

RAIRD

Remote Access Infrastructure for Register Data

Documenting the life-cycle of every data point

Ørnulf Risnes,
Norwegian Social Science Data Services (NSD)
<ornulf.risnes@nsd.uib.no>

EDDI 2015
Copenhagen 3 December 2015



Statistisk sentralbyrå
Statistics Norway



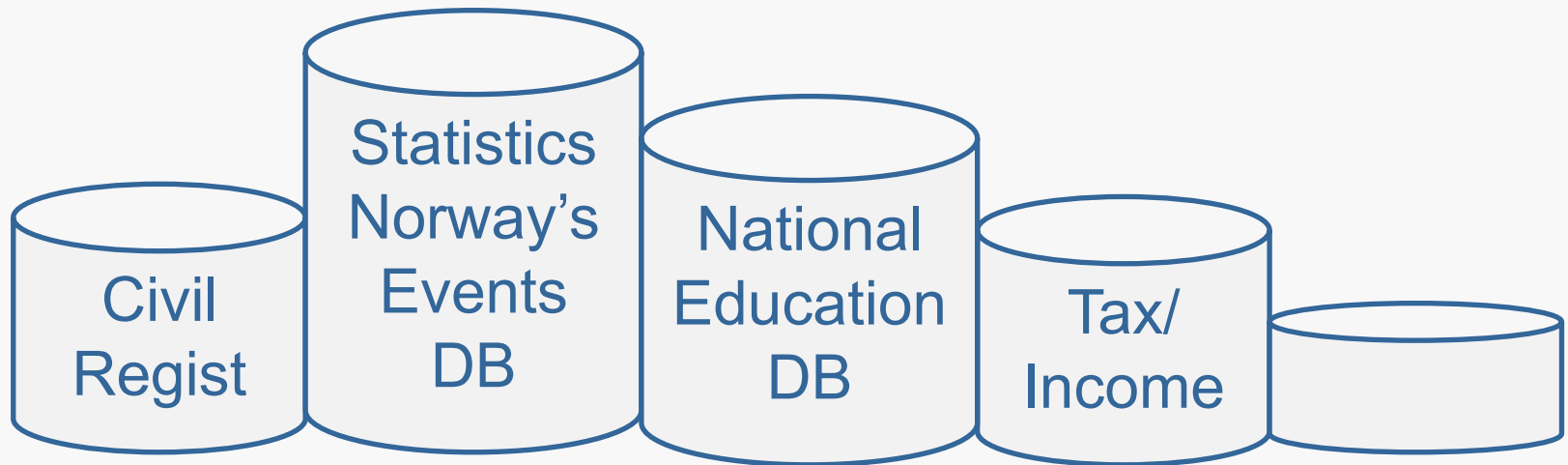
RAIRD
(See <http://raird.no>)

NORD-i
(Data Curation Platform)

DDI Moving Forward
(Data Description)

Goal of RAIRD:

