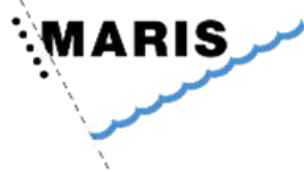


27-29 May 2024 



imdis

# International conference on Marine Data and Information Systems



27-29 May 2024 



International conference on **Marine Data**  
and **Information Systems**

ir. Marc Portier, VLIZ

**Using Data Fragments as the foundation  
for interoperable data access**

*about  
sustainable  
&  
interoperable*



credits

## Partners, projects and co-authors



*co-authors:* **MARIS (NL)** Peter Thijsse, Tjerk Krijger, Paul Weerheim - **BODC (UK)** Alexandra Kokkinaki, Gwen Moncoiffé - **Ifremer (FR)** Erwan Bodere

wake up inside FAIR-EASE

## on crossing (domain) boundaries and discovering "different species" of data / services / ....

- Did we not do this before / in that other project / last year?
  - e.g. ENVRI-FAIR, EOSC-LIFE, ...
- As we remix or extend groups...
  - we seem to redefine "harmonisation" each time
- What does this mean:
  - to the optimist – we are taking the next step towards ...
  - to the others – rearview graveyard
- But regardless of your level of optimism:
  - *Is this a sustainable approach?*



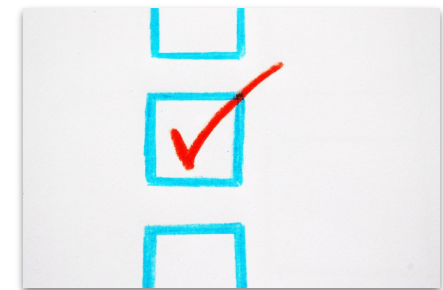
the myth behind

*my data is  
interoperable !*

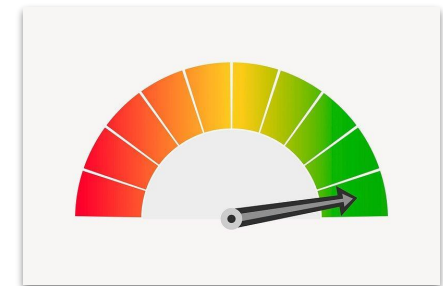
Interoperable



(that) Language makes us believe it is ...



... while smarter people will tell you it is ...



... both hiding the reality of data consumption:



## F.A.I.R. criticism

But we checked all the boxes !!!

- The data is accessible
- The license is open and permissive
- We minted the DOI
- We registered in the catalogue(s)
- The service(s) are available
- The API(s) are documented
- ...

What else do we need to do?

*Well, maybe we should do less?*



~ applying the "Kategorischer Imperativ" (Immanuel Kant) to engineering ?

$\Sigma$ -ing up?  $\Rightarrow$  What if we all keep doing X?

“ While the sum of many individual knowledge graphs is a bigger knowledge graph, the sum of many individual APIs is not a bigger API, but rather a mess of interactions in which neither the server nor the client can improve the situation much within the traditional framework of thinking about APIs. ”

$\Rightarrow$  *We do **\*not\*** need more APIs !*  $\Leftarrow$

– Ruben Verborgh,  
UGhent, IDLAB, KNoWS team



machine to machine vs machine to developer



## web-apis ignoring the hypermedia constraint

The paper / chapter (almost) nobody reads:

→ [https://ics.uci.edu/~fielding/pubs/dissertation/rest\\_arch\\_style.htm](https://ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm)

