



TWC



↔ <http://bit.ly/lebo-issues-2012>

# *Building the Web of Data about the Web of Data*

*Towards Visual Strategies on the Web*



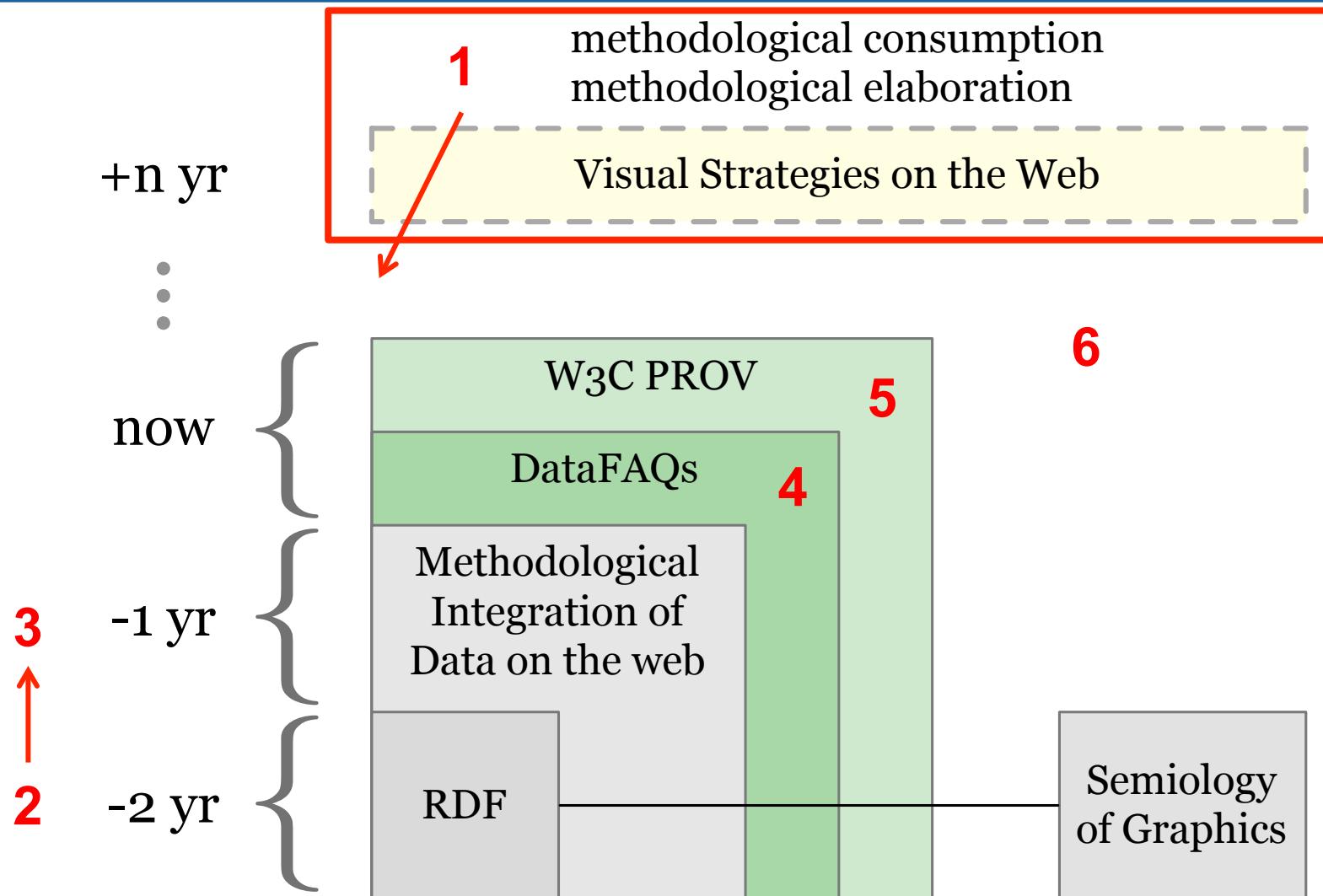
Timothy Lebo  
Tetherless World Constellation  
Rensselaer Polytechnic Institute



Rensselaer



# Two more pieces...





# The Pursuit of Context: an Archaeological Endeavor

*What are the **things** that you're talking about?  
How do those things **relate**?*

RUBES by Leigh Rubin



At last, the mystery of the  
Mayan calendar revealed.



A few years later, or the **same moment**  
somewhere else on the web.



