

# Dataset Dynamics Demo

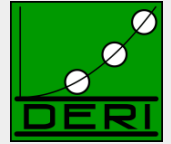
*Keepin' up with a LOD of changes*

Jürgen Umbrich & Michael Hausenblas

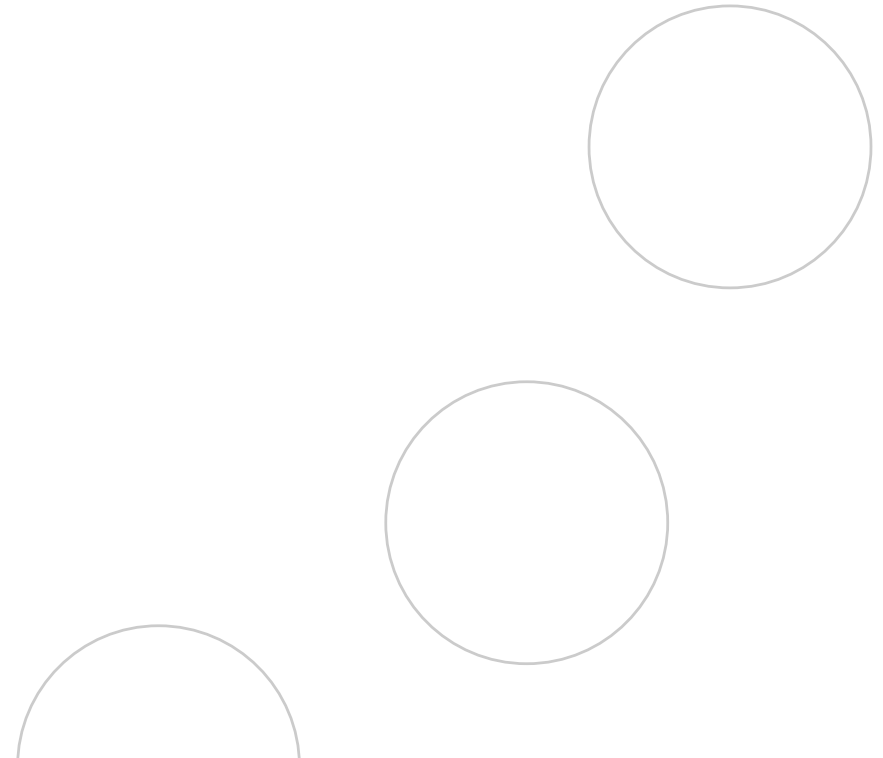
<http://linkeddata.deri.ie>



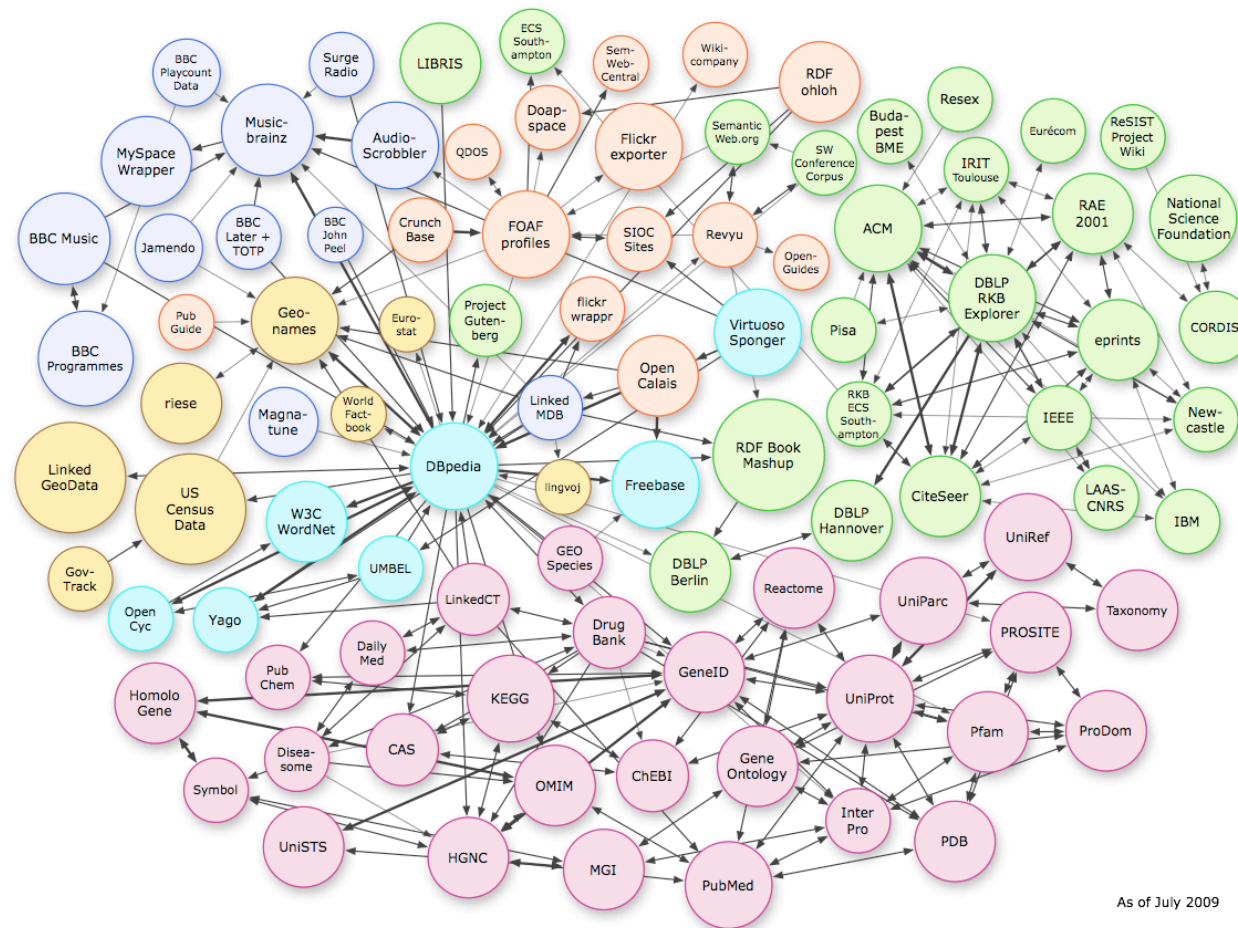
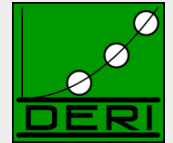
# Outline



- The Problem – datasets changes
- Solutions (?)
- LDC09: DCM/DCW demo



# The Problem I

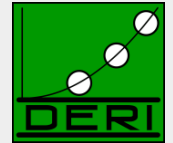


As of July 2009

In late 2009: we have over **100** open datasets providing over **6.7 billion** RDF triples, interlinked by some **142 million** links ...

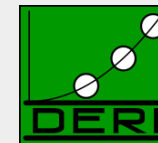
*courtesy of Richard Cyganiak and Anja Jentzsch*

# The Problem II



- **Datasets DO change as ...**
  - ... new resources are added
  - ... new links are set
  - ... old stuff is deprecated/removed
- **The changes may happen because ...**
  - They are planned (e.g., new data is available from the origin data source, such as in the Eurostat case)
  - It is an inherently dynamic data source (for example Twitter, etc.)
  - Stuff gets fixed, corrected, due to quality reasons
  - Of many other reasons ;)

