



↔ <http://bit.ly/lebo-issues-2013>

Content-Preserving Graphics



Timothy Lebo
Tetherless World Constellation
Rensselaer Polytechnic Institute



Rensselaer



Motivation

- Technical barriers **slow analysts** down in *creating, understanding, trusting, and re-purposing* results.
- Linked Data offers a huge **potential** for establishing explicit, understandable connections within and across data sources.
- Linked Data is **ready to use**, **but not yet useful**.
- Linked Data currently suffers from an **irony of use**.

Hypothesis: Applying Linked Data principles to the analytical process **reduces the time required** for **analysts** to *create, understand, trust, and re-purpose* any existing result.



Practical Challenges in Visual Analytics

Among **35** data analysts from **25** commercial organizations:

- Most tedious and time-consuming task is **discovering** and **wrangling** data
- Analytical results are **static**
- Analytical results are **shared** via email, a shared file system, or during group meetings
- Difficulties discovering when **relevant** data becomes available
- Visualizations avoided because considered **a barrier to underlying data**

S. Kandel, A. Paepcke, J. Hellerstein, and J. Heer. Enterprise data analysis and visualization: An interview study. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2917–2926, Dec. 2012.



An Example: How to *understand, trust, and re-purpose* this result?

UCS Satellite Database | [Logout](#)

www.ucsusa.org/nuclear_weapons_and_global_security/space_weapons/technology

Union of Concerned Scientists

[\[About\]](#) [\[Issues\]](#) [\[Join\]](#) [\[Act\]](#) [\[Donate\]](#)

Nuclear Weapons & Global Security

Center for Science and Democracy

TEXT SIZE [A](#) [A](#) [A](#) [PRINT](#)

[Home](#) » Nuclear Weapons & Global Security

UCS Satellite Database

9-13 Satellite Database Downloads
Includes launches through 8/31/13

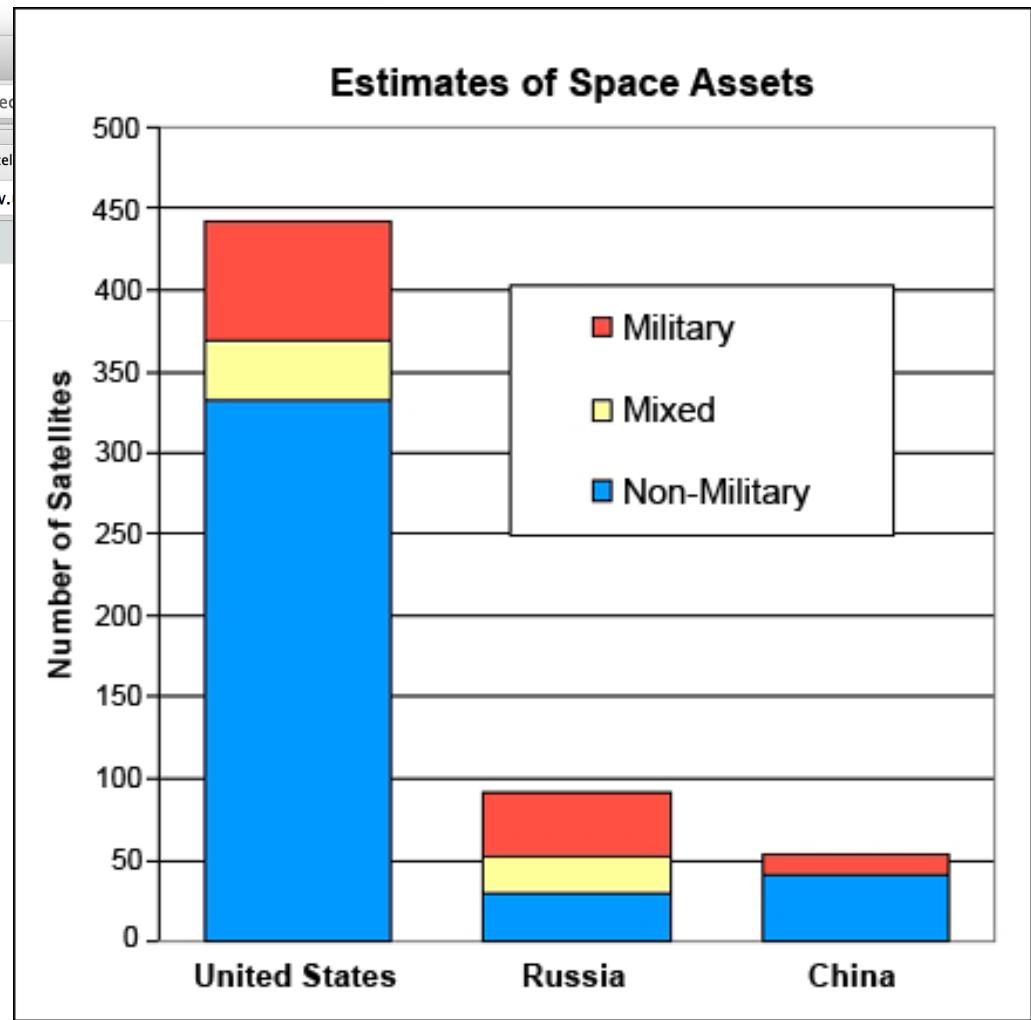
| | |
|---|---|
| Database (Excel format) | Database (text format) |
| Changes to the database (pdf) | Quick Guide to the Database |
| Common misconceptions (pdf) | User's manual |