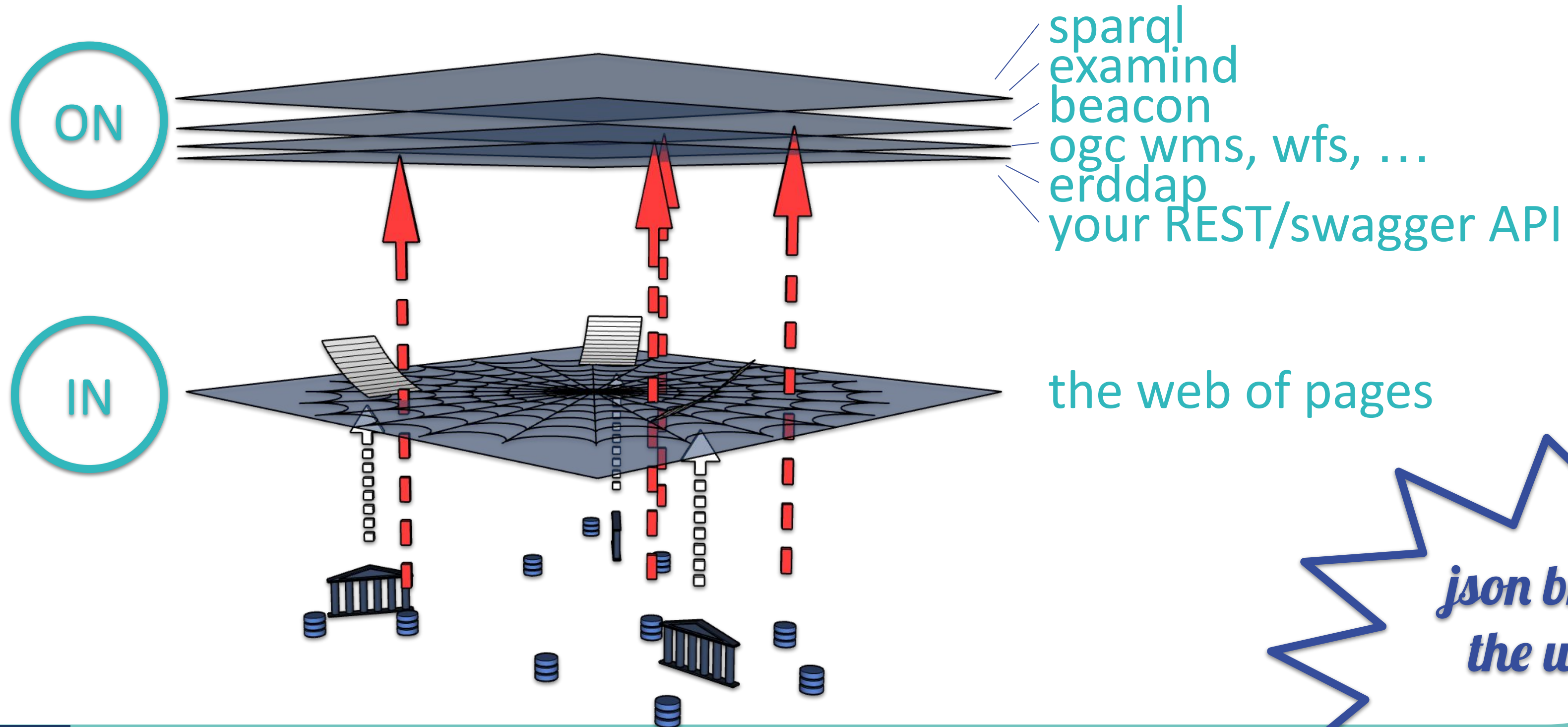
- from THE "Representational State Transfer (REST)" paper
- Roy Fielding, 2000 - PhD dissertation
- hypermedia-as-the-engine-of-application-state
- on (web) architectural style & "constraints"
- the hypermedia constraint: resource > representation > link > (repeat)

*one big data-space ~ a well knitted fabric*

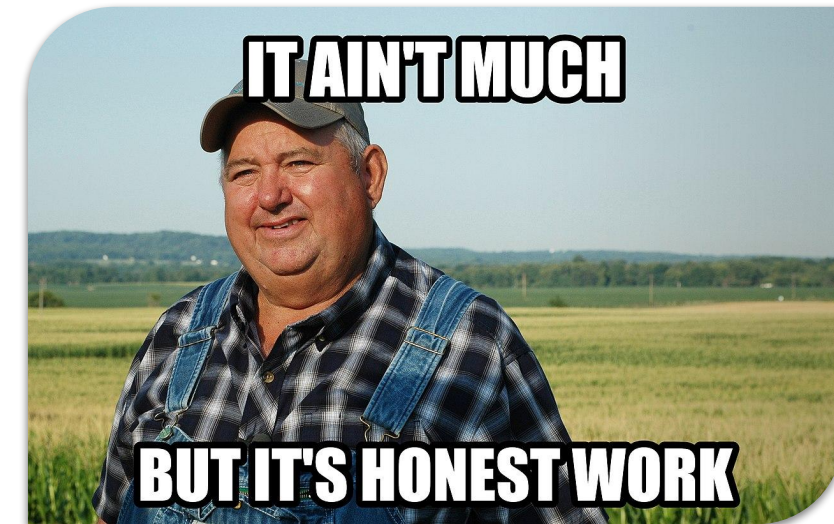


so this idea did not transition well:

## web of pages > web of data



## working the web



The web (HTTP) is not just about some data transport layer (like TCP/IP) that simply delivers your data / services to browsers out there

Instead it is a protocol-standards-and-principles-set, you need to embrace and **work WITH (IN) it, not (ON) TOP of it**