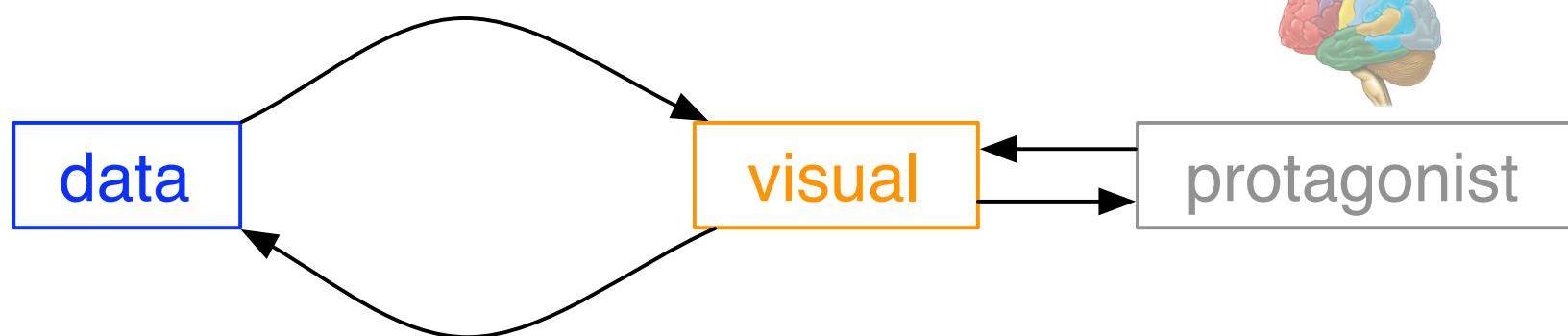
# Objective

Protagonist is more informed about the world

Protagonist cannot consume **data** directly

Data must be transformed into **visuals** for consumption

**Vision** is the predominant method of consumption

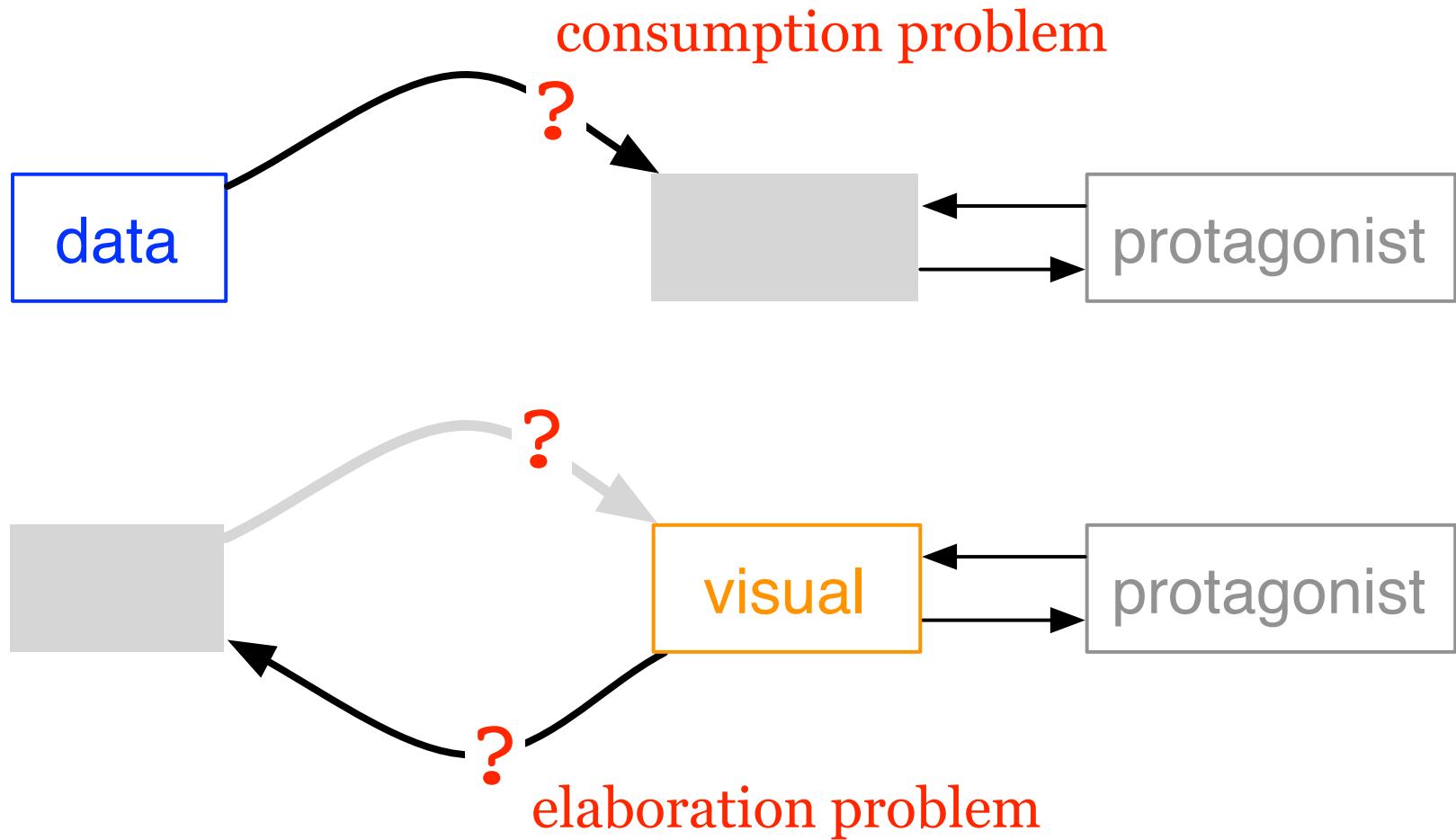


Inquiries of **visual** content require elaboration of the source **data**



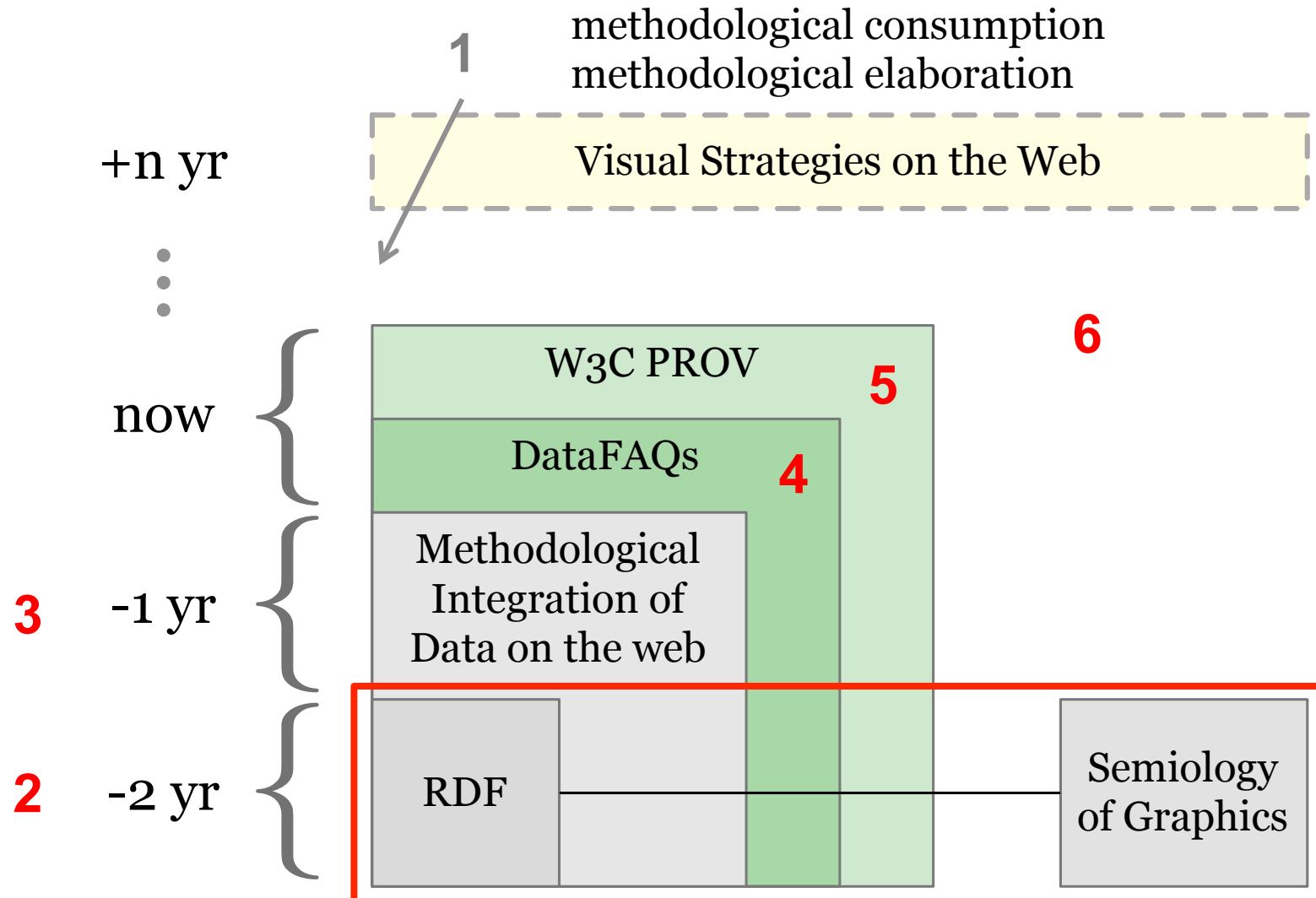


# Two Symmetric Challenges





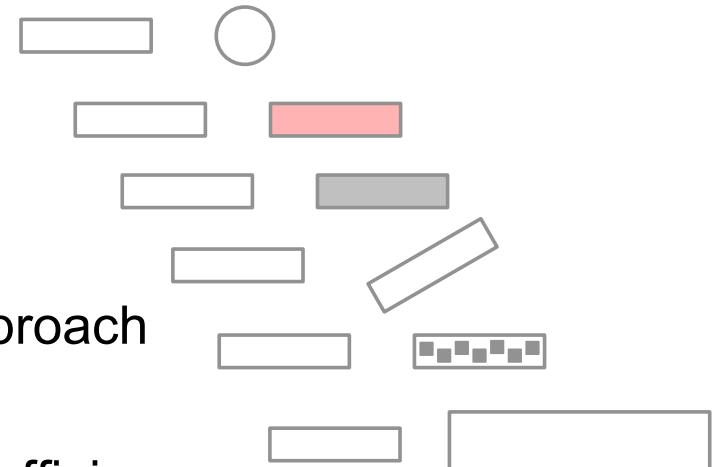
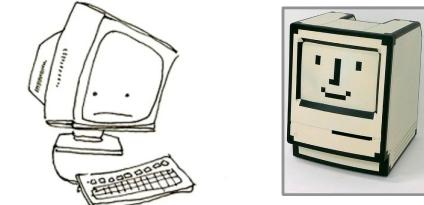
# Foundations





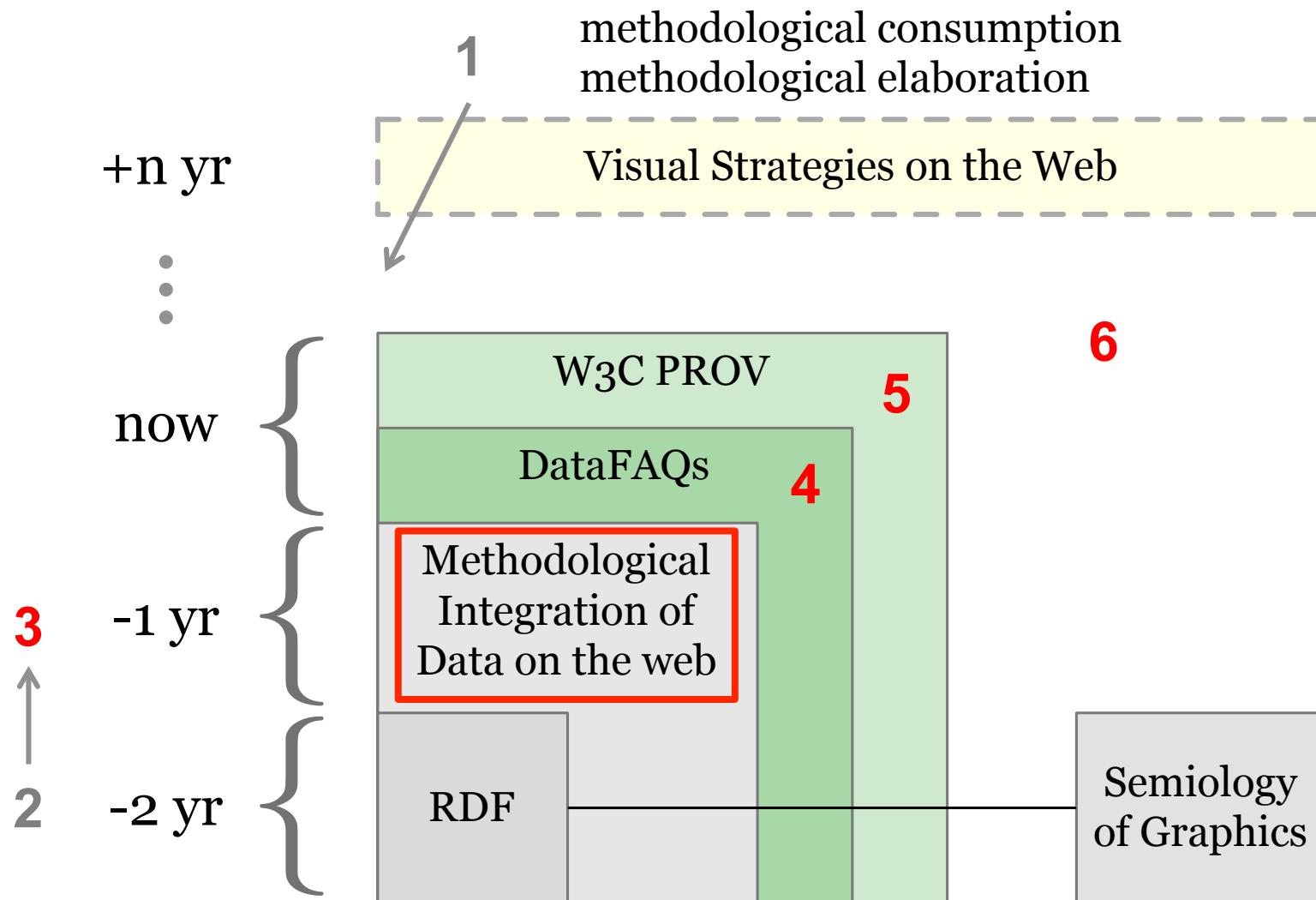
# Basis for Visual Strategies on the Web

- RDF: My Data Hero
  - What are the **things** you're talking about?
  - How do those things **relate**?
  - Can I **get more information** about a thing with JUST its name?
  - Can I **merge** your data with **any other data** in the same format?
- Bertin's Semiology of Graphics
  - Set of Information
  - Analysis of the Information
  - Domain vs. Graphical sign-systems
  - Imposition, implantation, perceptual approach
  - Stages in reading a graphic
  - Level of reading, instant of perception, efficiency





# Methodological Data Integration





# Good RDF and Bad RDF

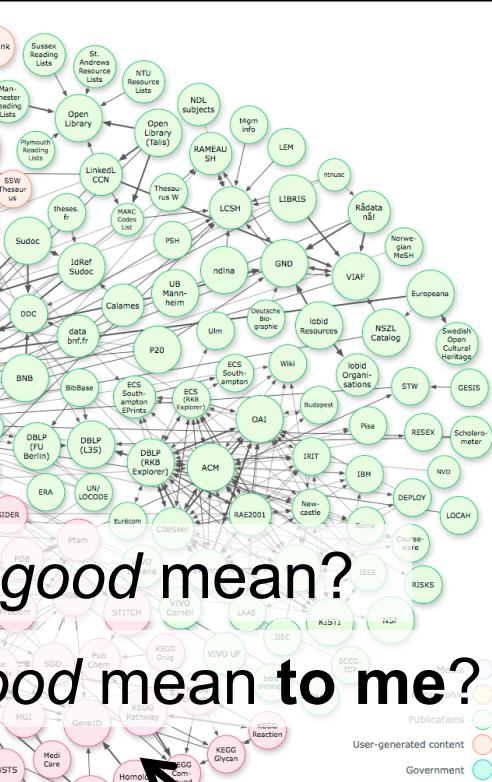
"Hawaii",  
"75-6129  
"96740",

What about the rest of it?