Get Database Updates
Sign up to be notified when

The UCS Satellites Database tracks more than 1000 objects in space around Earth.

Successes

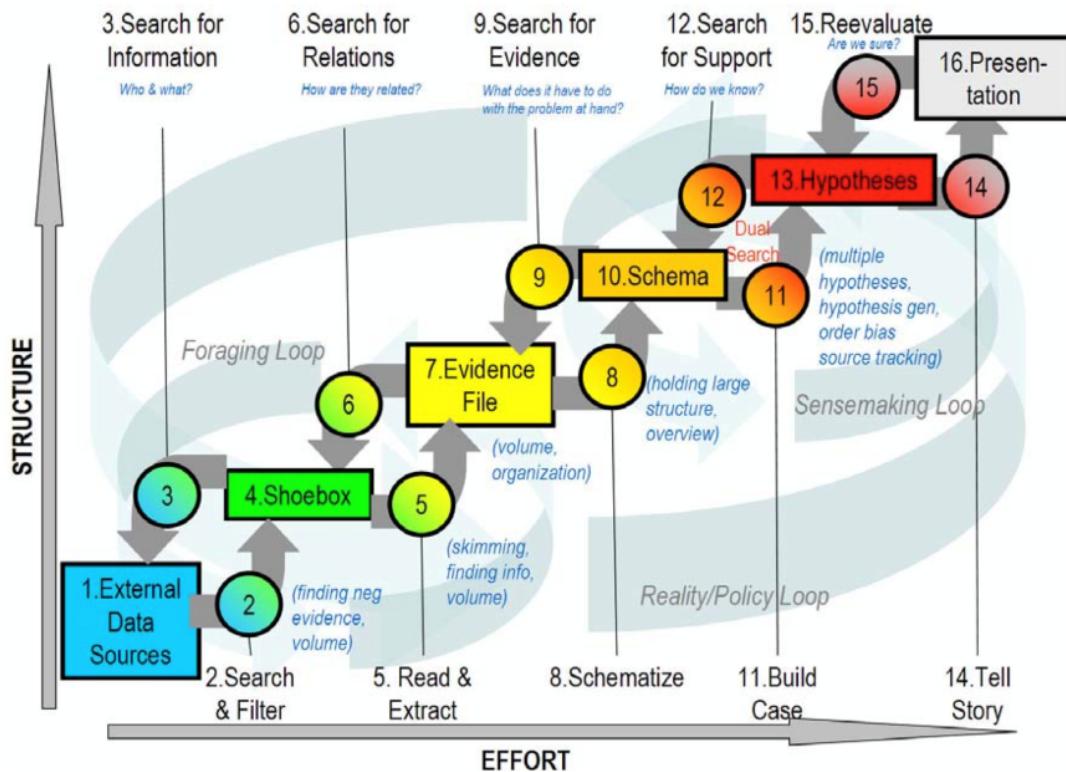
Food and Agriculture





The Analytical Process is Tortuous

Analytical Process: expend **effort** to establish **structure**



- **Bottom-up** seeks, searches, filters *givens*
- **Top-down** adds *structure* based on mental model
- **Iterative, ad hoc**

