ❑ Publications about cod in the Belgian and Dutch EEZ between the years 2000 and 2020
- ❑ Datasets of researcher X using the FRRF on board of RV Simon Stevin
- ❑ Dates and conditions (cdt) of the last 5 sightings of the orca near the Belgian Coast
- ❑ ...

Motivatie

$$\#DataScientists \times Efficiency = Innovation Yield$$

What data scientists spend the most time doing

See 2016 [CrowdFlower Data Science Report](#)

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%



when



18%



Motivatatie & inspiratie

Open Science:: Hoe kunnen we de kost van (data) science verlagen?

- Het **volume en diversiteit** aan oceaan informatie en digitale resources **stijgt snel**
- Maar, geen lokale, nationale en regionale digitale systemen zijn nog altijd **niet gelinkt** op een duurzame manier.
- Op dit moment missen we de **‘digitale lijm’** om oceaan stakeholders met elkaar en met de data systemen te verbinden.
- Er is nood om de onafhankelijke digitale initiatieven te aligneren en hen te verbinden to een divers maar **geïnterconnecteerd data systeem**.

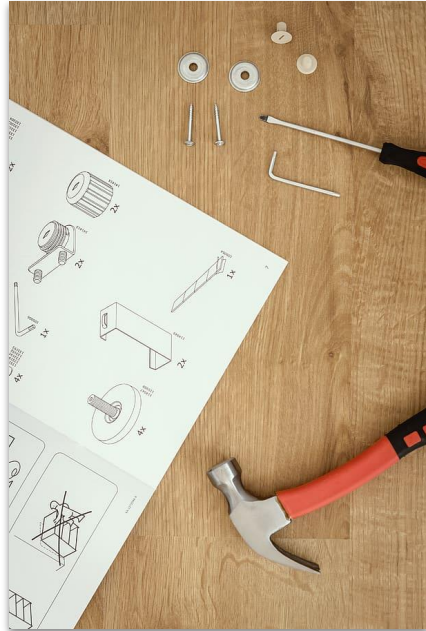


- Het web als collectieve hub
- Elke component is verantwoordelijk voor eigen bijdrage
- Linked Open Data en Semantische web technologie als drivers

**There is only one web.
Not a separate one for Science.**

Marc Portier

data, metadata and semantics



- **data** = *the stuff*
- **metadata** = *the stuff about the stuff*
 - provenance (who, where, when, what, how)
 - citation & attribution
 - general findability (keywords, some links)
- **semantics** = *stuff about what the stuff is about*
 - making it understandable = **interoperable**
 - making it machine-actionable
 - often overlooked
 - assumed self-explaining (to humans only)
 - historically no focus on this



«your new bed»

MALM

299,- €

190 x 210

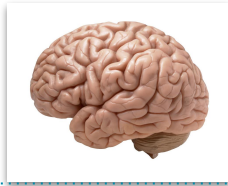
THE
INSTRUCTION
KIT





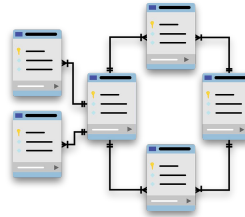
Verlies van semantiek == verlies aan kwaliteit

scientist & her data



❖ expert insight

local structured storage



❖ explicit semantics?
❖ standard semantics?
❖ implicit semantics?
❖ lost semantics?