What does good mean?

What does good mean to me?

```
ds4383:th
raw:co
raw:co
raw:co
raw:column_3 "75-6129 Alii Drive",
raw:column_4 "Kailua-Kona";
raw:column_5 "96740";
raw:column_6 "-155.9819183";
raw:column_7 "19.61436844" .
```



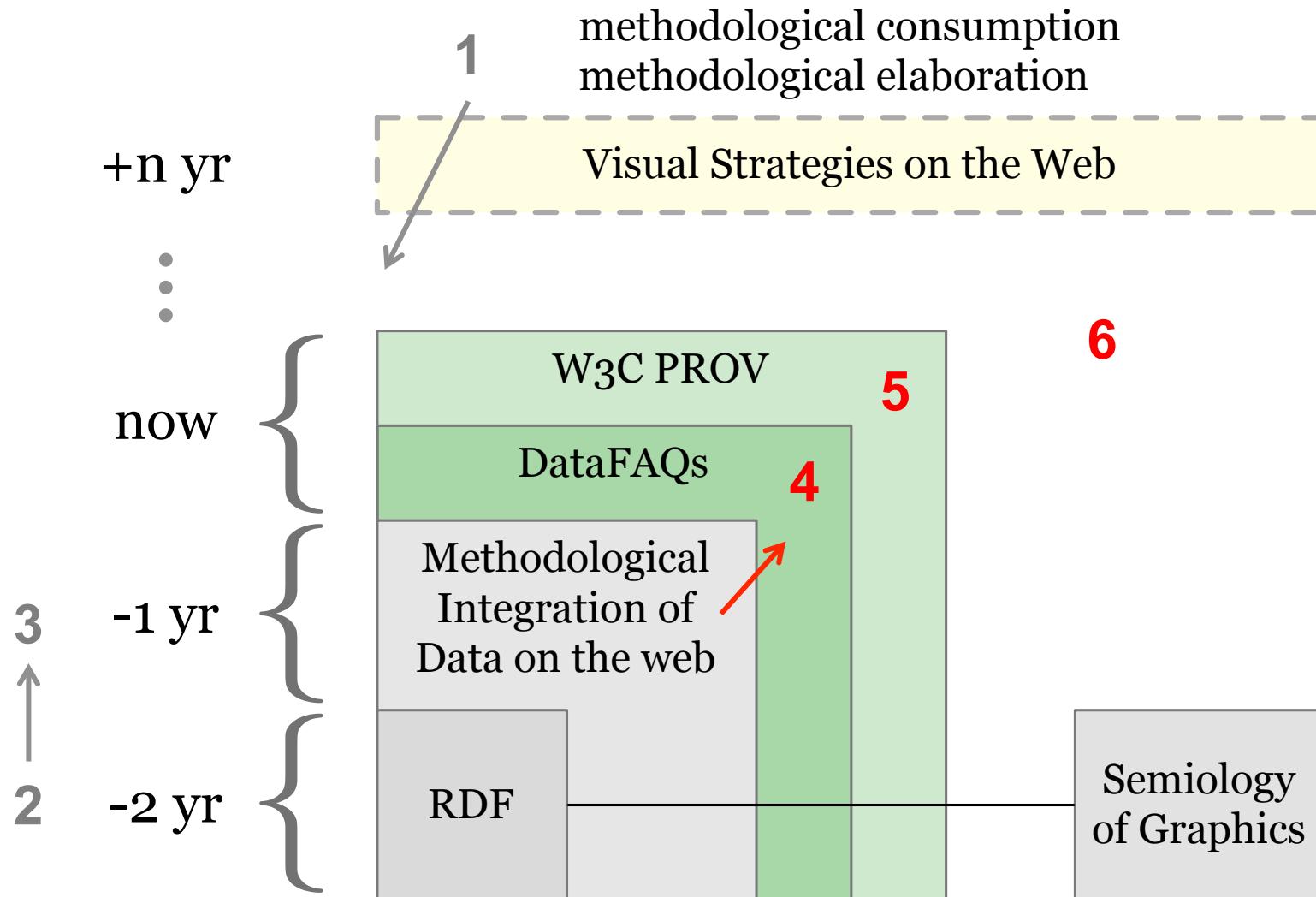
owl:sameAs  
<http://sws.geonames.org/5855797/>,  
govtrackusgov:HI,  
dbpedia:Hawaii .

arket;  
1367;  
den Market Place";  
g 19.6 .

ed\_state:Hawaii;  
6129 Alii Drive";  
Lua-Kona";  
40" .



# Evaluating Web Data





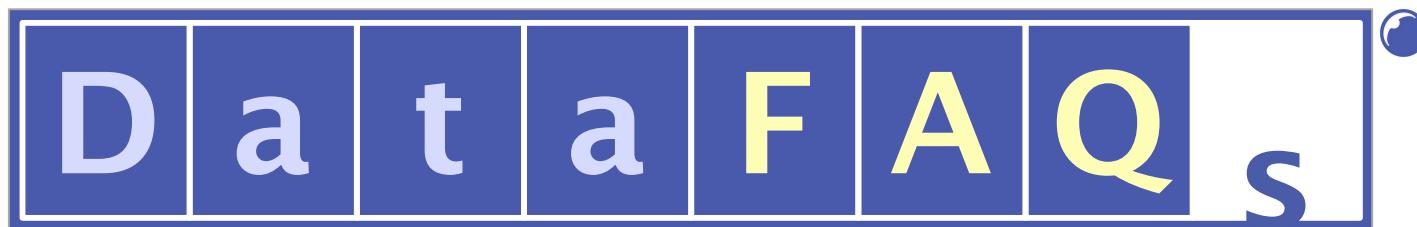
# Motivations for DataFAQs

- Data level
  - Which parts of Linked Open Data are good, bad?
  - How can they get better?
- Community level
  - What does good and bad data *mean*?
  - Explicit, observable representations
- For Visual strategies
  - Easier to reason with *good* data
  - Need to select *which* data to show before we can determine *how* to present it
  - Visualization quality parallels data quality