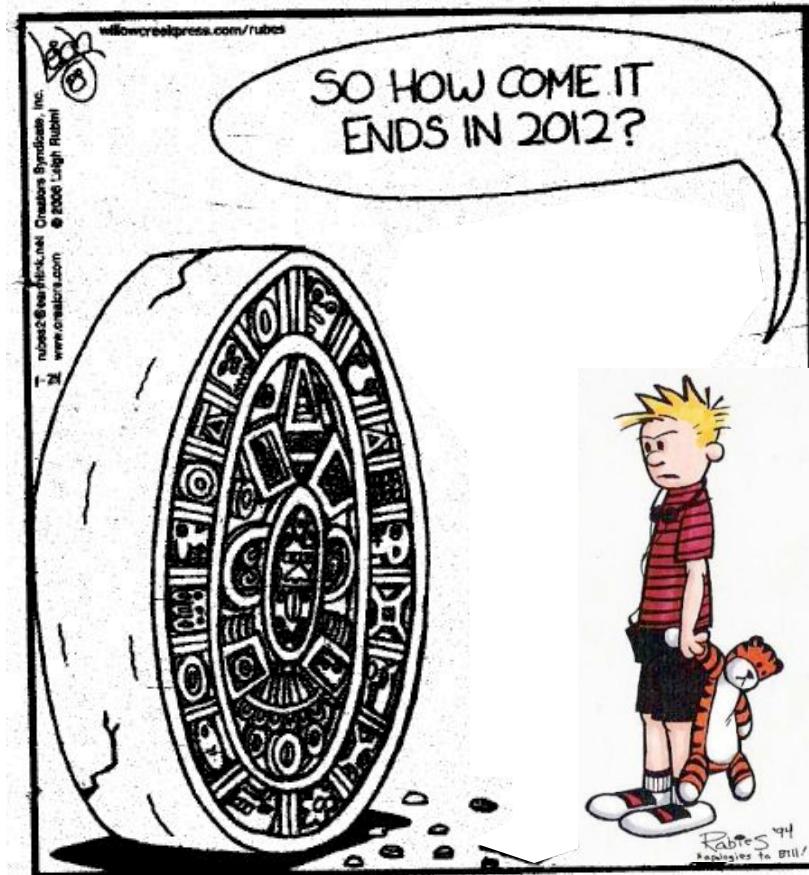
P. Pirolli and S. Card The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis*, 2005.



Interpreting Results is an Archaeological Endeavor



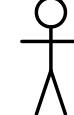
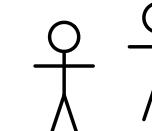
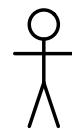
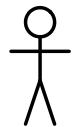
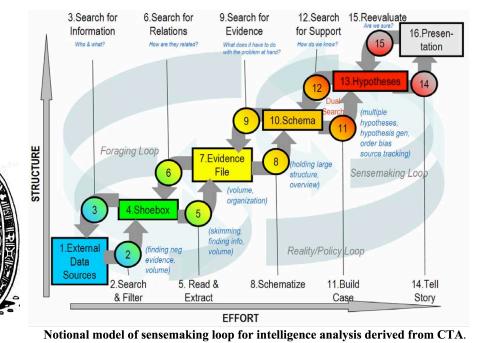
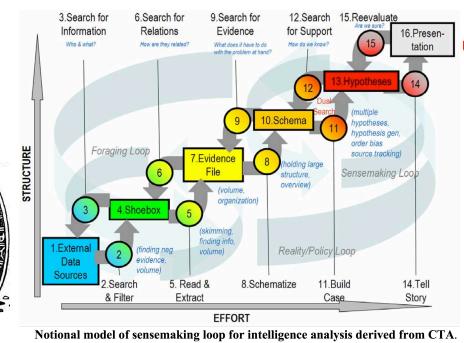
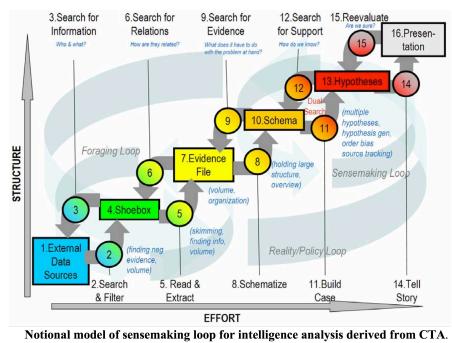
At last, the mystery of the Mayan calendar revealed.