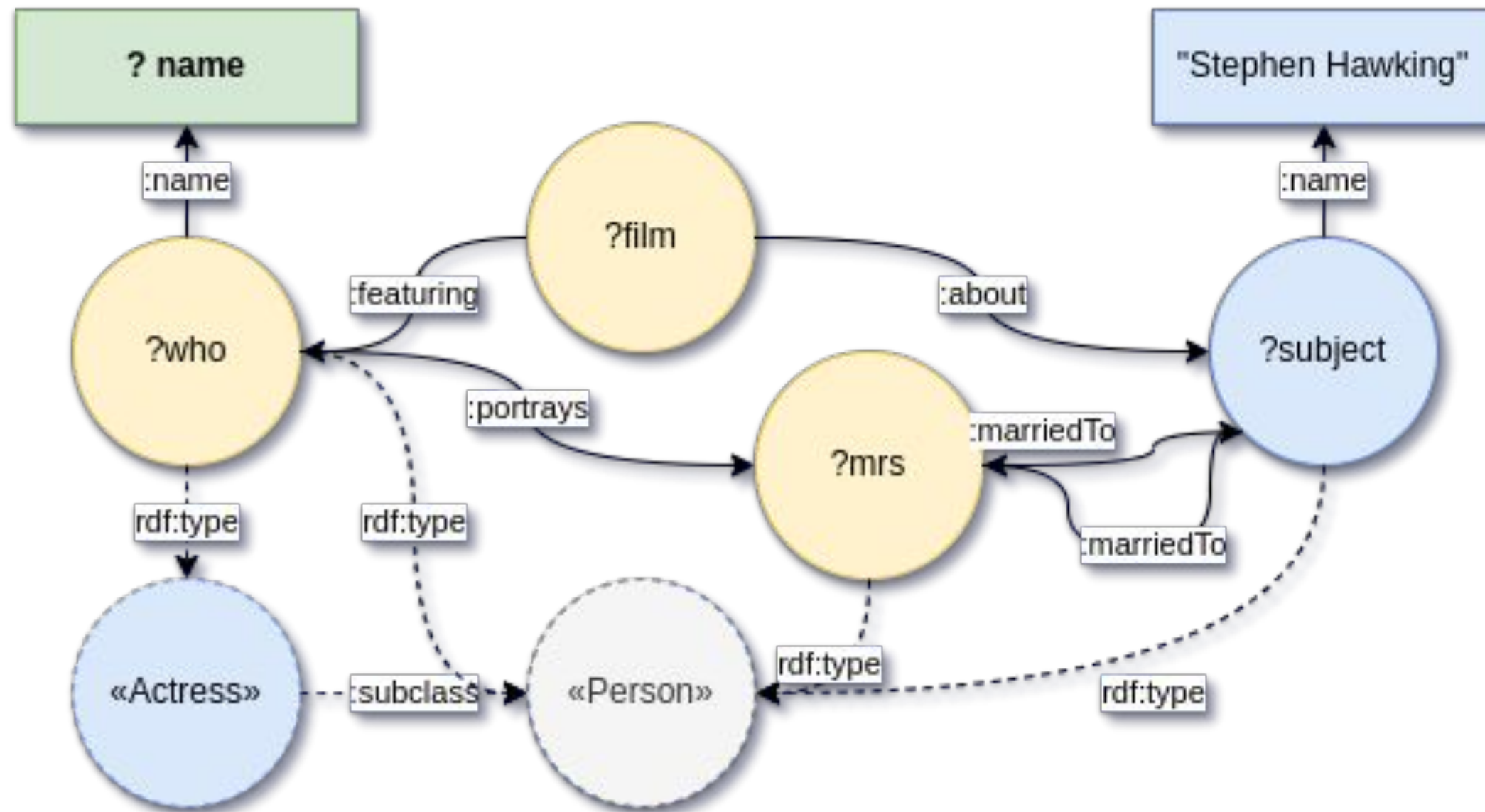
*Like woodwork ...*

*learn to work with the 'grain' of the material  
learn to love and understand the beauty of the web*



who is the actress playing the wife of Stephen Hawking?