D | a | t | a | F | A | Q | s

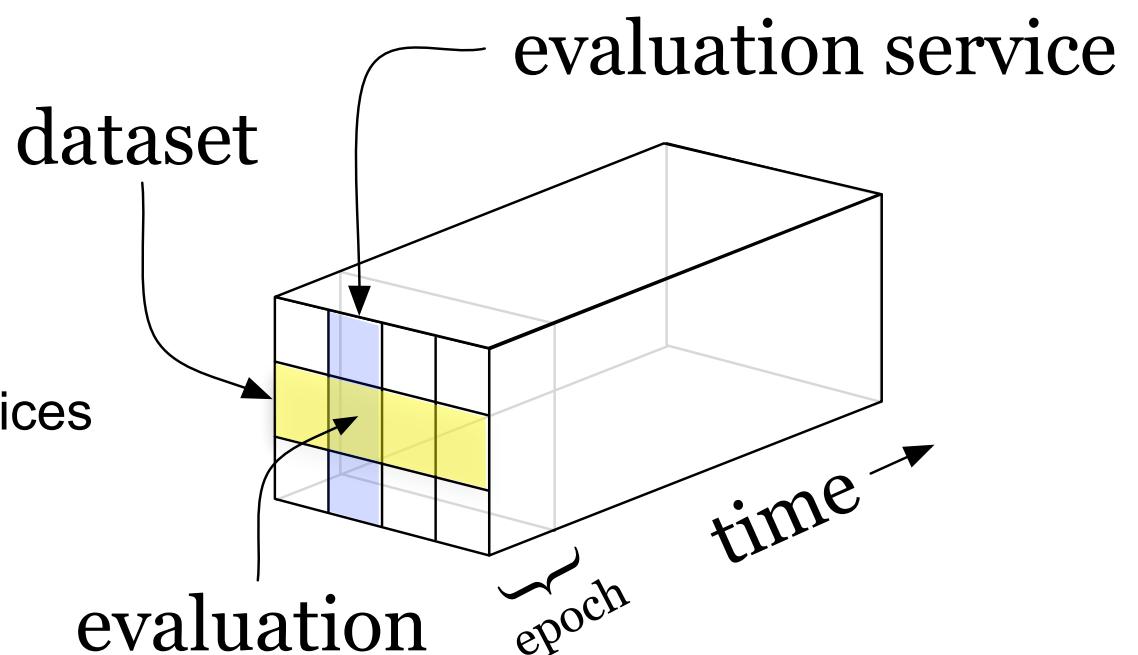


# DataFAQs essentials



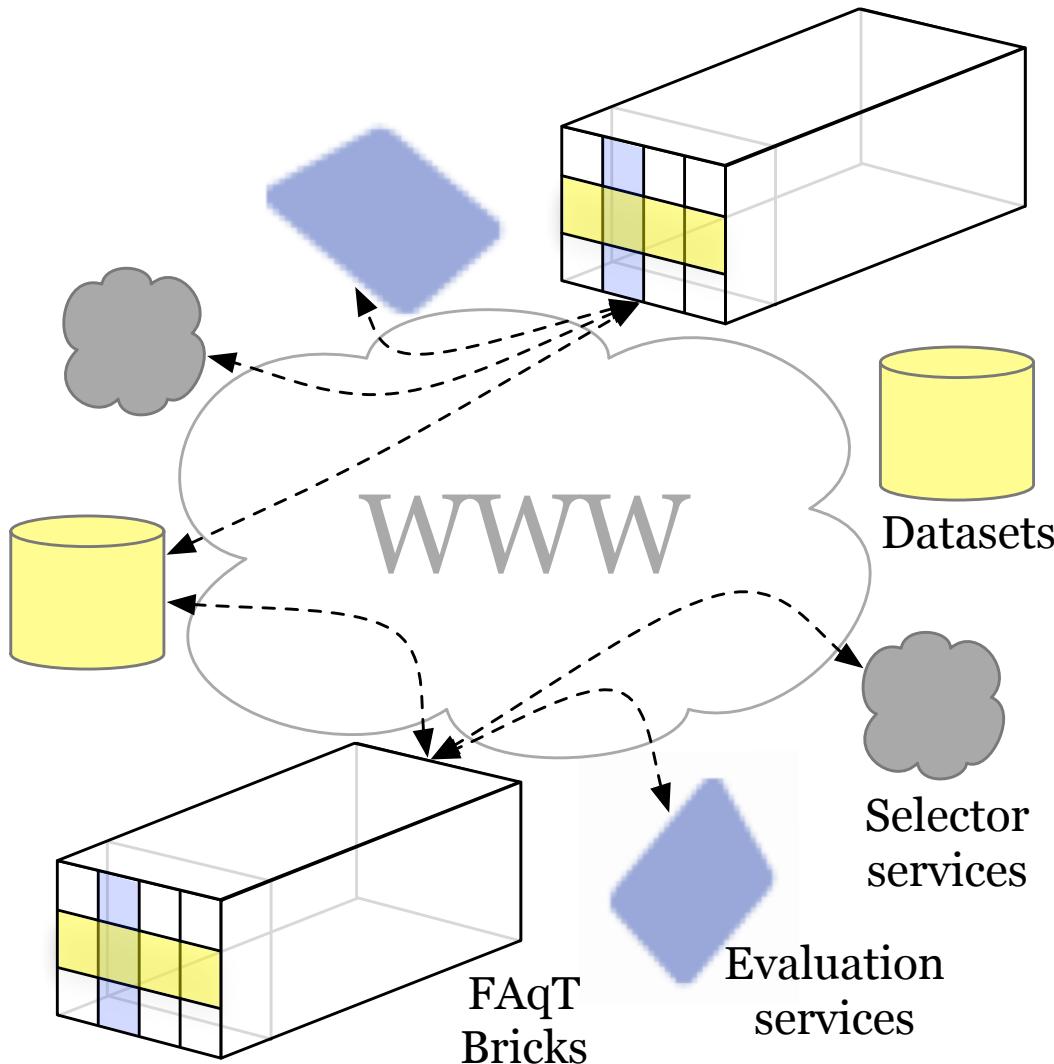
## LINKED DATA QUALITY REPORTS

1. Choose Datasets
2. Choose Evaluation Services
3. Explore Evaluations





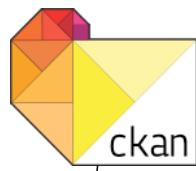
# DataFAQs is distributed



- You create and maintain your own **FAqT Brick**
- You evaluate others' **datasets**
- You uses others' **evaluation services**
- You and others **access** your **FAqT Brick**



# Collecting Webs of Data about the Web of Data



datasets

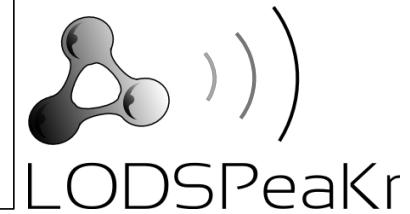
$d_1$

$d_2$

$d_3$

evaluation services

	$f_1$	$f_2$	$f_3$	$f_4$
$d_1$				
$d_2$		evaluations $e_{2,2}$		
$d_3$				





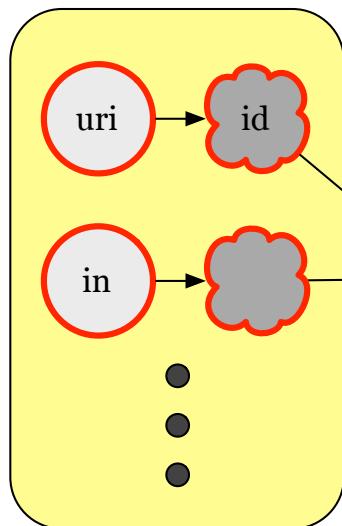
# Selecting Webs of Data about the Web of Data