# The Problem III



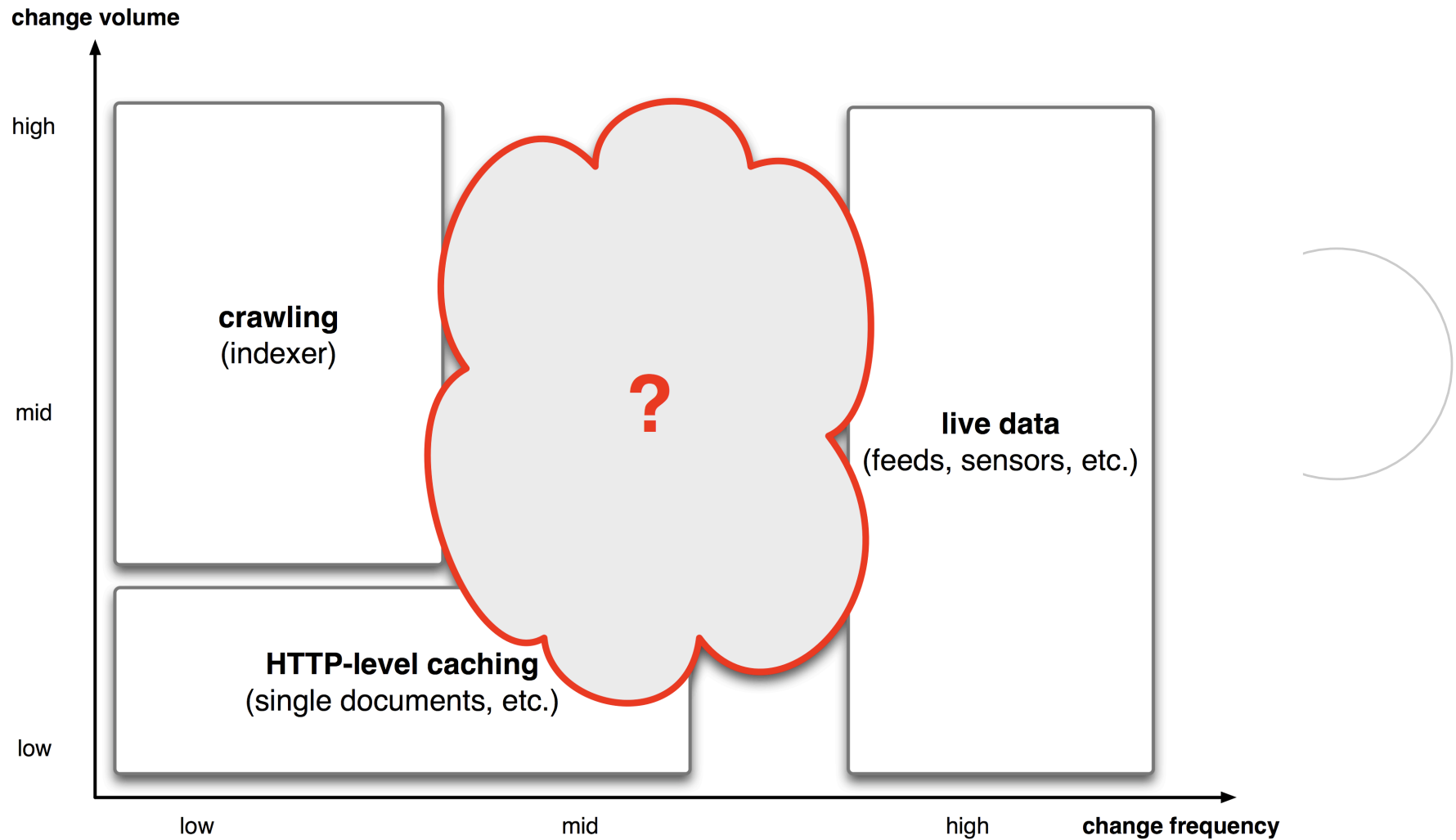
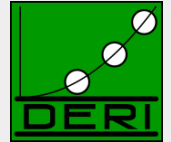
The key question now is:

*How does a consumer of the dataset (such as other datasets linking to it, applications using it, 3<sup>rd</sup> party services, etc.) learn:*