Assume we understand the question

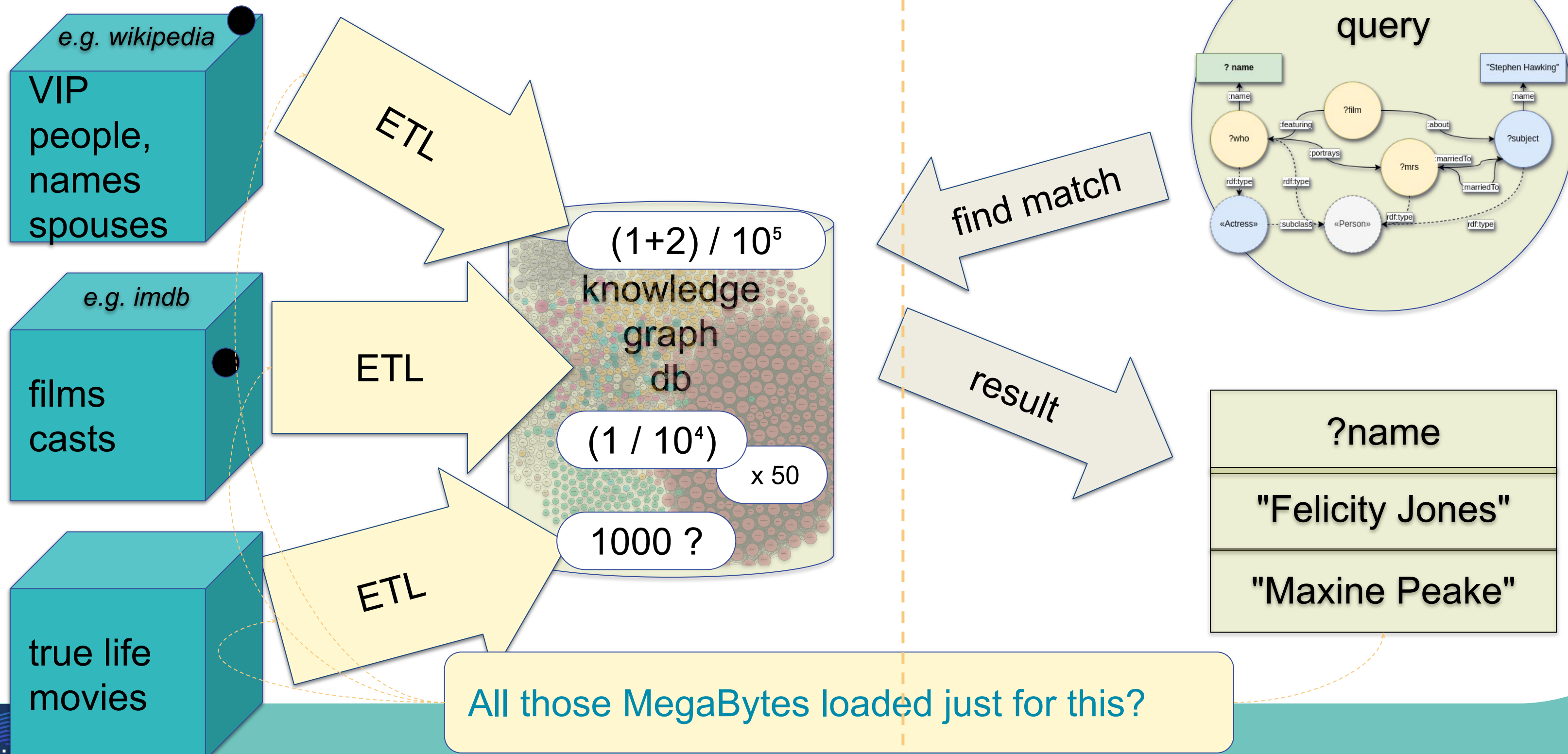


**sparql**

```
SELECT ?name WHERE {  
  ? who :name ?name;  
    :portrays ?mrs.  
  ? film :featuring ?who;  
    :about ?subject.  
  ? subject :married_to ?mrs;  
    :name ?s_name.  
  FILTER( ?s_name="Stephen Hawking")  
}
```