Increase research on data from administrative registers



Statistisk sentralbyrå
Statistics Norway

Initial exposure to the problem



50-60 SAS-files

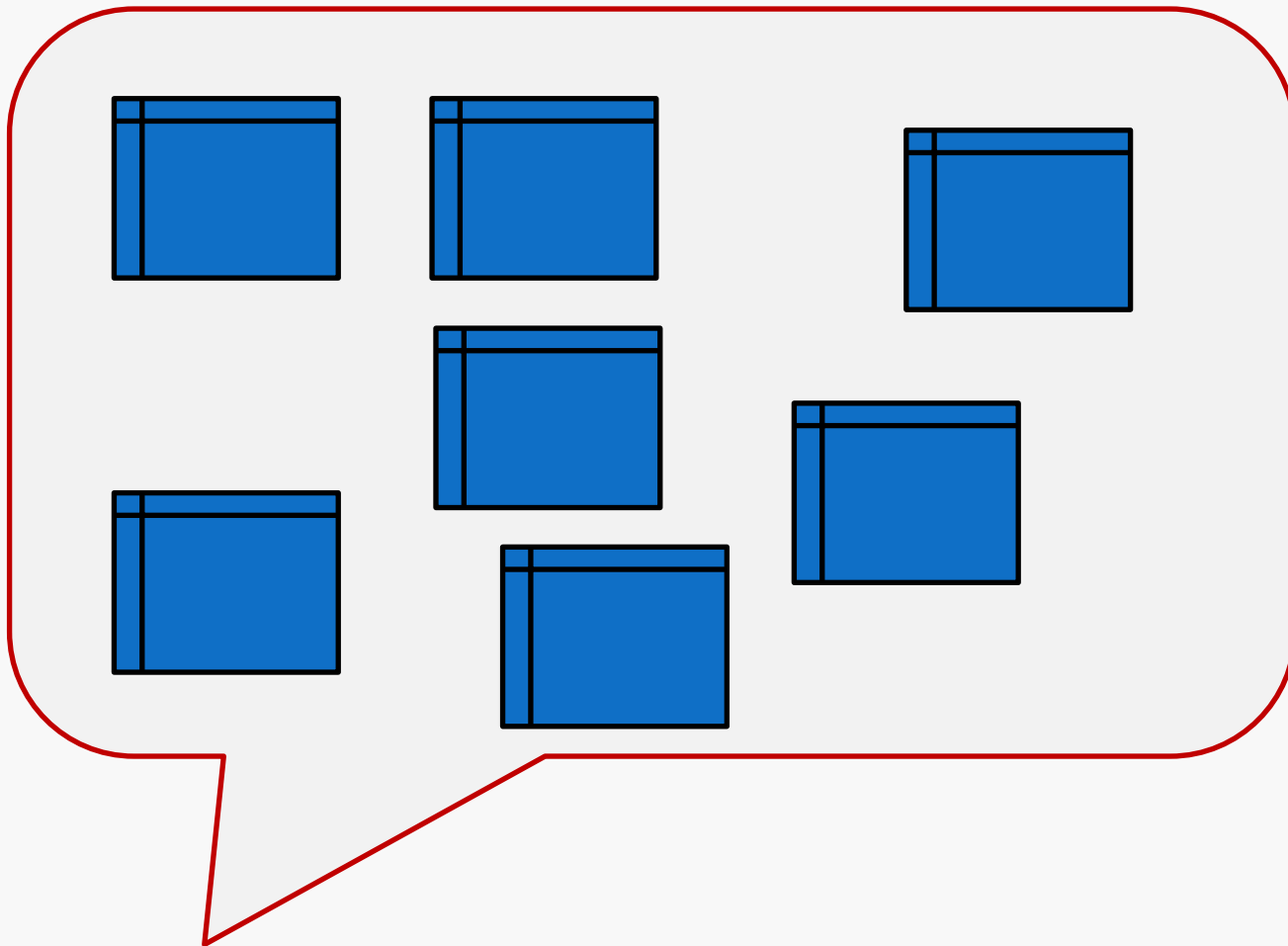
and a PDF



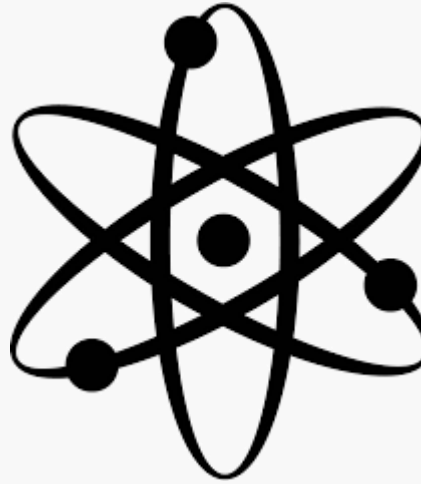
More abstractions

or

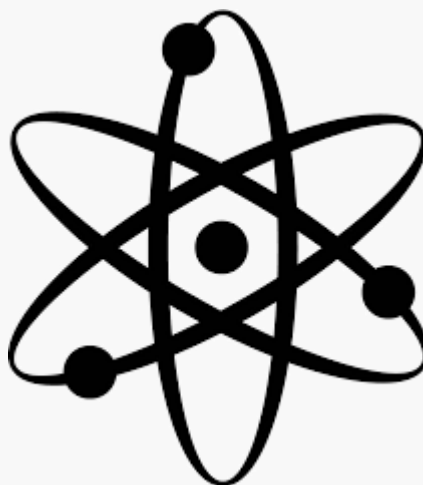
Deeper understanding?



GSIM had «Datum»



