

REUTERS/Pawel Kopczynski

## DATA INTEROPERABILITY AND LINKED DATA TECHNOLOGIES

Dave Weller
Chief Enterprise Architect
Thomson Reuters
EFSA Conference, Milan, 15 October 2015





### THOMSON REUTERS AND OUR DATA CHALLENGE

# 6 Million

**UPDATES PER SECOND** 

across 50 million instruments in

THOMSON REUTERS

**ELEKTRON**"

1.2 Billion
CITATIONS in
WEB OF SCIENCE™

9 Billion

DOCUMENTS in WestlawNext®

500,000

PHOTOS AND IMAGES are captured and published by



1 Billion

**PEOPLE** reached by Thomson Reuters news and insights every day 30,000

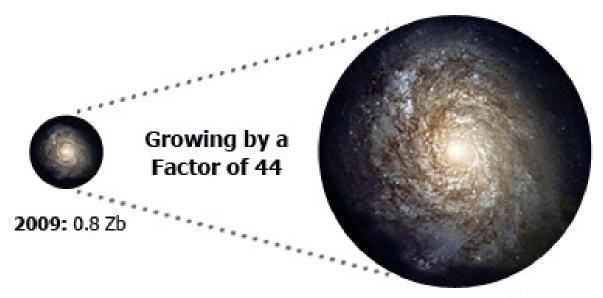
SOURCES

of data from feeds, filings, firms, and funds



## THE EVOLUTION OF DATA IN THE WEB

- Massive growth of open data
   Opportunity is huge... in the web...
  - But race to keep track and create understanding is being lost
- - But bets are off. As it stands the race is being lost



2020: 35.2 7b

Photo credit: Debriefing Software



## THE SEARCH FOR MEANING



- **European Federation of Sea Anglers**
- **Egyptian Financial Supervisory Authority**
- **European Food Safety Authority**
- Stella Polare:

**Congress Centre** 

**North Star** 

**Italian Navy Yacht** 

## IDENTIFIERS – THE KEY TO LINKING DATA

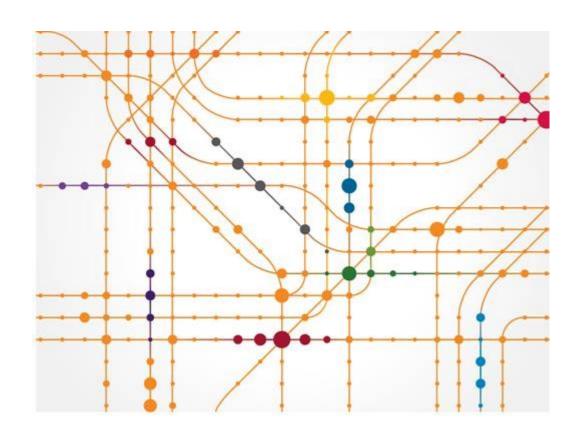
- At web scale, identifiers are critical
- Data discovery relies on open data and open identifiers
- Important roles for:
  - Data Consumers
  - Data Publishers
  - Identifier Publishers



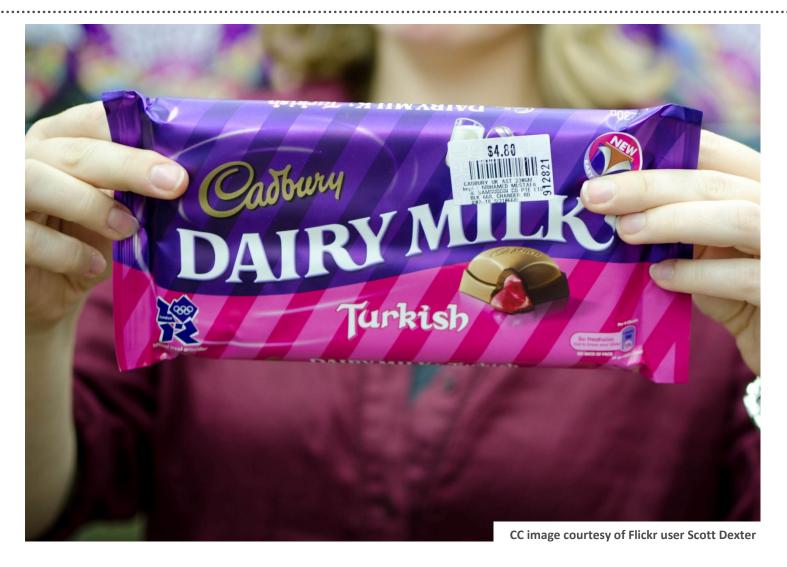


## INFORMATION IDENTITY AS A NEW DISCIPLINE

- Information identifiers are not just primary keys
- Identifier schemes explain how unique labels are assigned to things
- Essential to fulfill the potential of open data