our current approach

# A cannon to kill a mosquito?

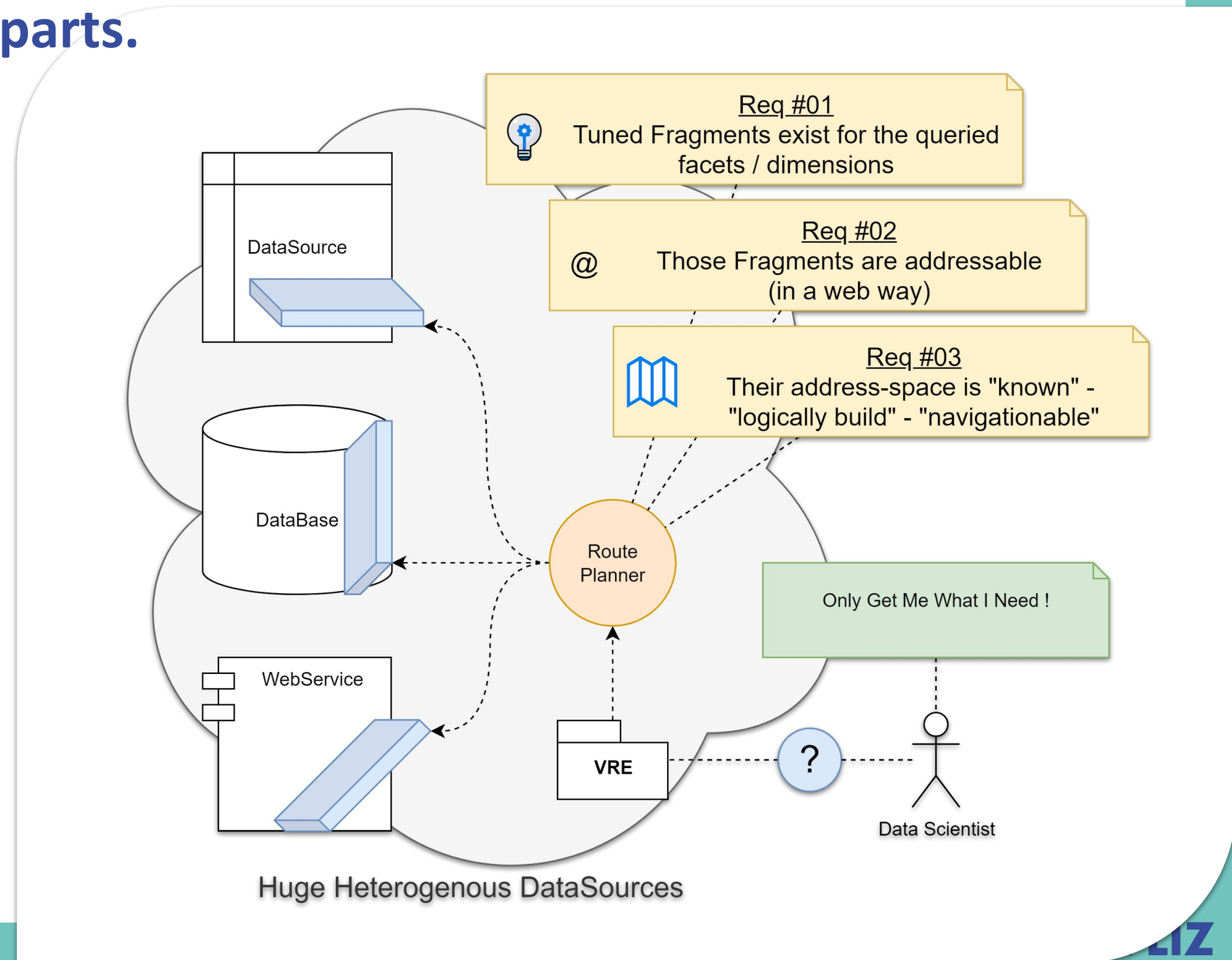




on sliding that data back IN the web

## Finding the shortest path to only the relevant information parts.

- LOD of fragments (exposed deep paths into subsetting services)
- navigation paths between
- routing algorithms
- routing services



serving smart (and powerful) clients

## fragment negotiation

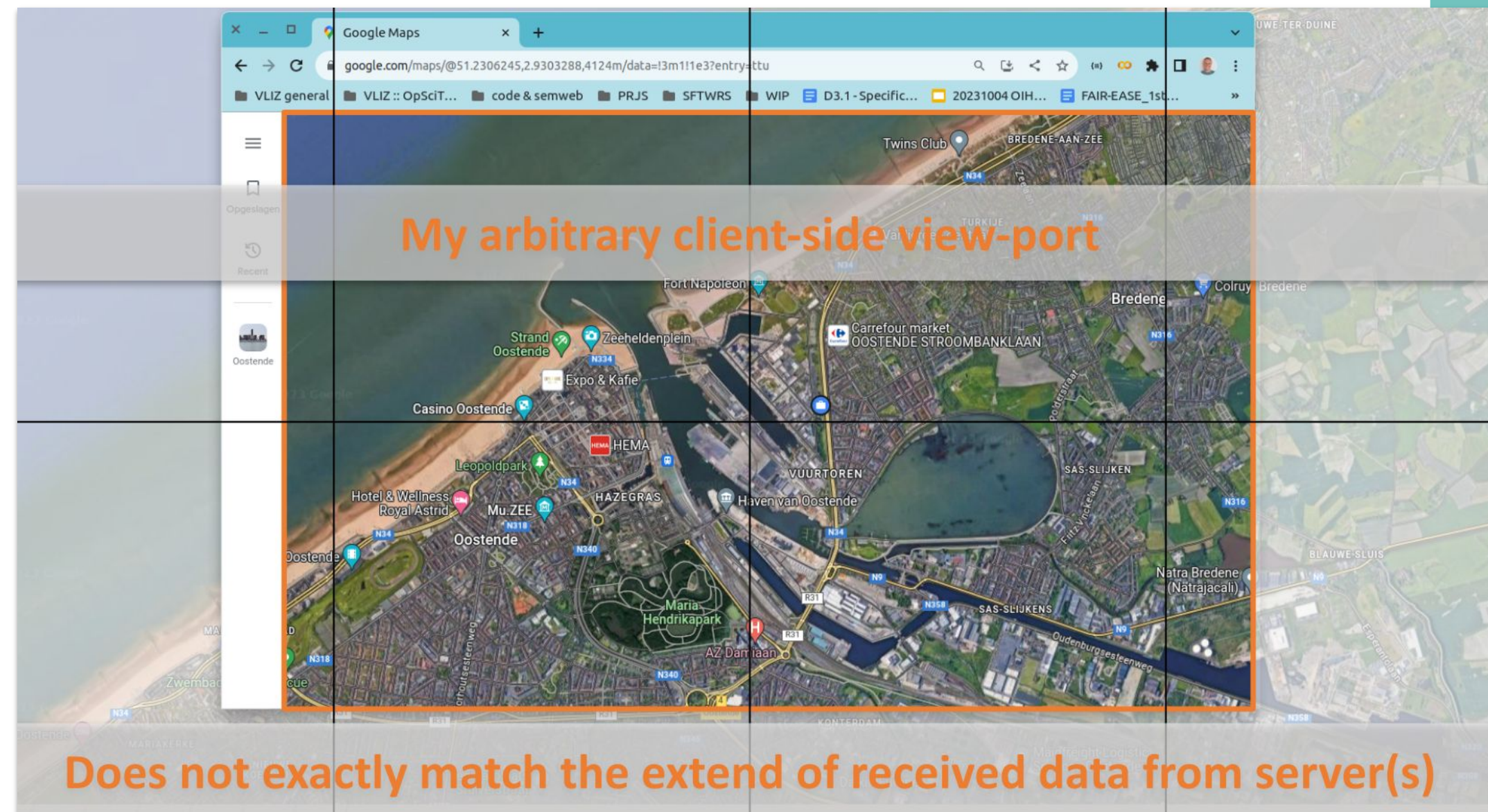
our servers / services  
do **not** need to be