Datum



Datum
- the key to
understanding (all?)
data

Datums are everywhere



Multi-level event-history data

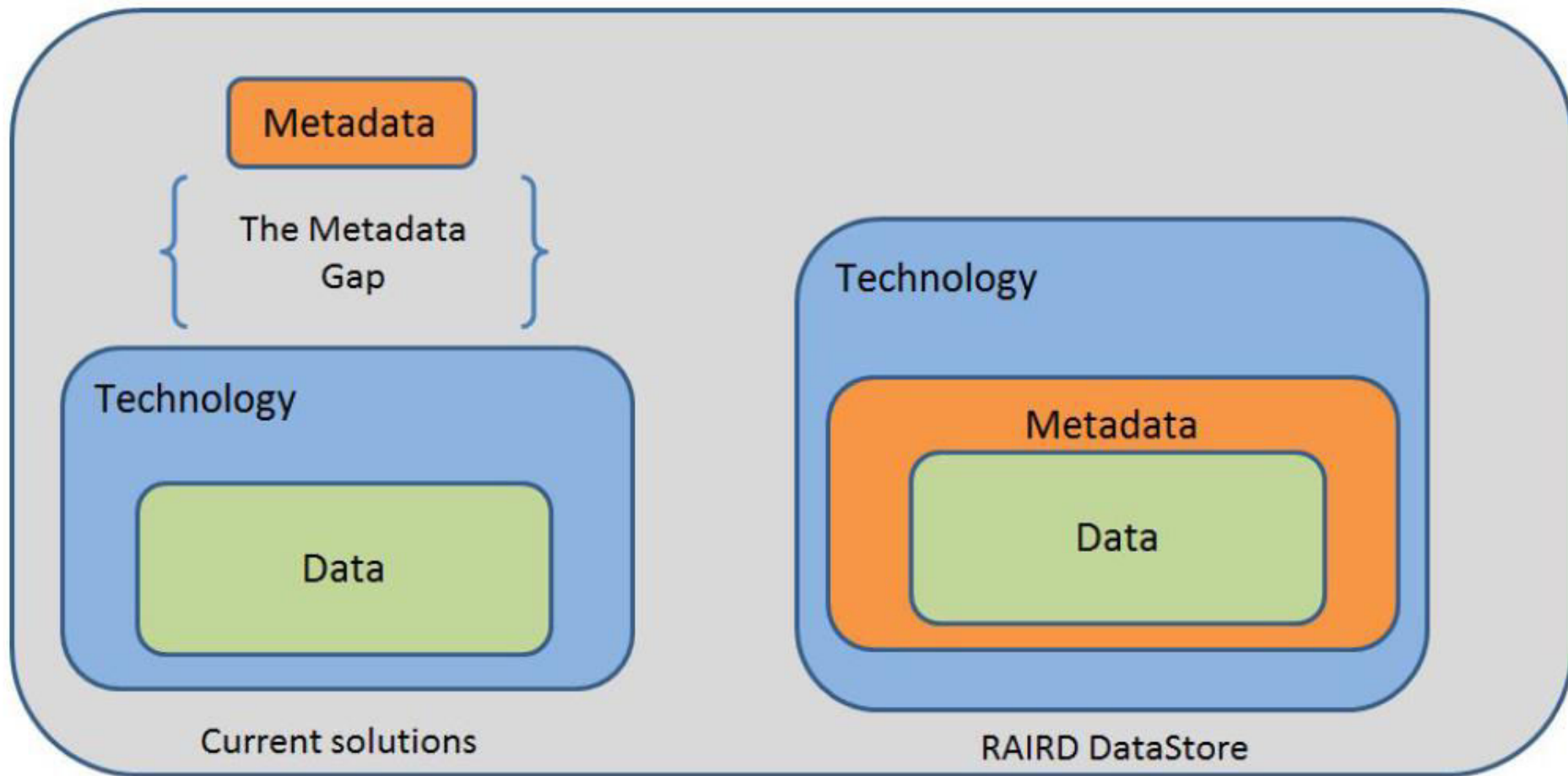
Longitudinal

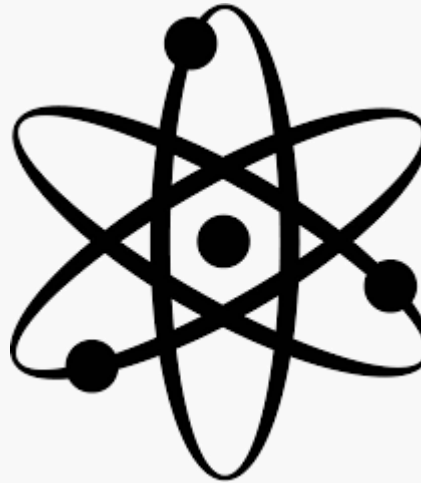
Cross-sectional

Aggregate

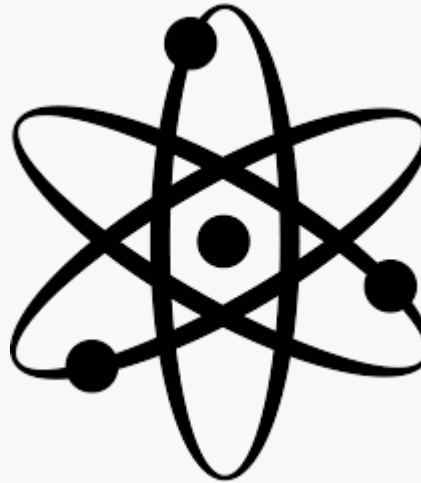
DDI Data Description to the Datum-level?