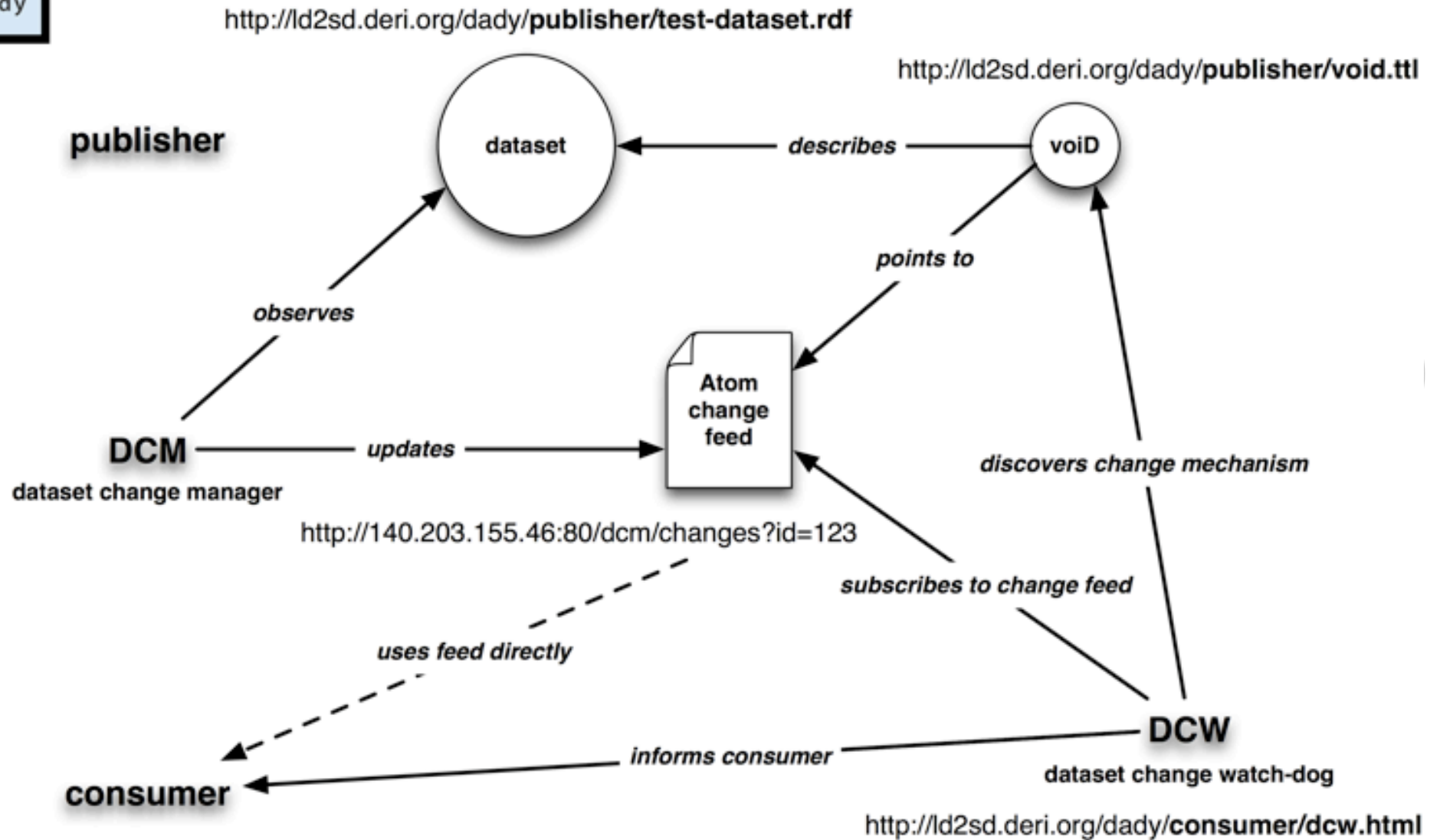
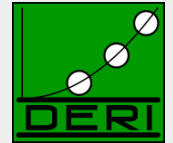
WHEN something has changed?

WHAT has changed?

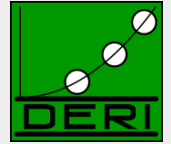
# Solutions?



# LDC09: DCM/DCW demo



# Links and Feedback



- <http://esw.w3.org/topic/DatasetDynamics>
- <http://code.google.com/p/dady/wiki/Demos>
- General discussion:  
[public-lod@w3.org](mailto:public-lod@w3.org)
- Feature requests, etc.:  
[dataset-dynamics@googlegroups.com](mailto:dataset-dynamics@googlegroups.com)
- Twitter, identi.ca, etc.:  
#dady #linkeddata