- doing everything
- for everyone

client-side libraries should be  
part of the solution

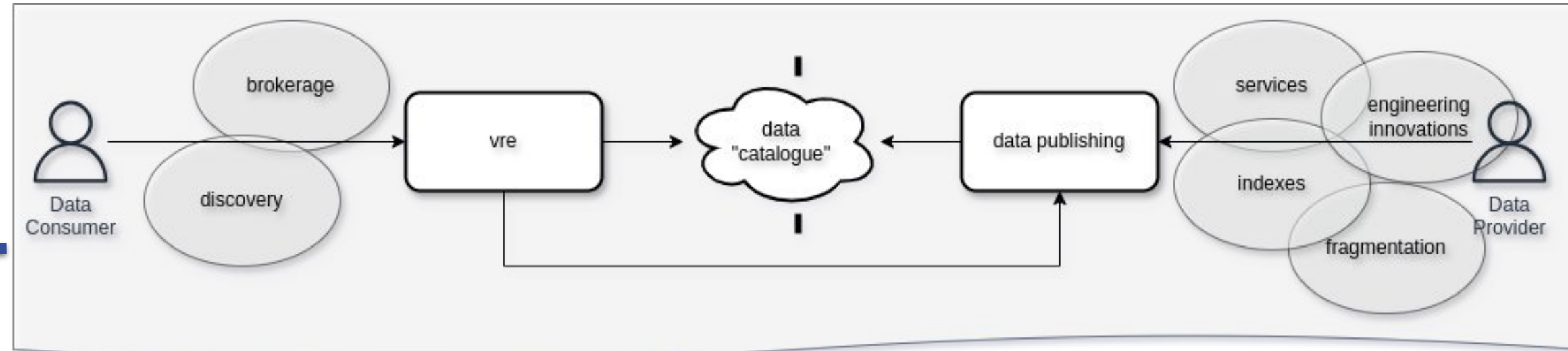
80% fitting fragments can be  
defined and provided (cached)

e.g. Google Maps?