R A I R D



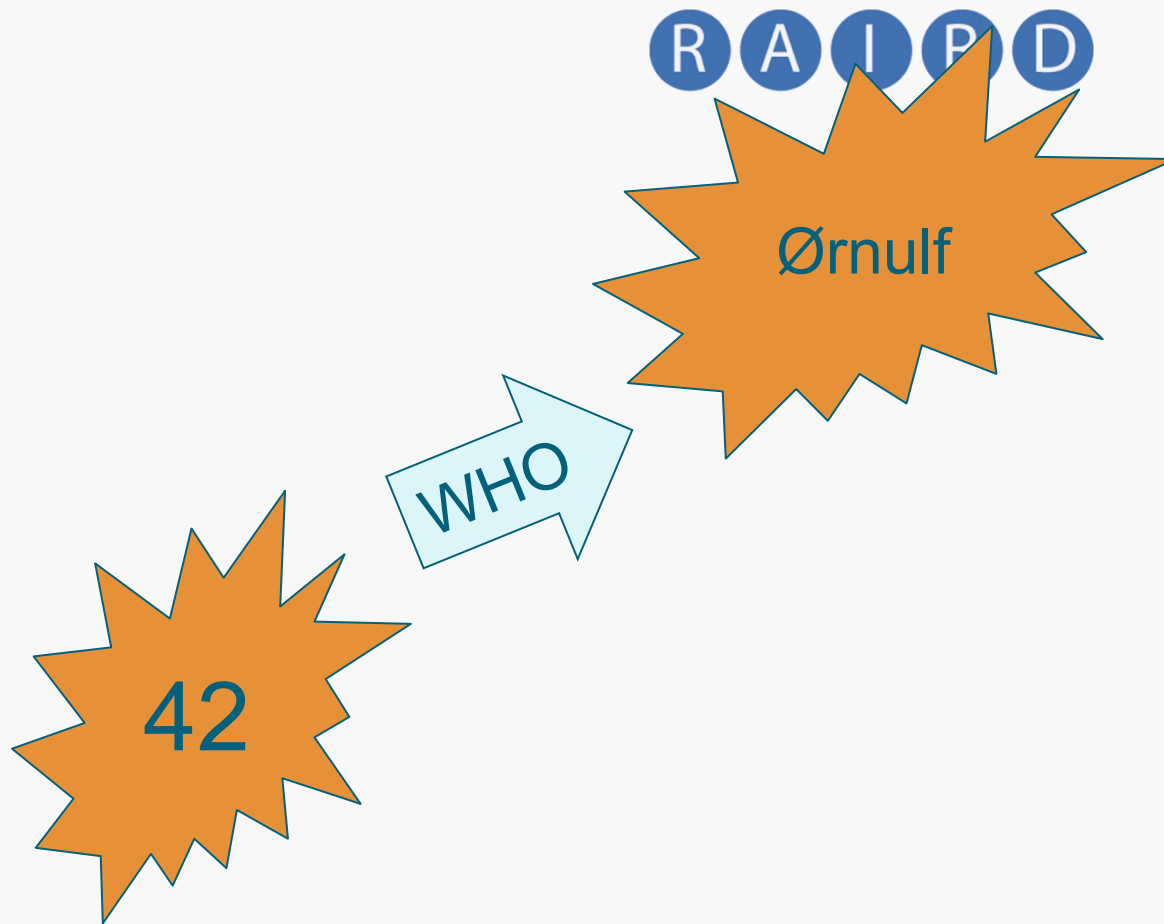


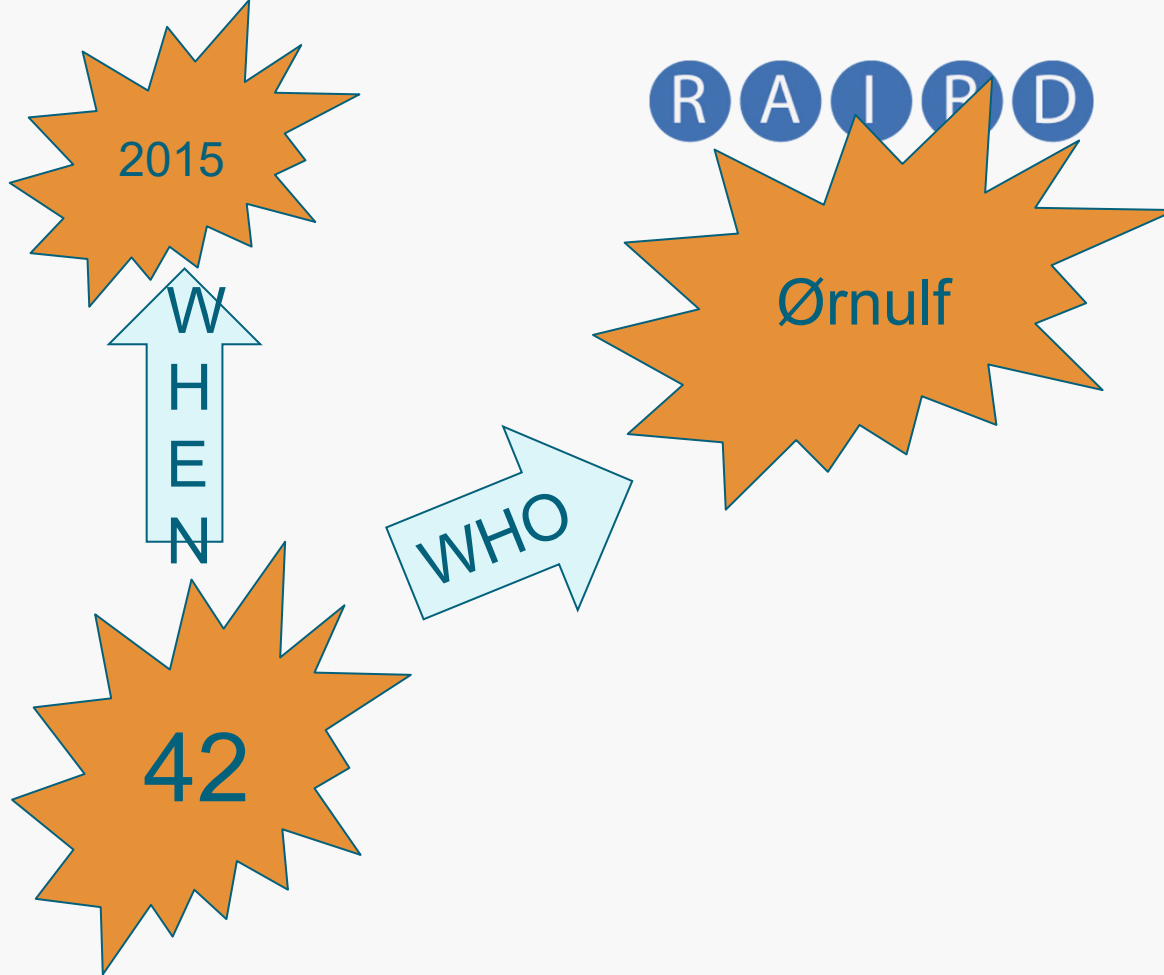
Bottom-up approach

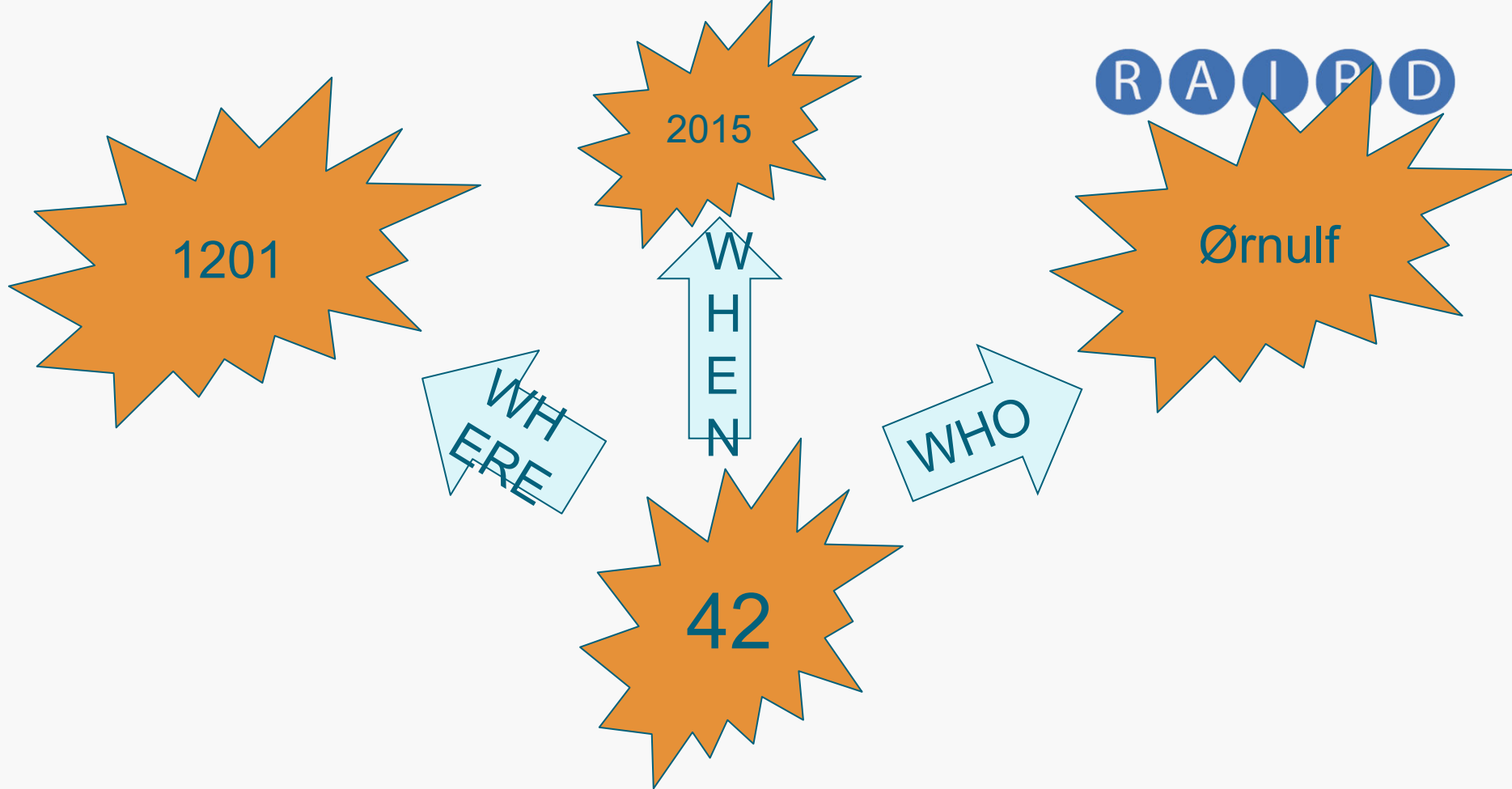


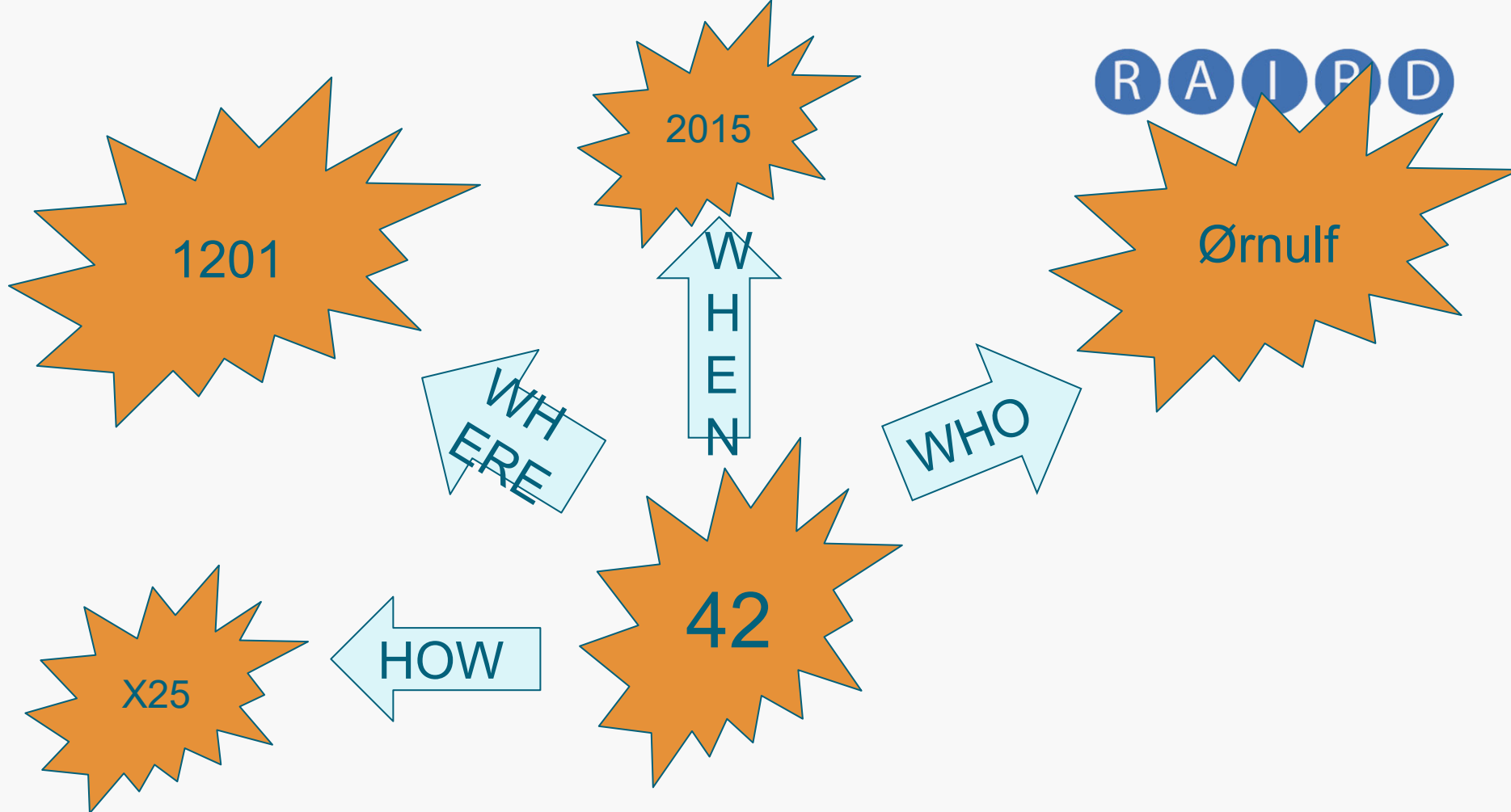
How Datums are related

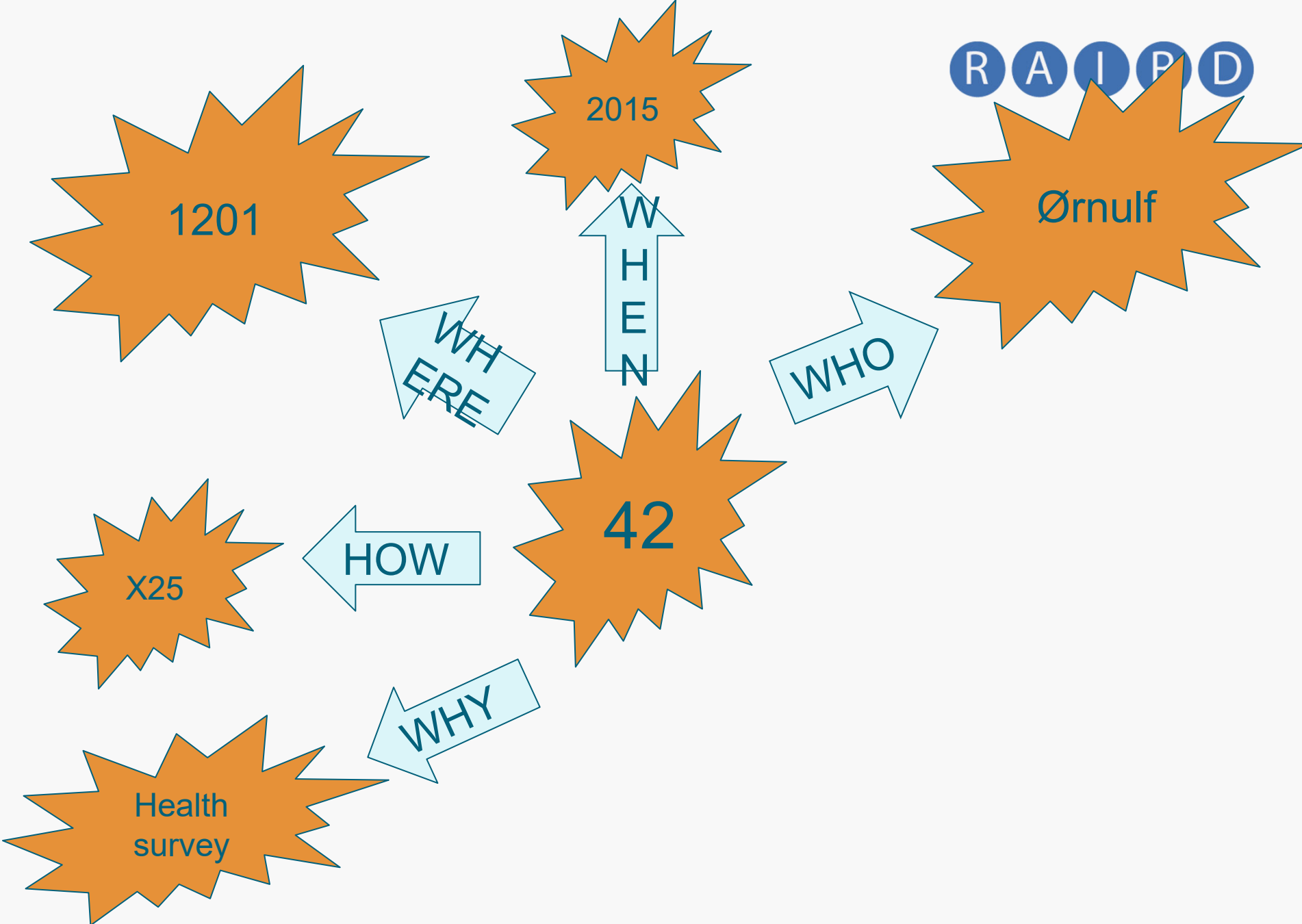