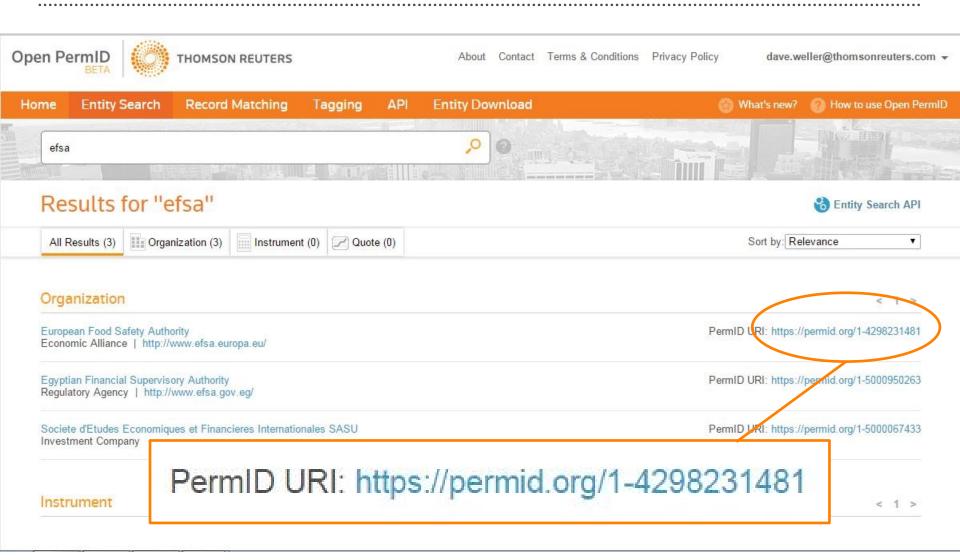


## WHAT IS THIS?





### OPEN PERM ID





## THOMSON REUTERS INTELLIGENT TAGGING

#### **Unstructured Content**

Content uploaded from news articles, blog postings, proprietary data, catalogs, social media, and more











Extract, classify, and tag metadata

#### ACME Corp. Announces Big Expansion

Figs., 2010. All 40 Vol. PC 20 A SEC OFFICE (1 ) a recognition for the things of the sale of shape of \$10.5 Vol. Section (1) and \$10.5 Vol. Section (1) and

the "MPD JAMEST MANY TO A SHARE THE PLAN AND THE PLAN AND

Benediction of the Control of the Co

and St. V. M. Stein, P. M. Stein, S. M. Stein, Companion, American State Stein, and Stein State Stein State State

Data MANE (1998) 1998 (1997) has proposed in the sky of the other lates of the second of the second

And the property of the proper

### Structured Metadata

Our unique identifiers leverage the deep knowledge in Thomson Reuters professional data, creating metadata to enrich your own content – and also mapping it to Thomson Reuters content to give the best of both worlds.



Events

Companies

People

Relationships



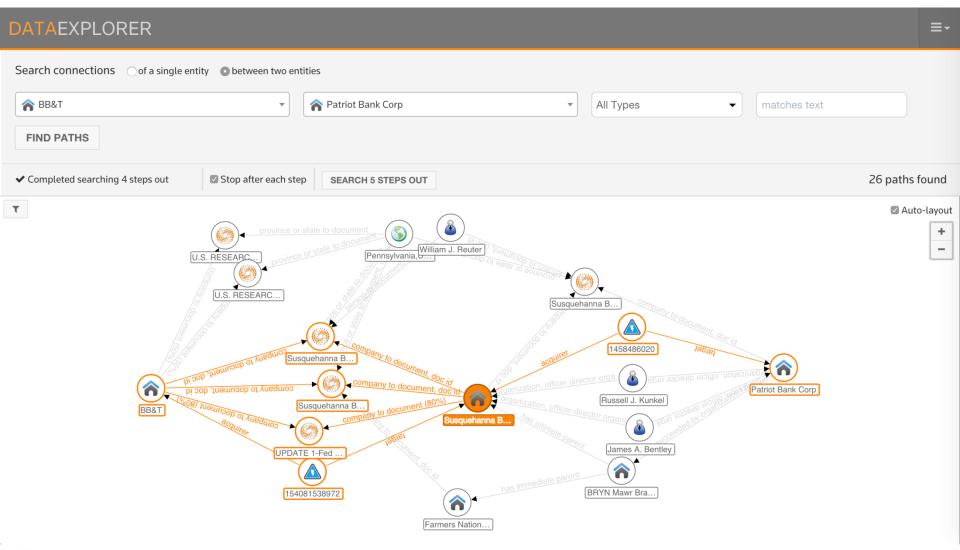


#### **Trusted Source**

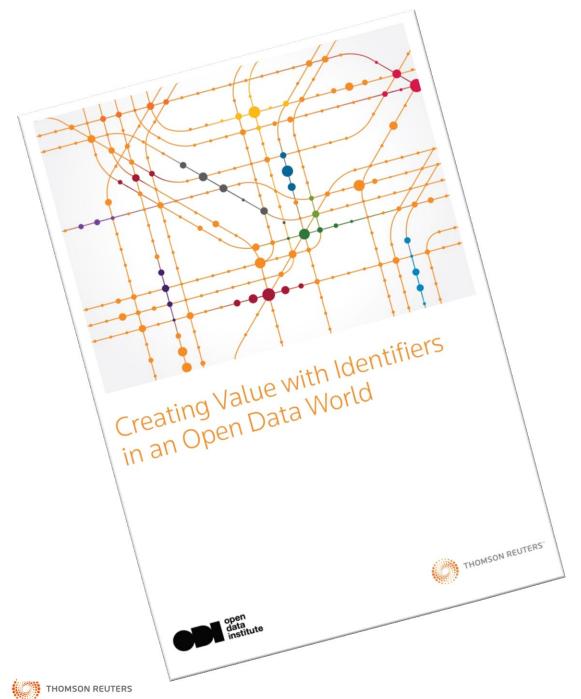
The Thomson Reuters key advantage is assigning unique identifiers, or **PermIDs**, which go beyond keywords, returning the right connections you'd otherwise miss.



## THOMSON REUTERS DATA FUSION







### **HELP TO GROW**

THE OPEN DATA COMMONS

### MANAGE THE COMPLEXITY

OF THE WEB

### **DISCOVER AND CONNECT**

TO OTHER DATA SOURCES

### **EXPLORE OPEN PERMID**

HTTPS://PERMID.ORG/