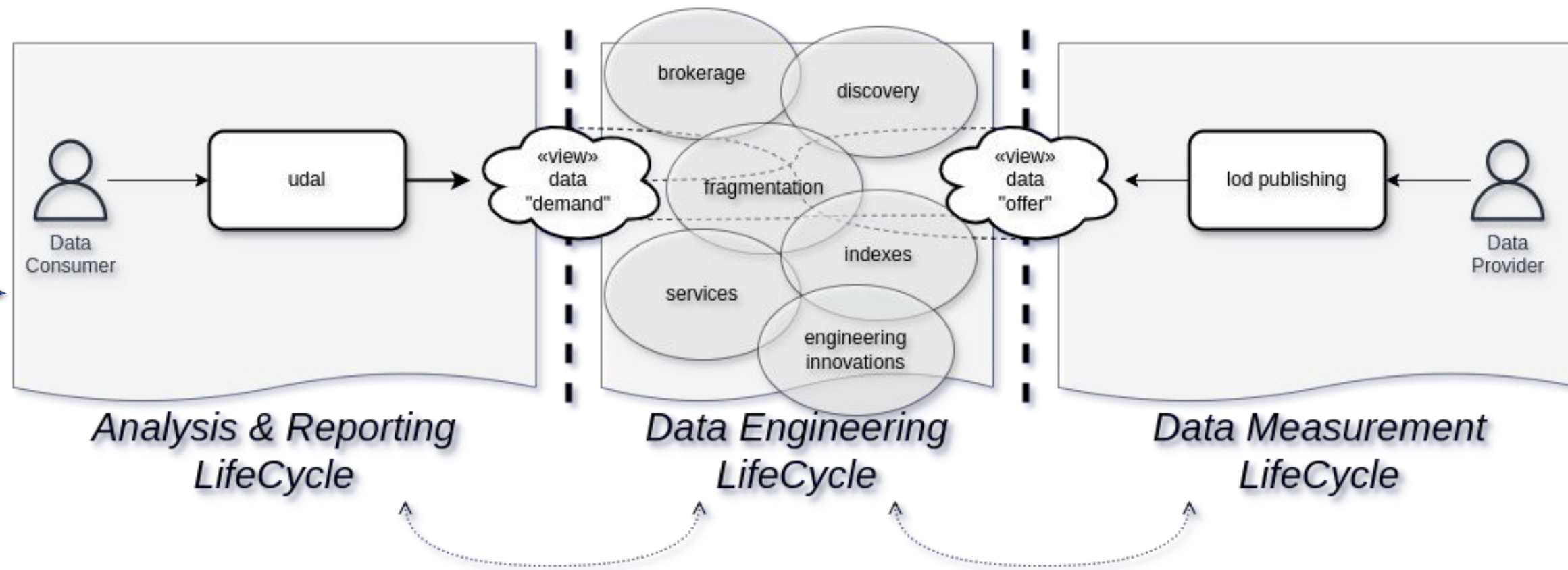


sustainable data flows, through decoupled clients

# UDAL



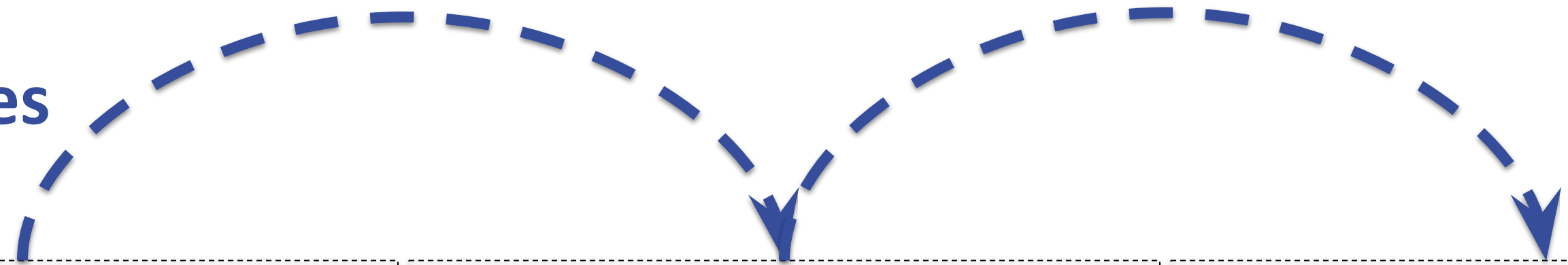
*One LifeCycle, impacted by all influences:  
Analysis, reporting, engineering, measurements, ...*