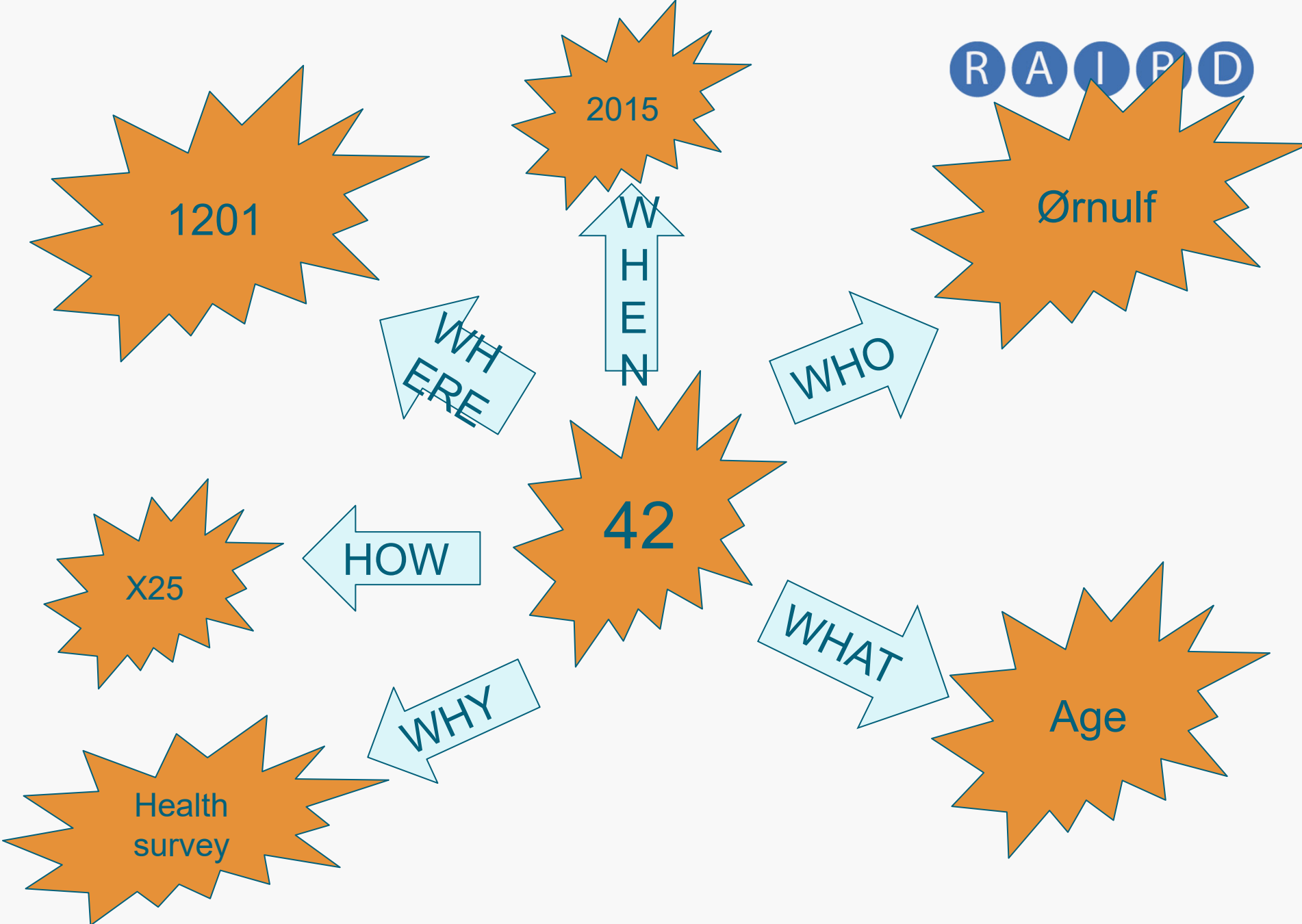












R A I R D

1201

2015

Ørnulf

W
H
E
N

W
H
E
R
E

W
H
O

42

H
O
W

X25

W
H
A
T

Age

W
H
Y

Health
survey

RAIRD

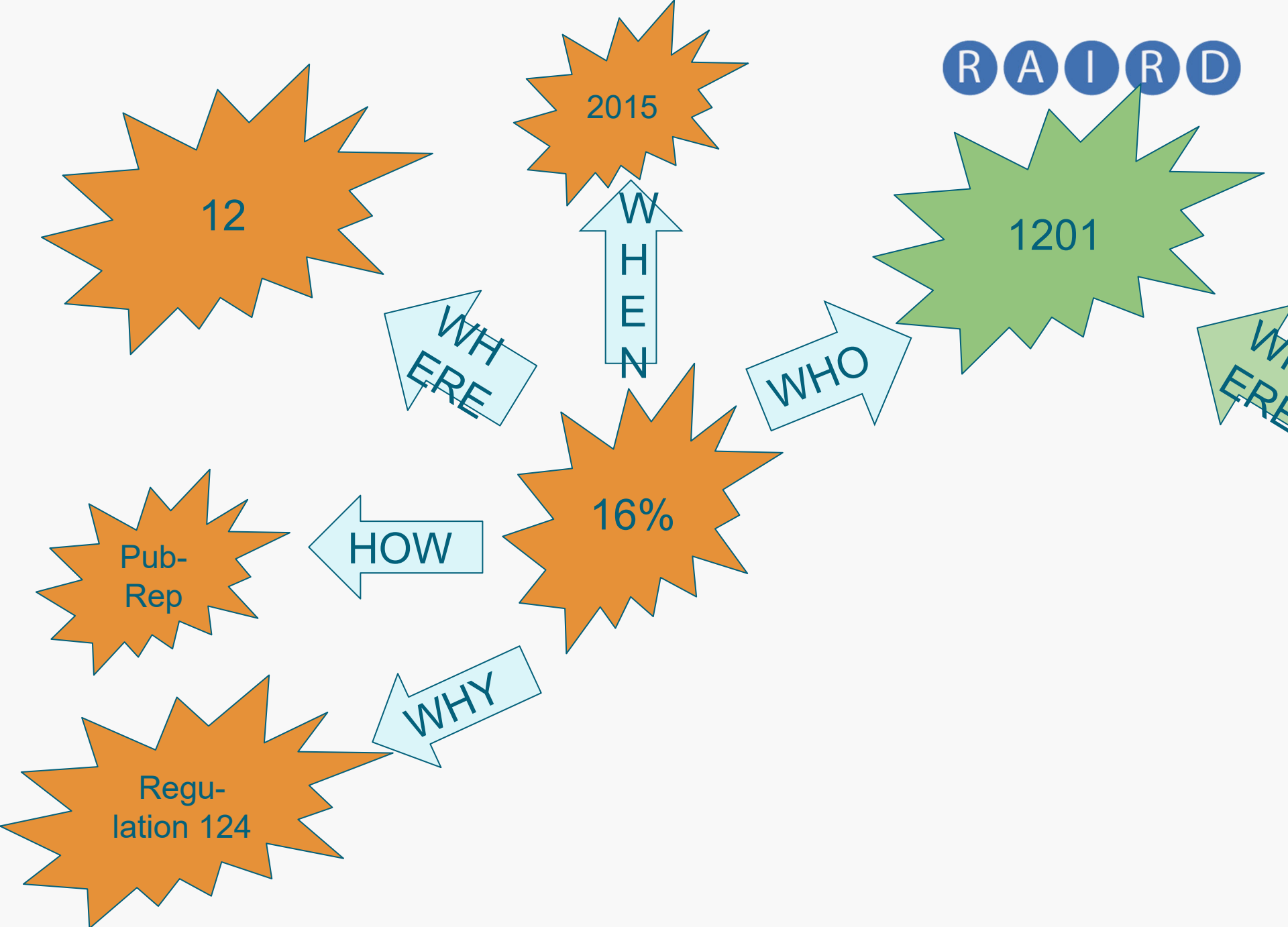
1201

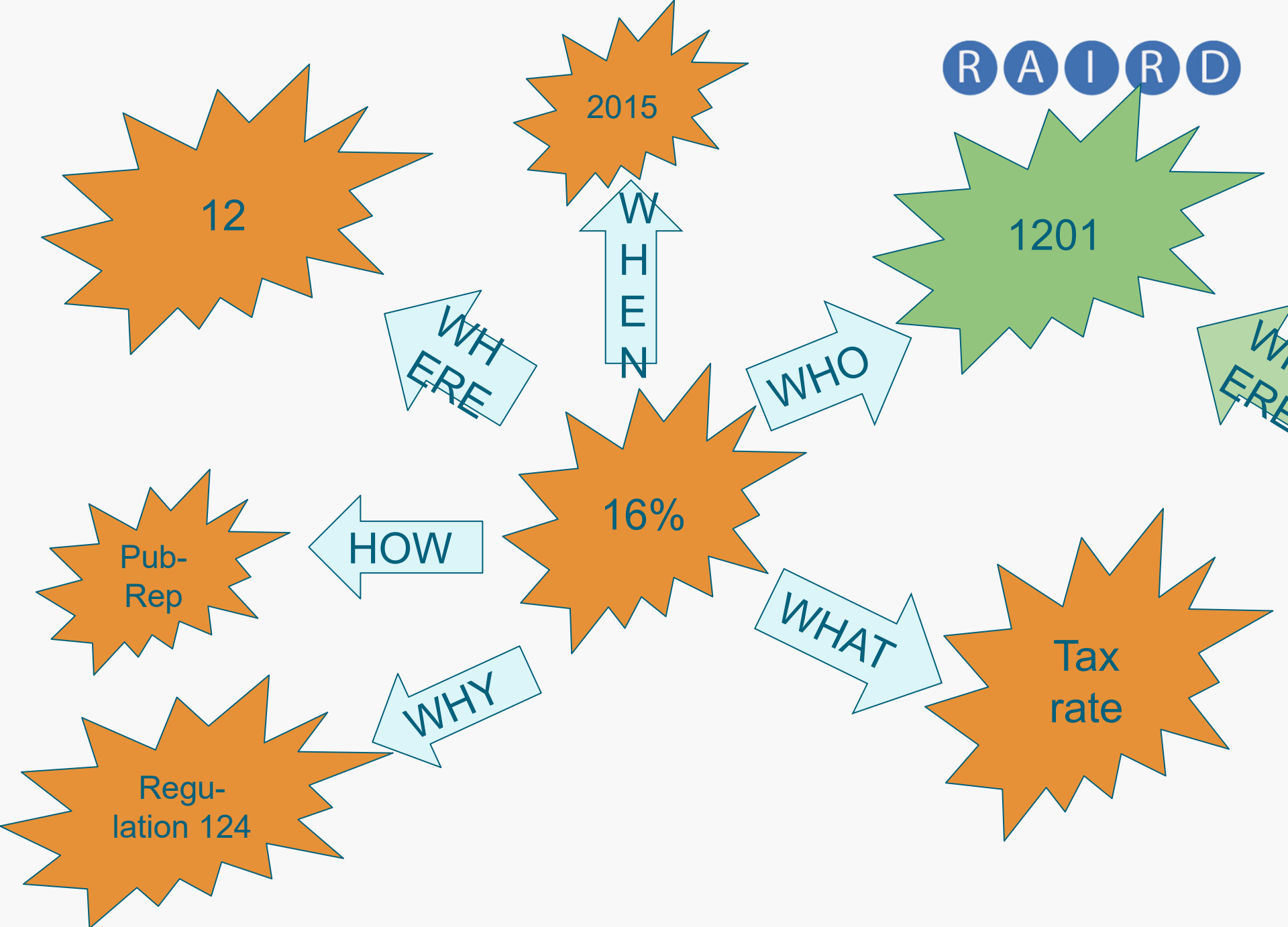
16%

W
ERE

RAIRD