RDF  
Descriptions

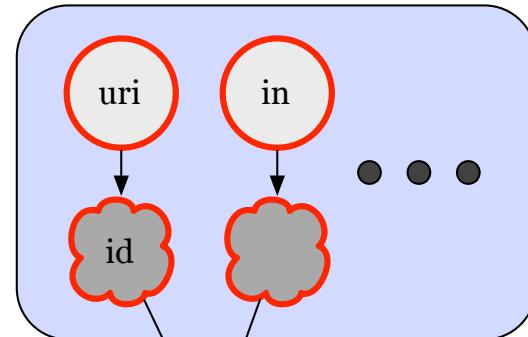
Selector  
services

Select  
Datasets to  
Evaluate

Add  
Dataset  
References



Select  
Evaluation Services



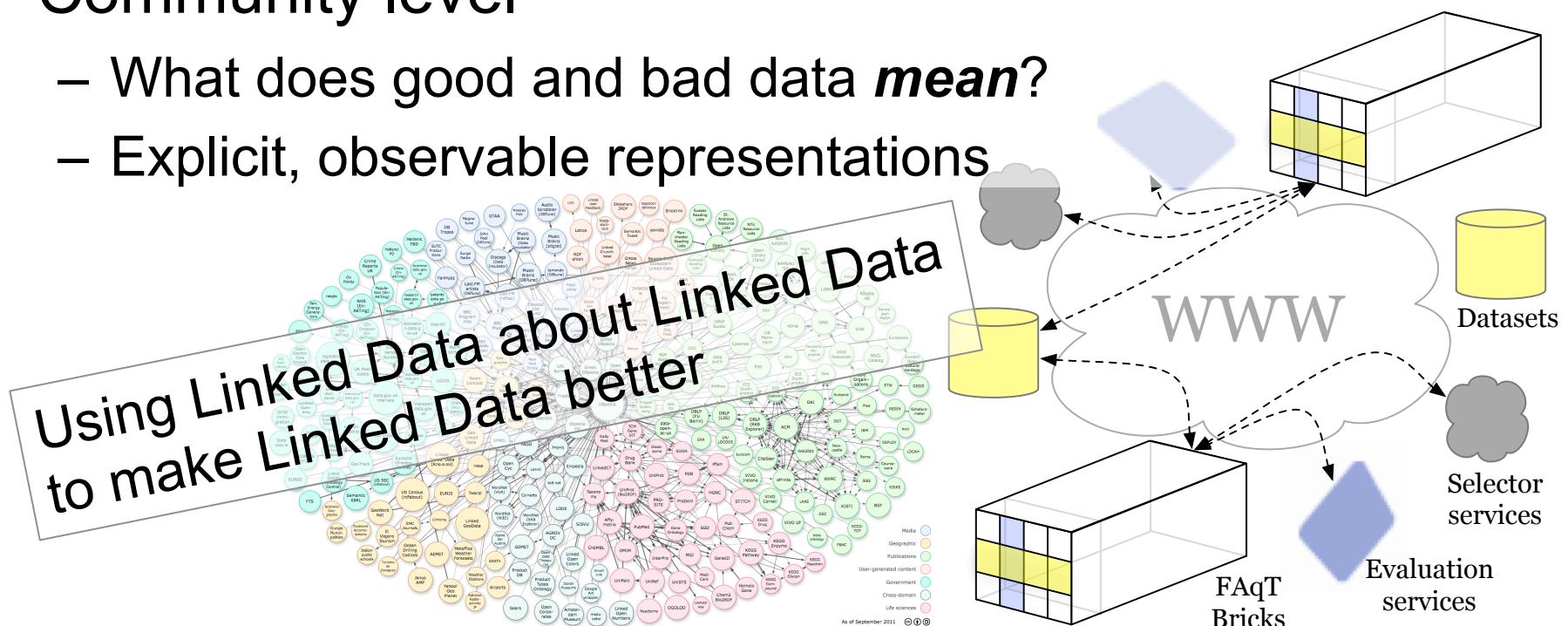
Evaluation Services

		Eval: C+	



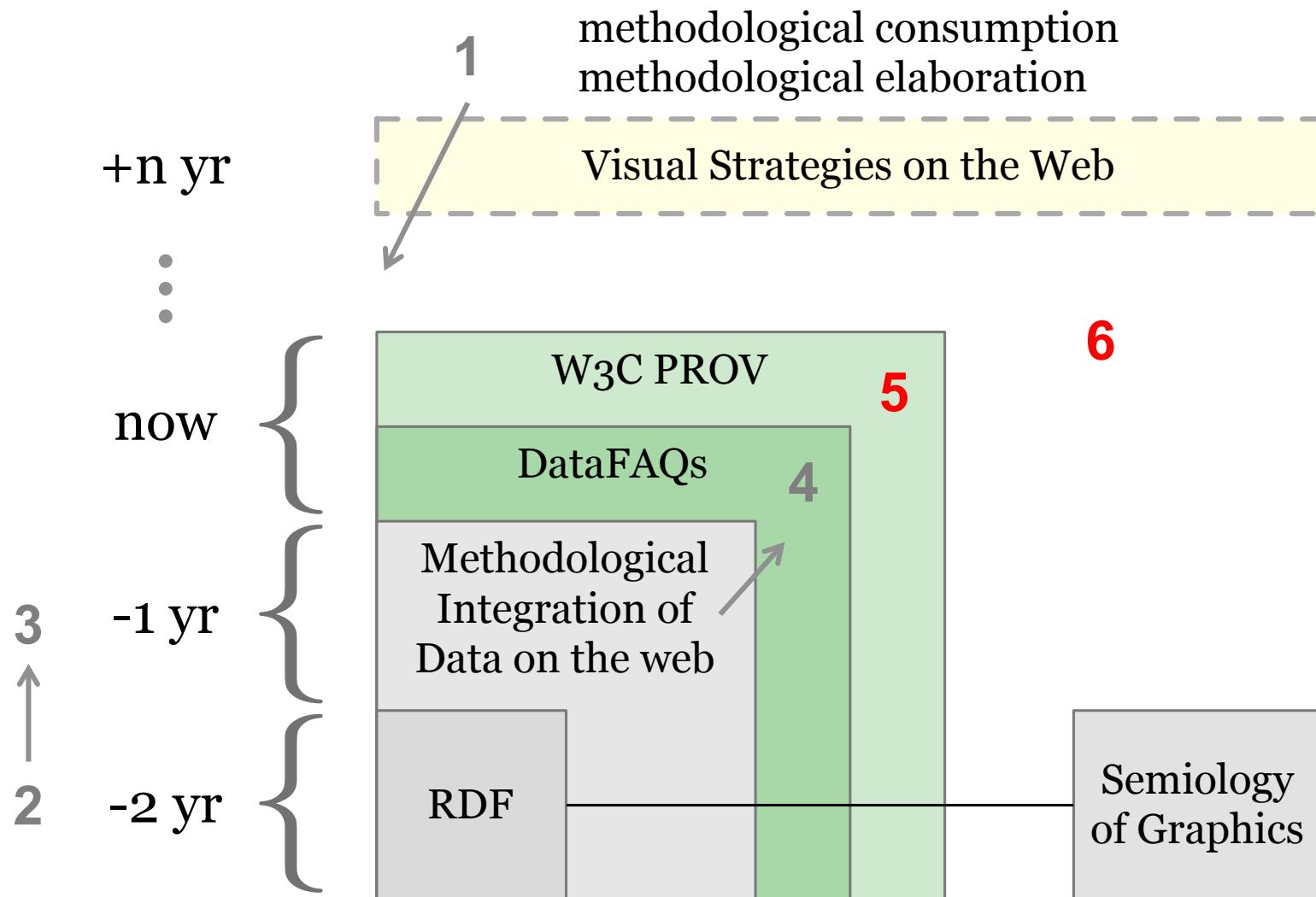
# Applying DataFAQs

- Data level
  - Which parts of LOD are good, bad?
  - How can they get better?
- Community level
  - What does good and bad data *mean*?
  - Explicit, observable representations



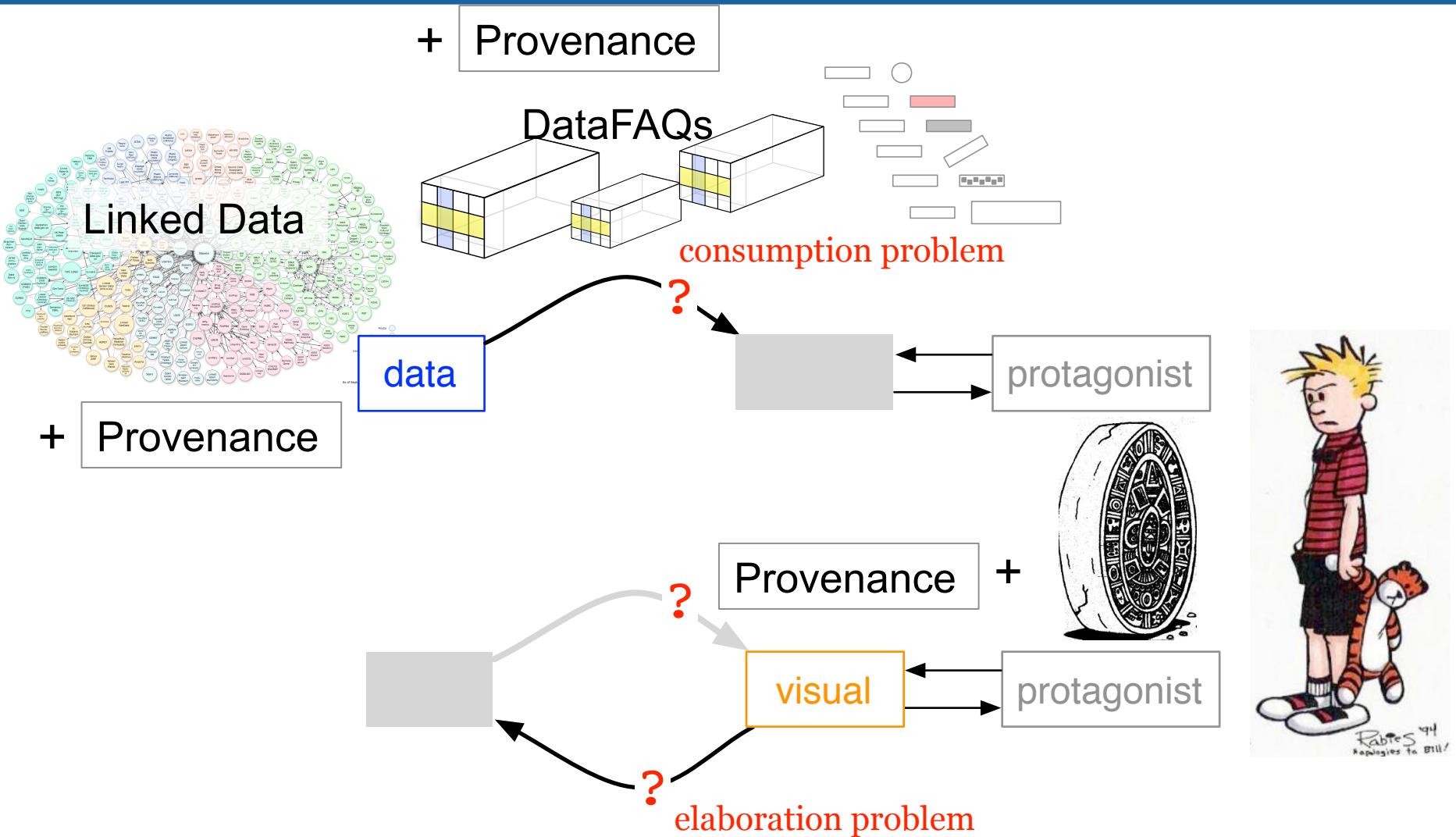


# Contextualizing Web Data





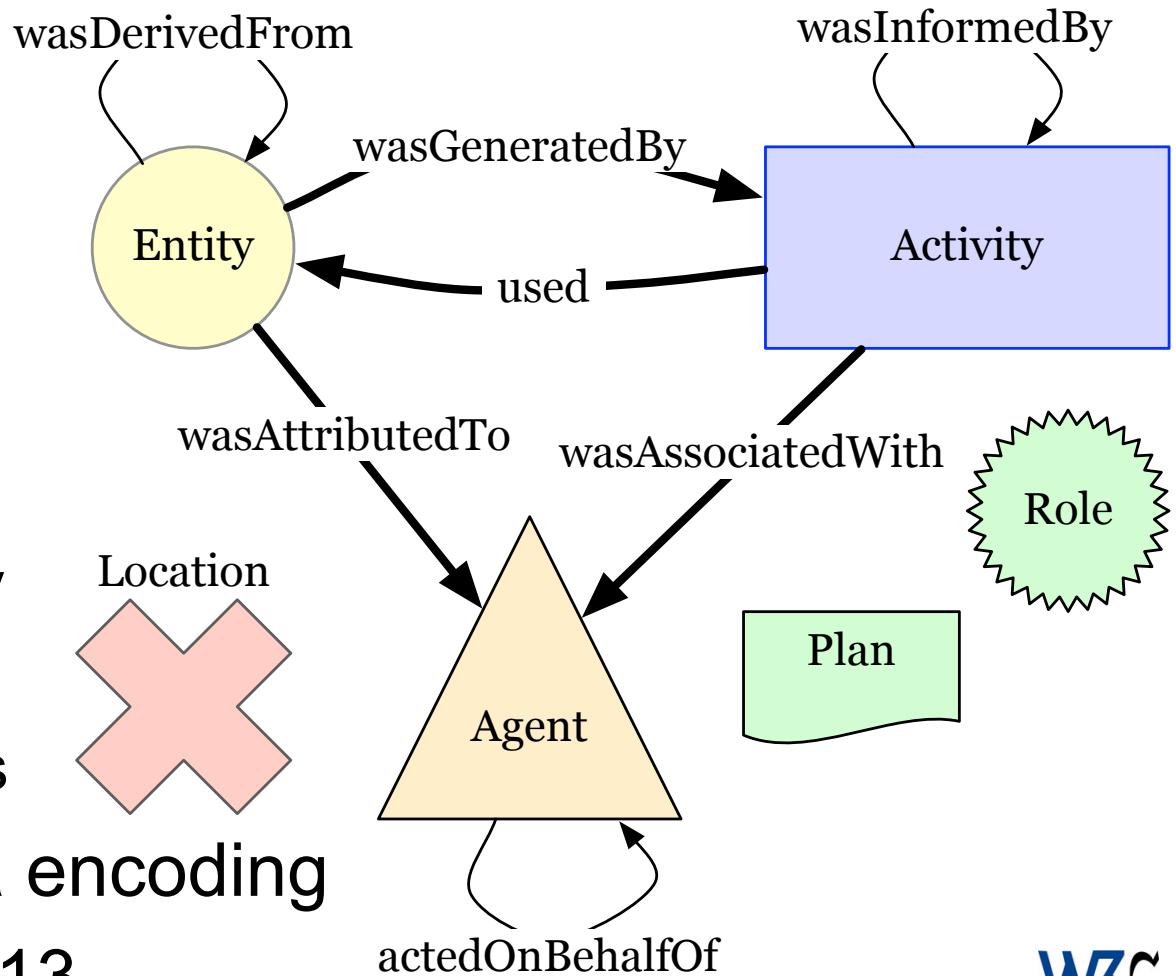
# Motivations for Provenance





# W3C Provenance Data Model (In Progress)

- Chains of
  - Entities
  - Activities
  - Both
- Responsibility
  - For Entity / Activity
  - Via delegation
  - According to Plans
- RDF / Linked Data encoding
- Final Rec ~Jan 2013