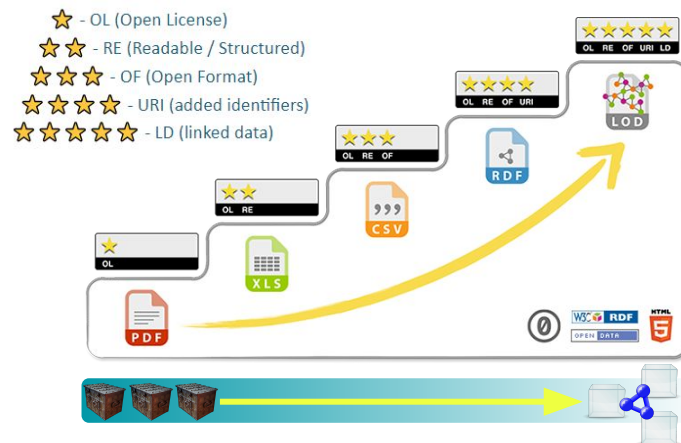
the unexpected outside view



❖ utterly clueless



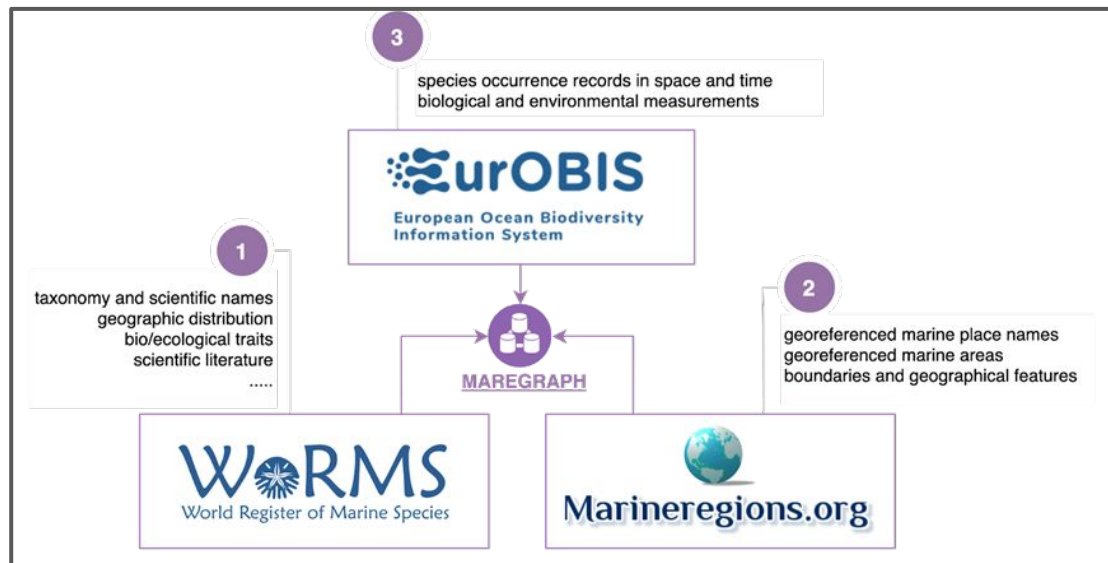
- Linked Open Data publicatie van IMIS informatie
 - Nieuwe weblinks volgens EU standaard
 - Persistent IDs
- Revamp front-end voor gebruikers (intern / extern)



TBL

MAREGRAPH*

towards an interoperable MARine knowledge GRAPH



- Verhogen van semantische, technische en legale **interoperabiliteit** van 3 datasets
- Open Linked Data publication
- OSLO track
- Opbouwen van de mariene data space als deel van de Vlaamse en Europese common **data spaces**

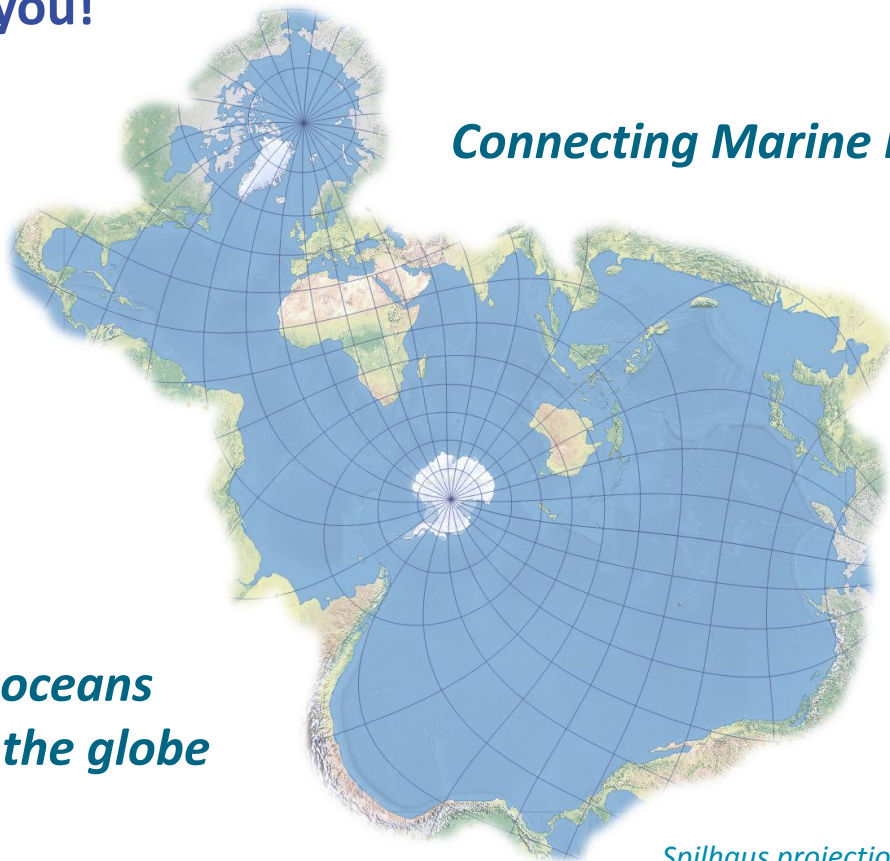
Thank you!

Connecting Marine Data



2021
2030

United Nations Decade
of Ocean Science
for Sustainable Development



*Like the oceans
connect the globe*

*Spilhaus projection
-- courtesy arcgis.com*

**Vacature:
Open Data Engineer**