Every Datum needs context

Every Datum is a potential
Metadatum

Datums and metadatums

R A I R D

Conceptual
Variable

Conceptual
Variable

Conceptual
domain

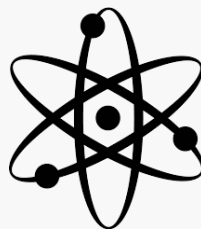
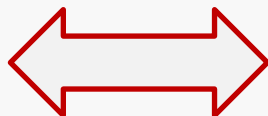
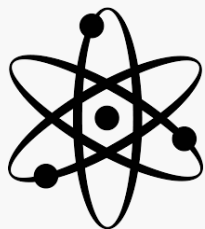
Represented
Variable

Represented
Variable

Value
domain

Instance
Variable

Instance
Variable



DDI Data Description
can bridge the metadata gap

**A consistent, logical model
for data AND metadata**

Datum-level revision control

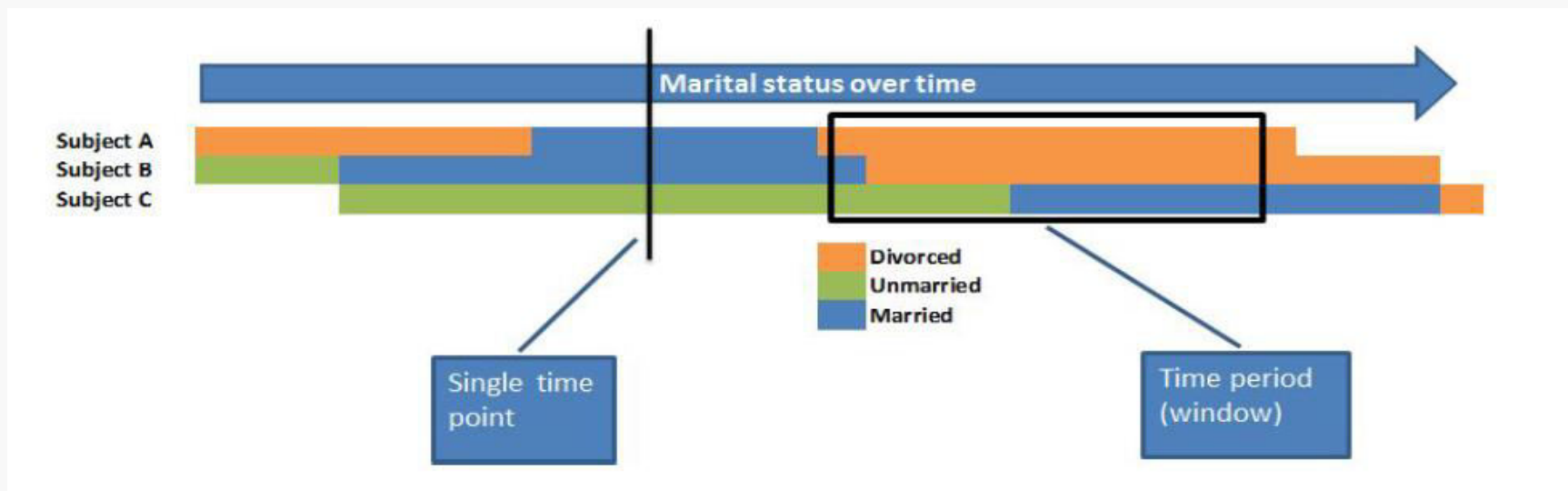