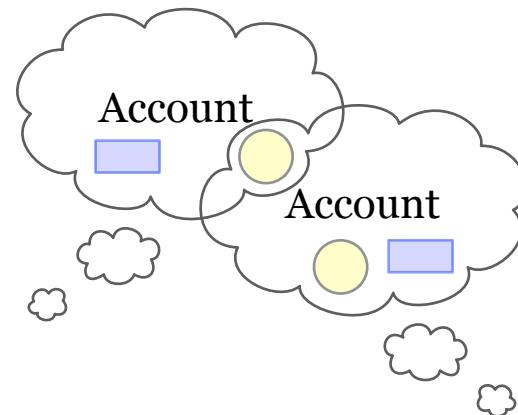


**W3C**

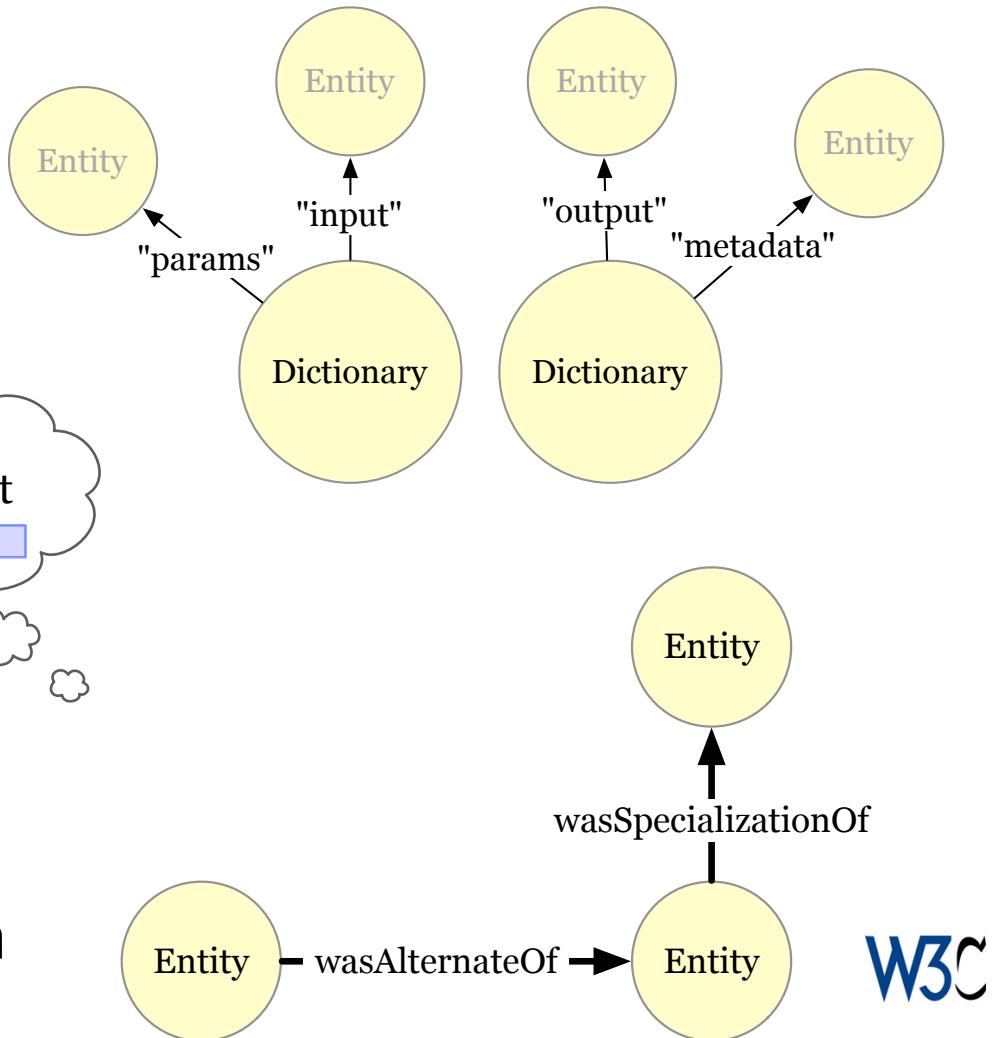


# W3C Provenance Data Model (In Progress)

- Composite structures
  - Dictionaries
  - Accounts

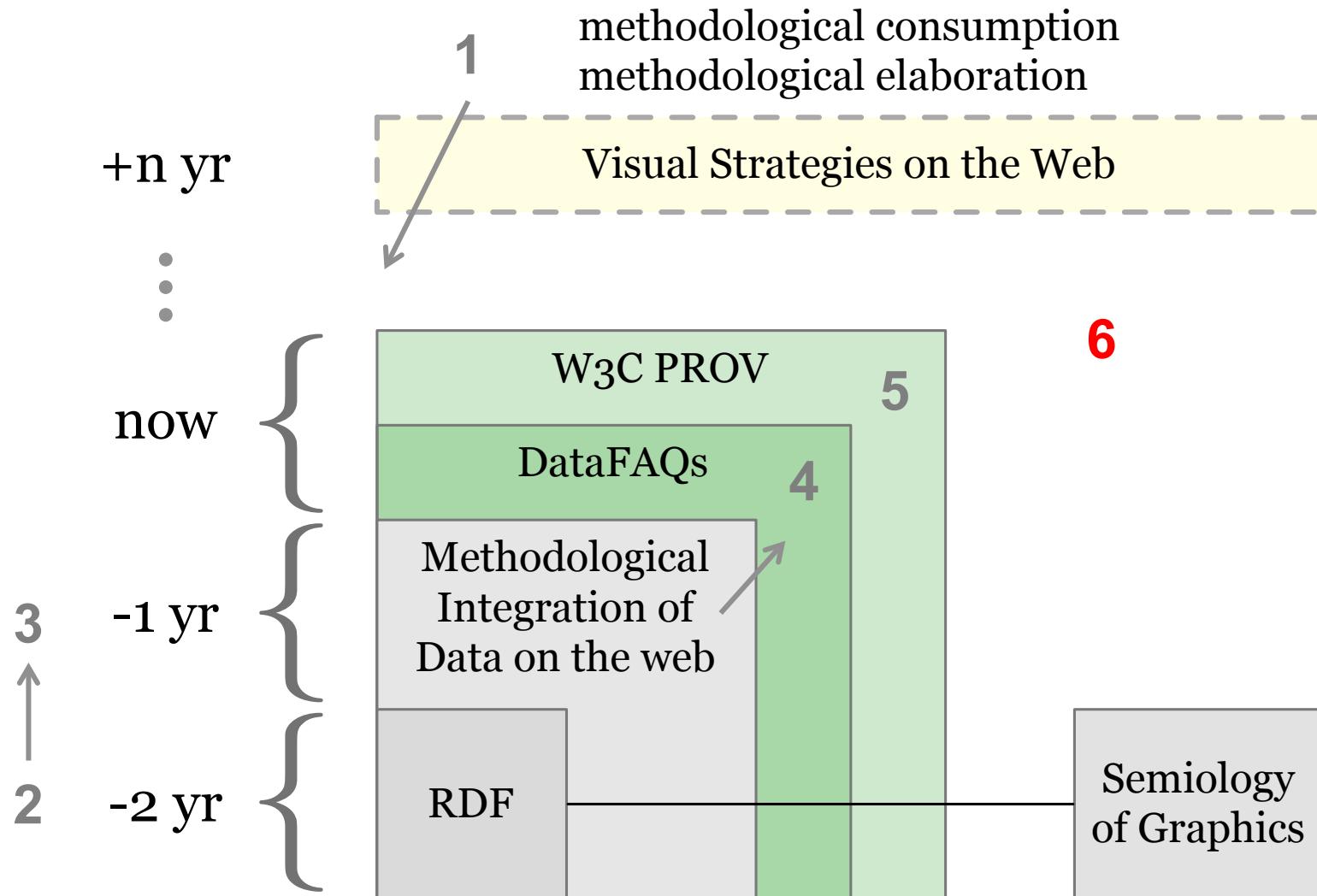


- Distinguishing levels of abstraction
  - Alternate / Specialization





# Summary

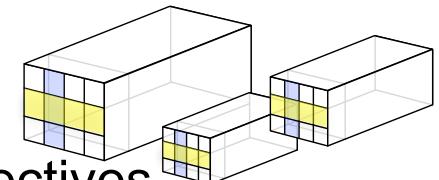




# Conclusion: The Fight for Context

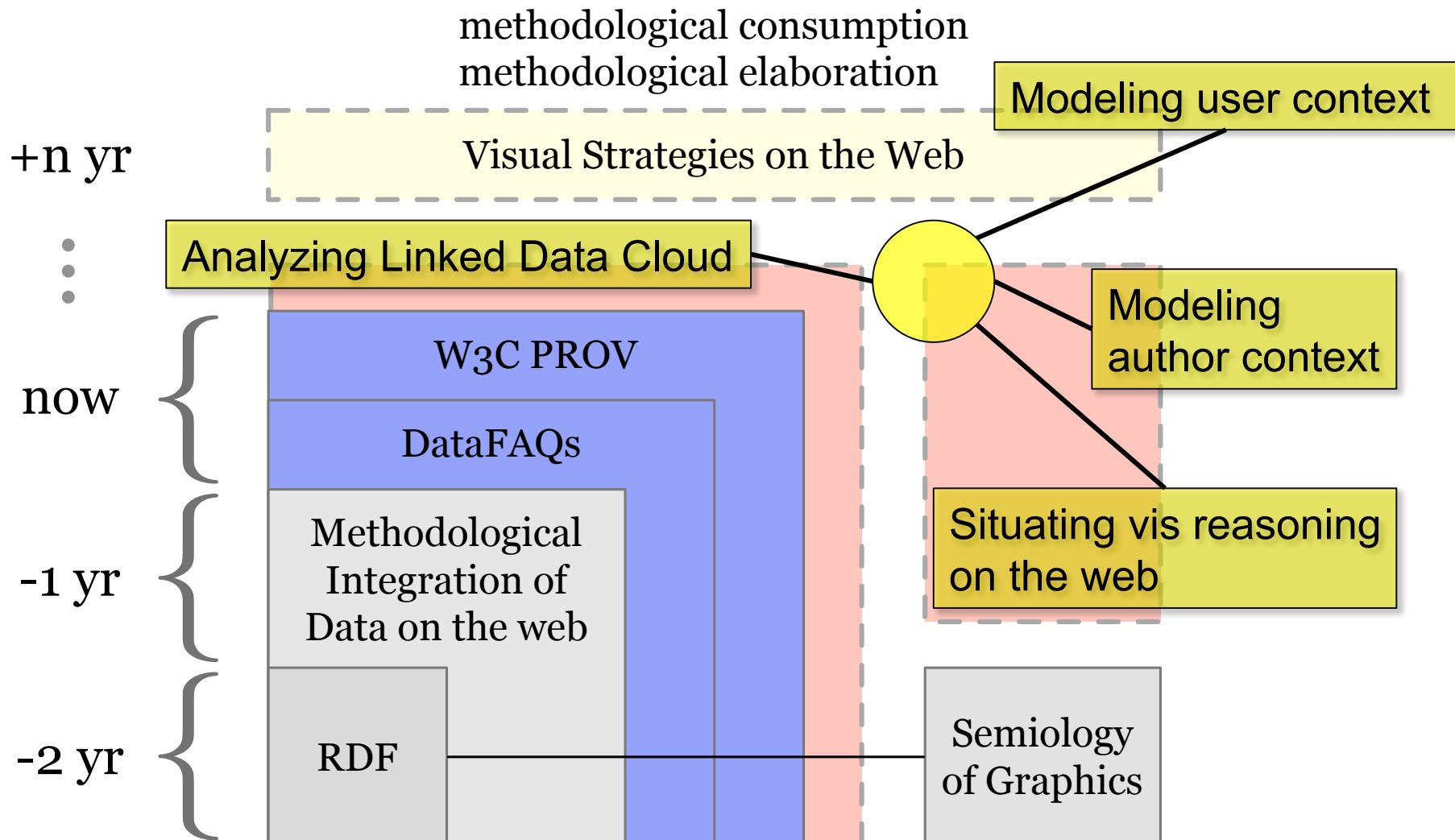
... and putting it on the web!

- Motivation with Calvin 
- Integrating data on the Web
  - Contextualizing entity names, relations  $G(\text{Title II}, \text{key}) = \text{50349700}$
  - Connecting isolated tokens across datasets
- DataFAQs quality evaluation framework
  - Evaluation services encapsulate contextualized perspectives
  - Analyses contextualize (dataset, evaluation service) pairings
- W3C Provenance Data Model
  - Entities, Activities, Agents, Accounts, Collections
  - Related to an historical context