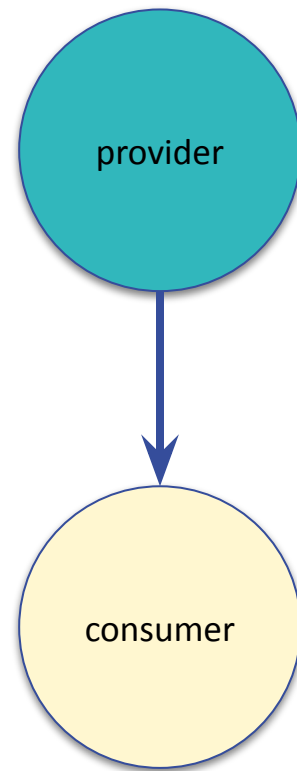


another phase in data-idea line-up

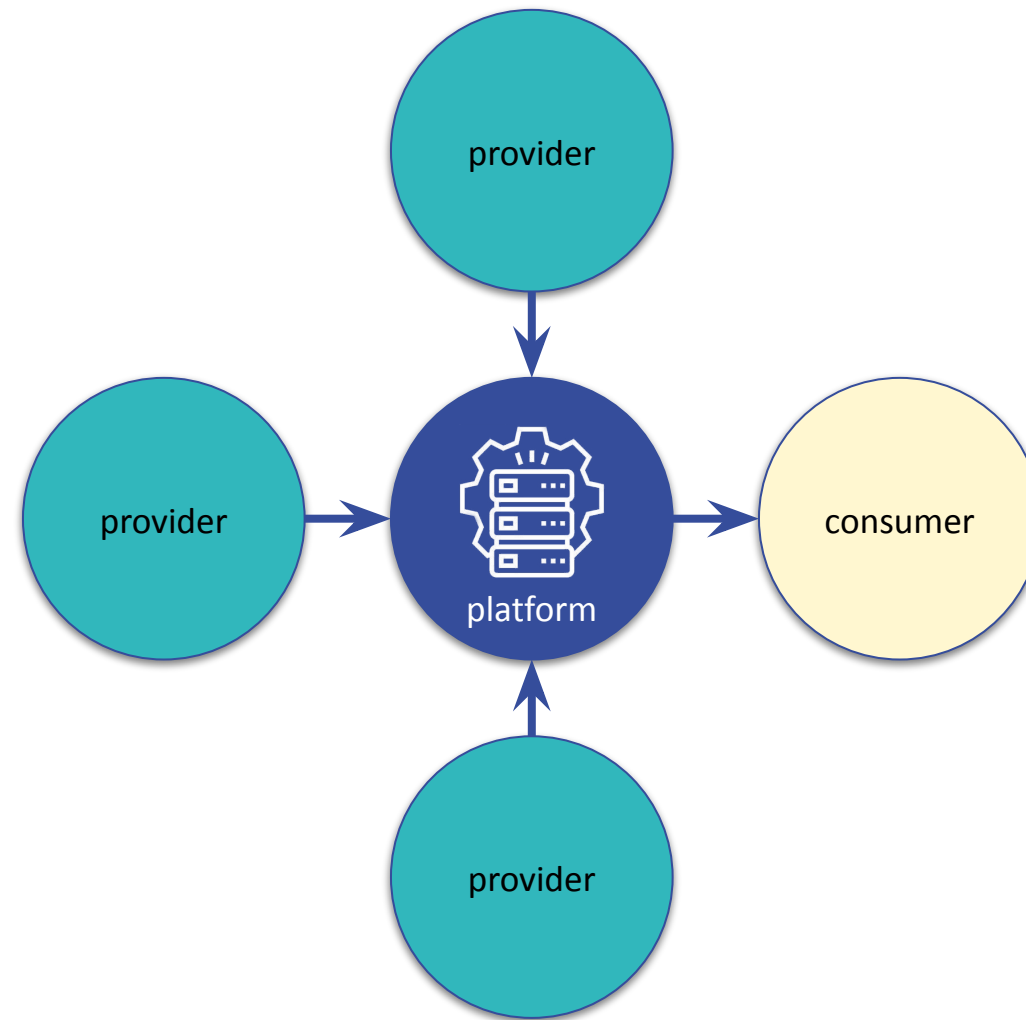
# data-spaces



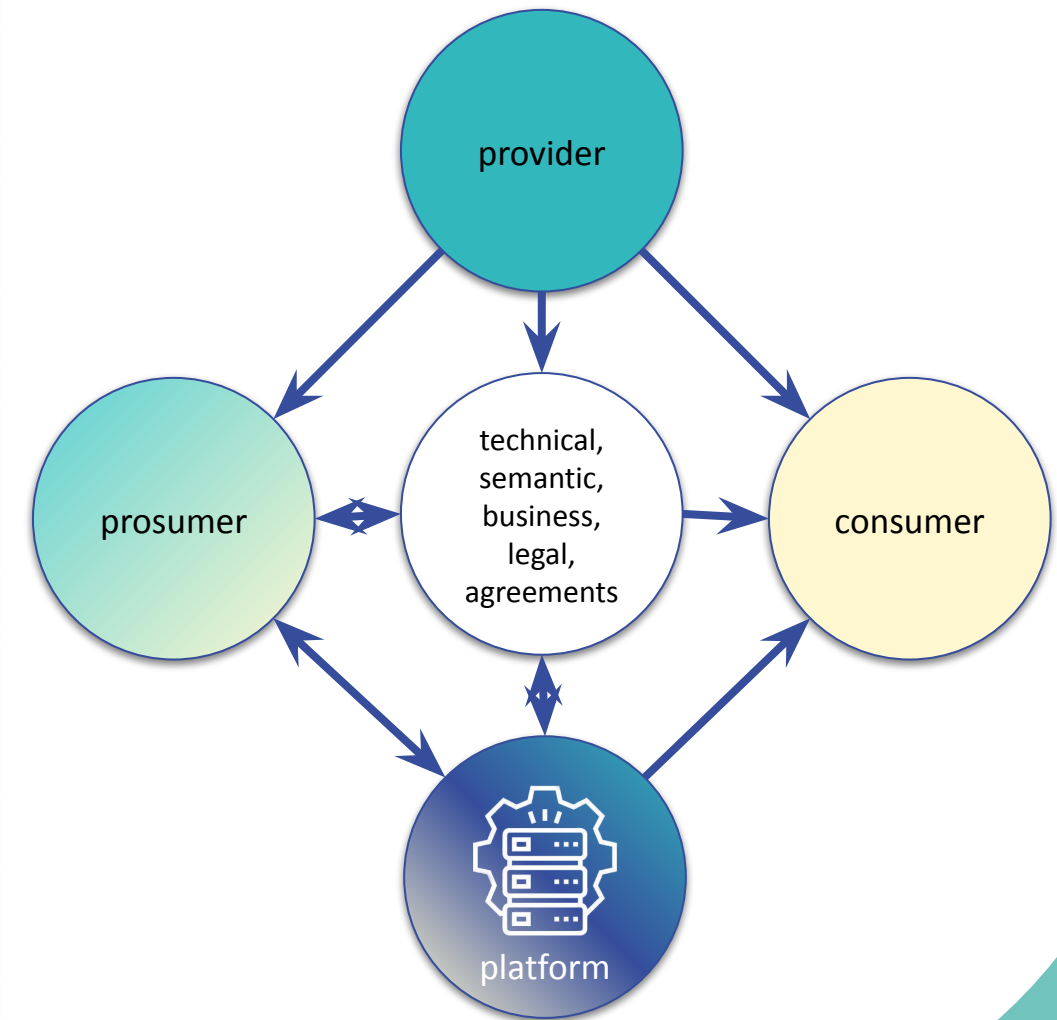
## bilateral / ad hoc



## platform



## dataspace



big changes ahead ...

## **towards an open and flexible data-market-place**

no more end-to-end control over the full data-pipelines  
lowering the cost, ensuring sustainability, instantly plugable data, ...

### **data-as-a-commodity**

sure, we are compelled to go ahead, but (as we experience on the web)

- trust (in a world where publishing cost is low, specially for low quality)
- economics of scale often in favour of the already privileged
- provenance to the rescue?

## recap

beware of the yet-another "ON" the web -syndrome / effort

- yet another webservice / api / protocol
- yet another aggregator / hub / portal
- yet another point-to-point connection

embrace the core web design & put data "in" the web

- describe datasets and data-fragments
- describe where to find them, and how to navigate to narrow down fits
- now think about "routing services"
- and clients that act as site-nav, jit-db, fragment negotiators ...

beware of extra challenges in the "data-as-a-commodity" world

27-29 May 2024 



imdis

# International conference on Marine Data and Information Systems