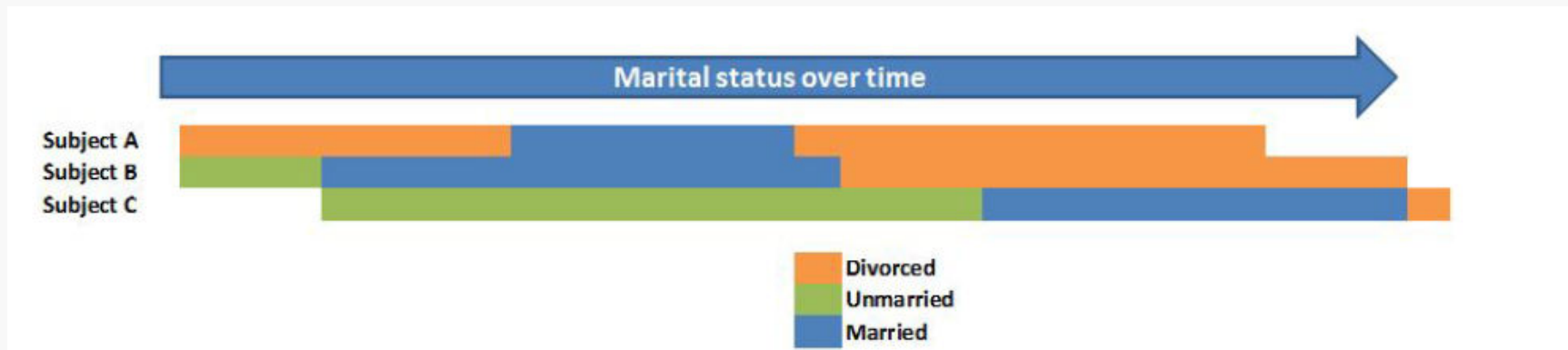
Git-like collaborative workflows

Automated testing

Continuous delivery

Audit trails and activity reports

Flexible query capabilities...



```
import mar_stat 2011-01-01
```

```
import mar_stat
```

```
2016-01-01 to 2009-04-01
```