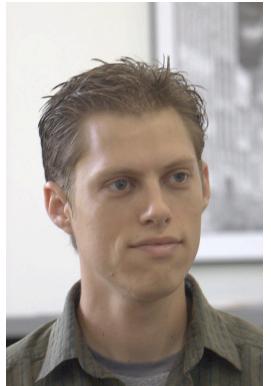


# Future Work





# Thanks!



- Yu Chen
- Dominic DiFranzo
- Jim McCusker
- Josh Shinavier
- Alvaro Graves



Deborah McGuinness

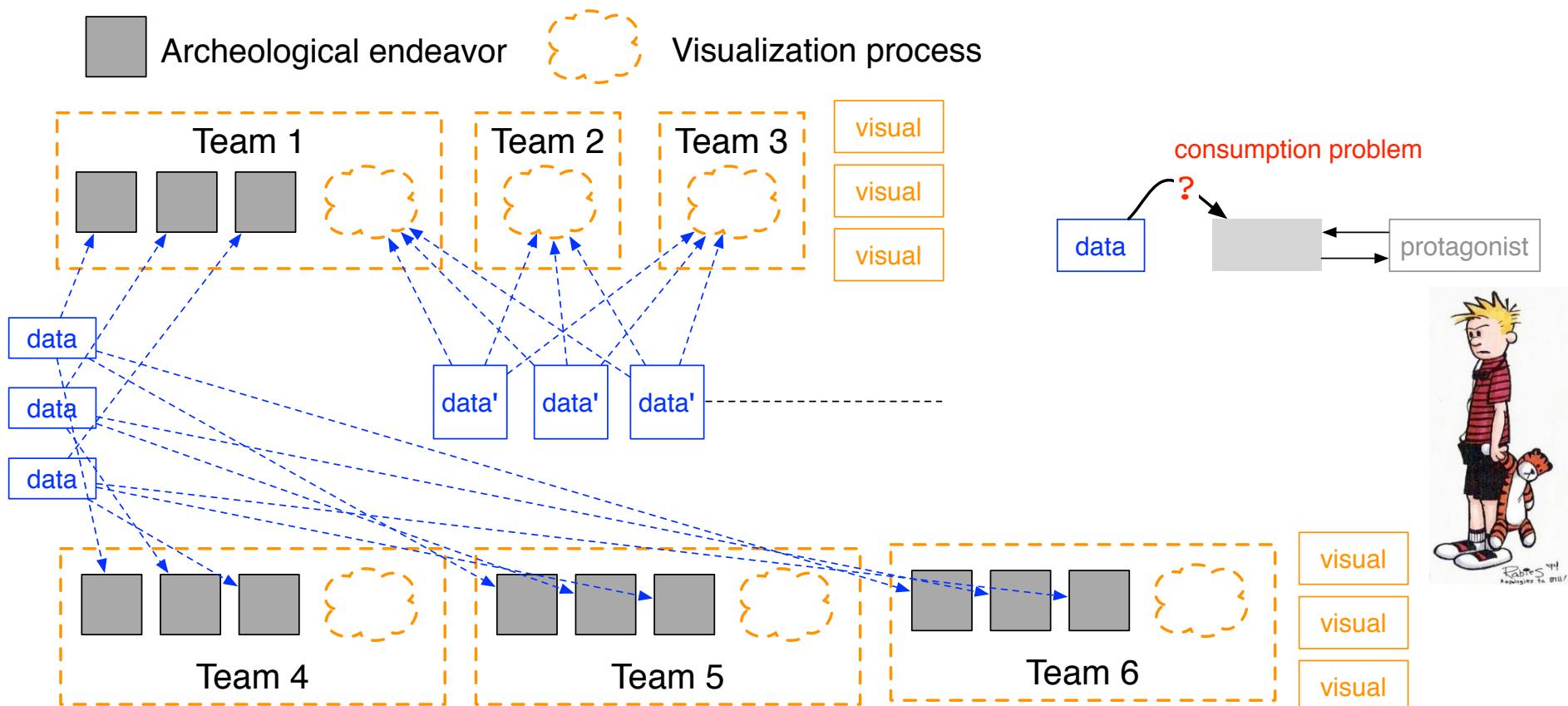


TWC

- backups

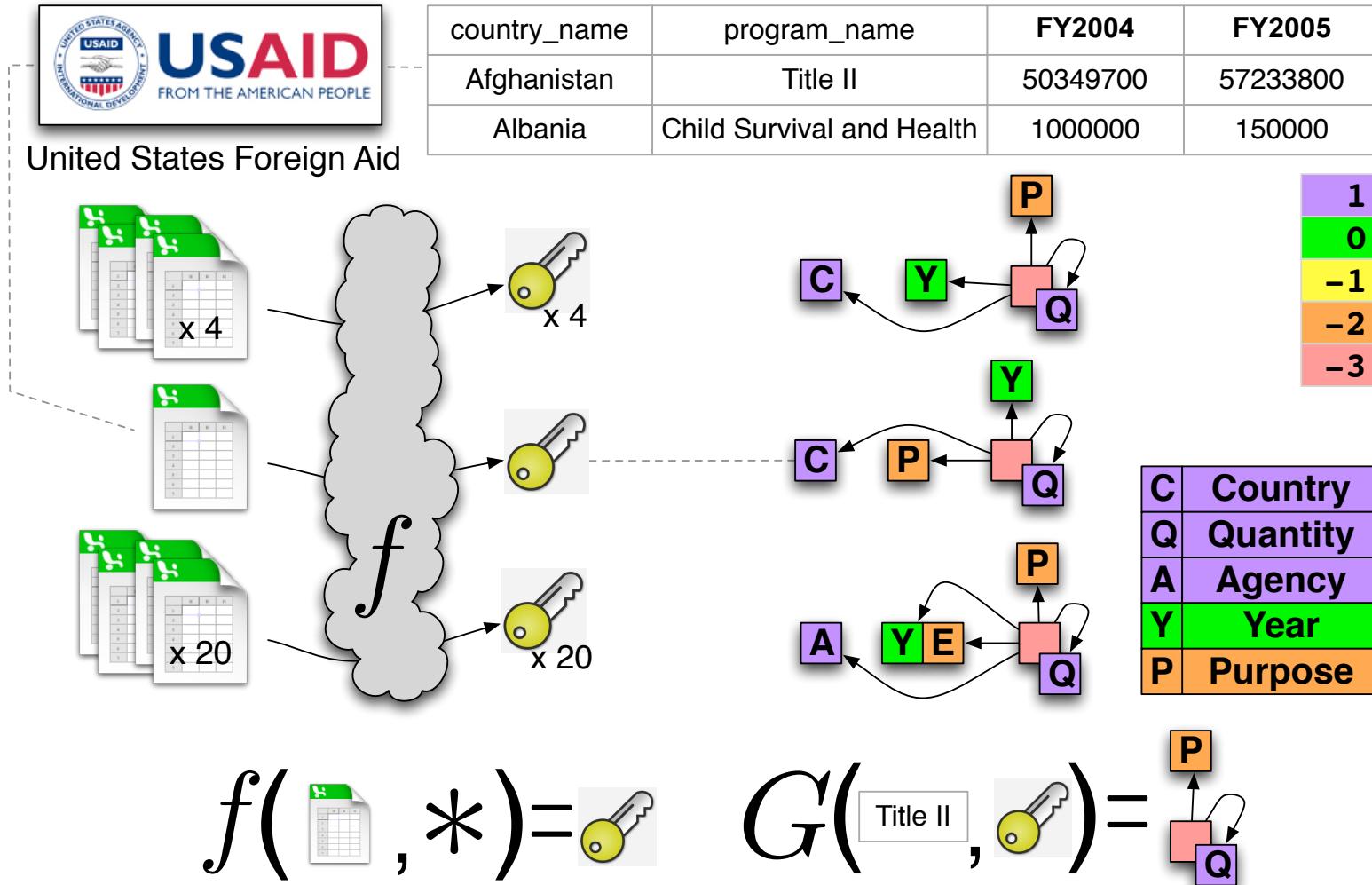


# Benefit of Contextualized Data: A Race to Visuals



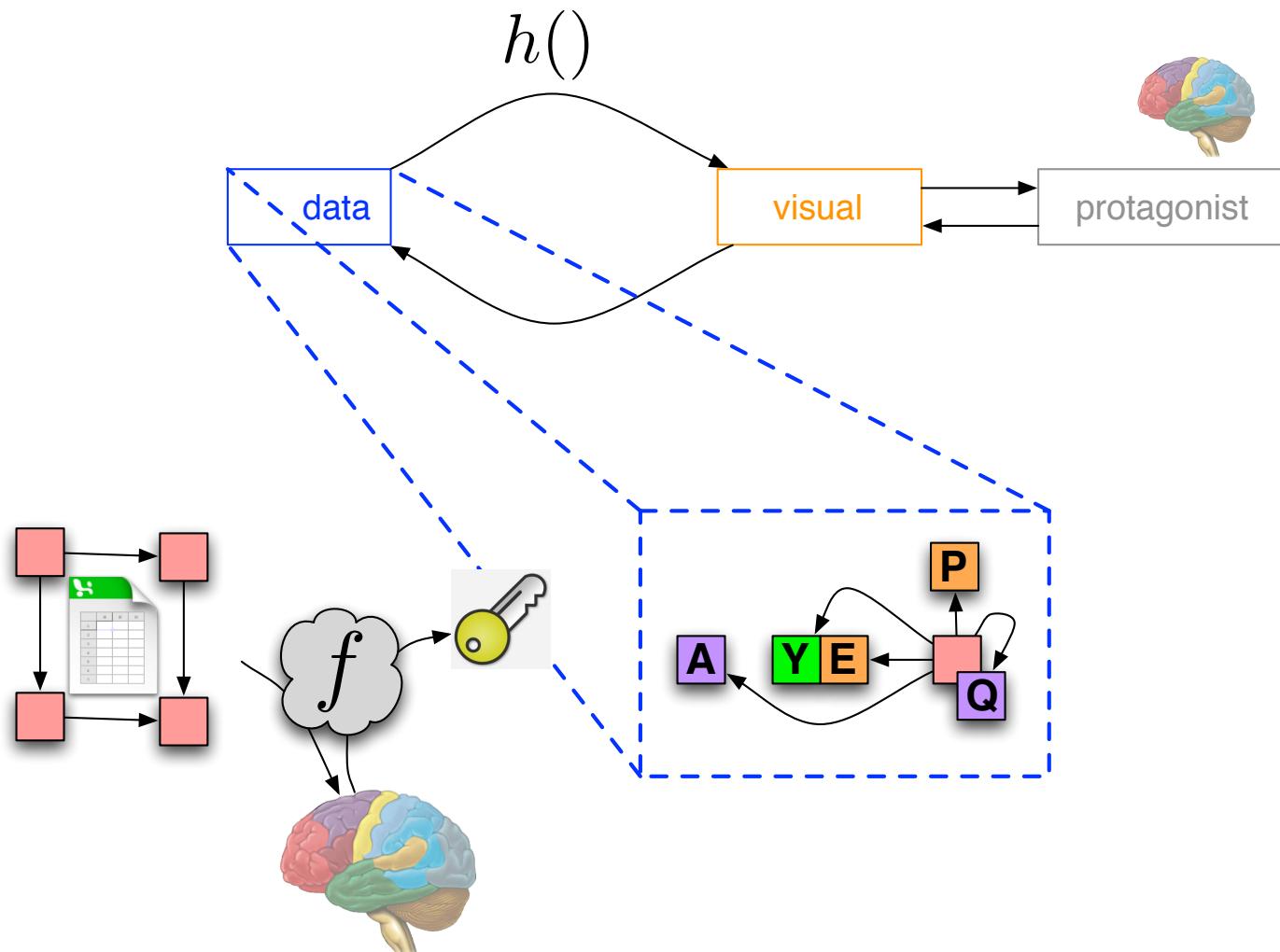


# Unifying Data with $f$ and $G$



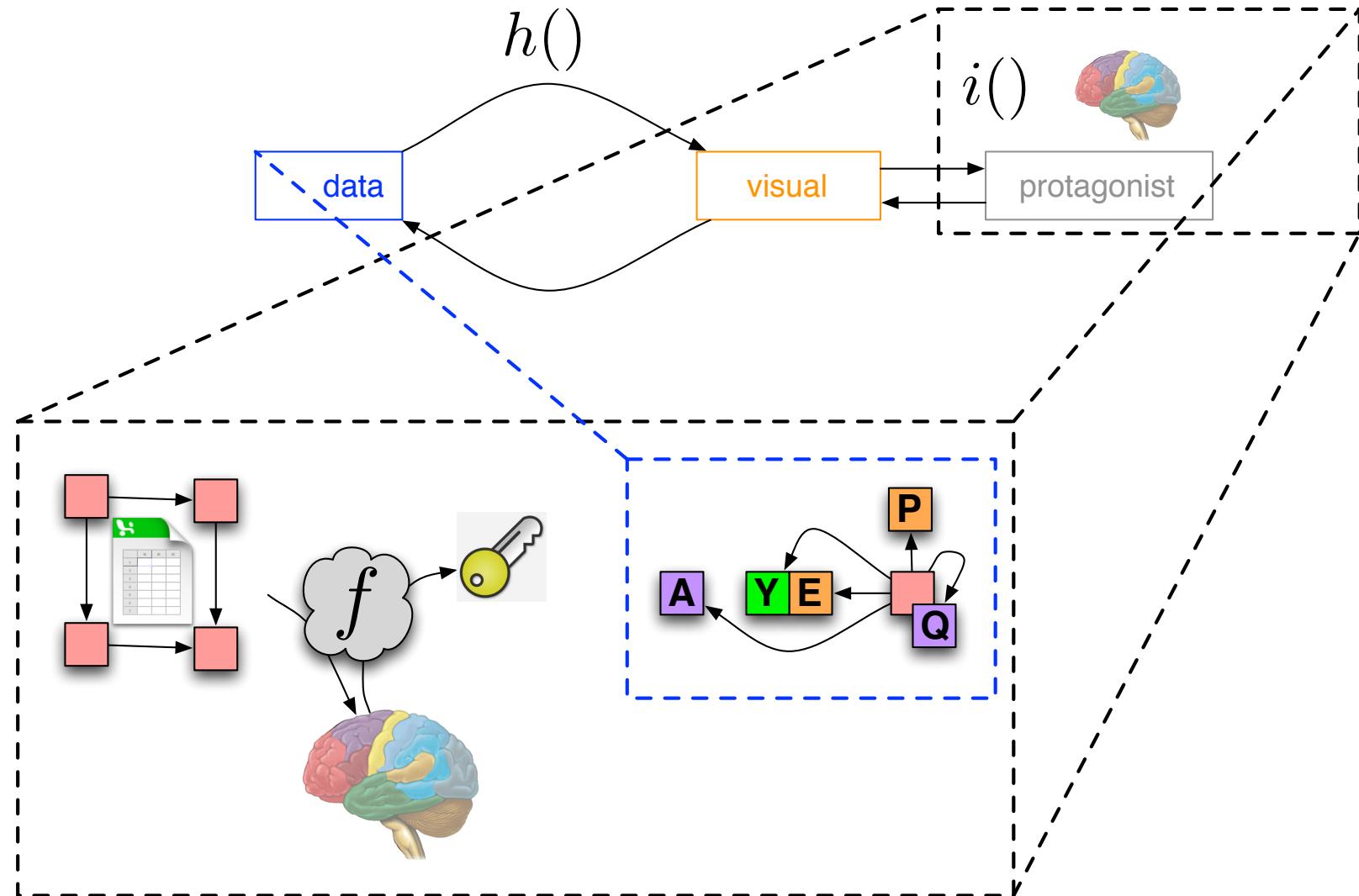


# Next Steps: $h()$





# Next Steps: $i()$





## 2 SioS (SADI in one Slide)

<http://sadiframework.org/registry/register/>

```
raptor -g -o turtle http://sadiframework.org/examples/hello
```

```
...
mygrid:inputParameter [
    mygrid:objectType
        <http://sadiframework.org/examples/hello.owl#NamedIndividual>;
    ...
];
mygrid:outputParameter [
    mygrid:objectType
        <http://sadiframework.org/examples/hello.owl#GreetedIndividual>;
    ...
];
...
```



```
...
# send.ttl
<http://purl.org/twc/instances/TimLebo>
a hello:NamedIndividual;
foaf:name "Tim" .
```

```
curl -d @send.ttl.rdf http://sadiframework.org/examples/hello
```

```
<rdf:RDF    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
               xmlns:hello="http://sadiframework.org/examples/hello.owl#">

    <hello:GreetedIndividual rdf:about="http://purl.org/twc/instances/TimLebo">
        <hello:greeting rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
            Hello, Tim!
        </hello:greeting>
    </hello:GreetedIndividual>
</rdf:RDF>
```