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



Compounded Costs of Reusing Previous Analytical Results





Outline

- Problem: Disconnected results from ad hoc analytical processes
- Solution: Linked Data?



Linked Data Avoids Archaeological Endeavors

Linked Data offers a huge **potential** for establishing explicit, understandable connections within and across data sources.



Include **links** to other URIs, so that people can **discover** more things.



Tim Berners-Lee
w3.org/DesignIssues/LinkedData

Use HTTP URIs to name things, so that people can **look up those names**.

When someone looks up a URI, provide useful information in **RDF**.



Available as **non-proprietary** format.
(e.g. CSV instead of Excel)



Available as machine-readable **structured data**.
(e.g. Excel instead of image scan of a table)

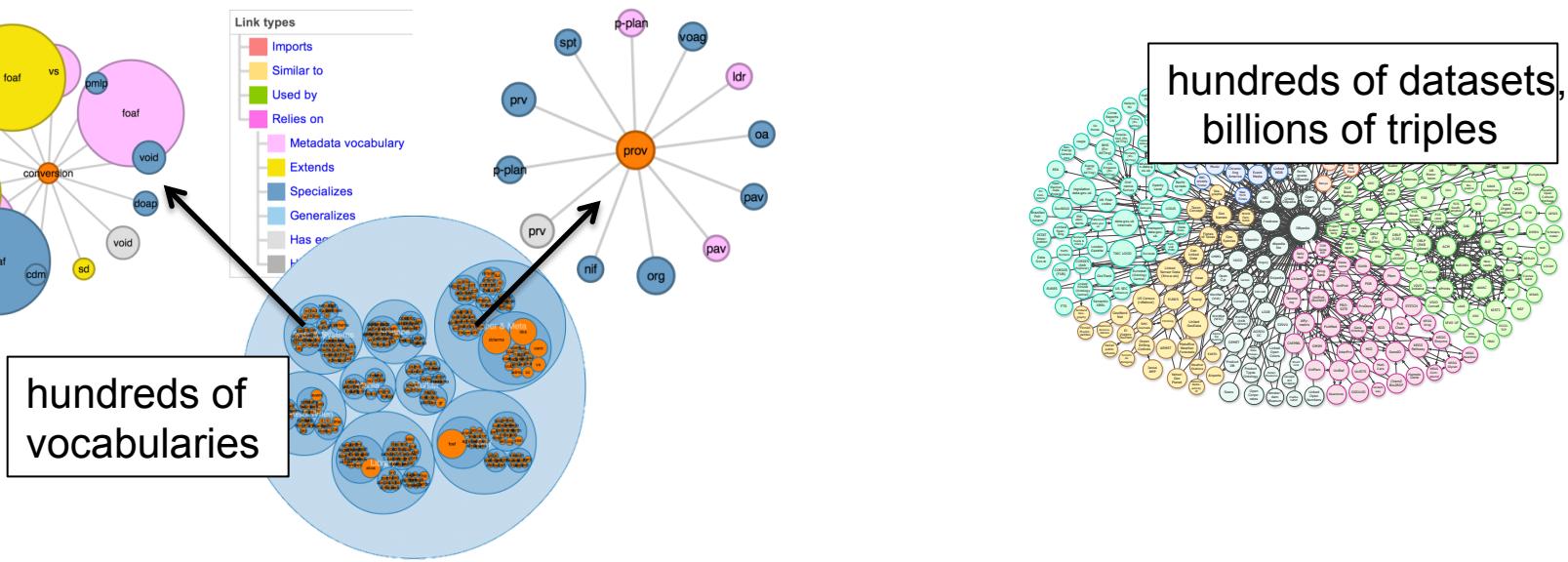


Available on the web (whatever format) but with an open license.

<http://5stardata.info>



Linked Data is *ready for use*, but not yet *useful*

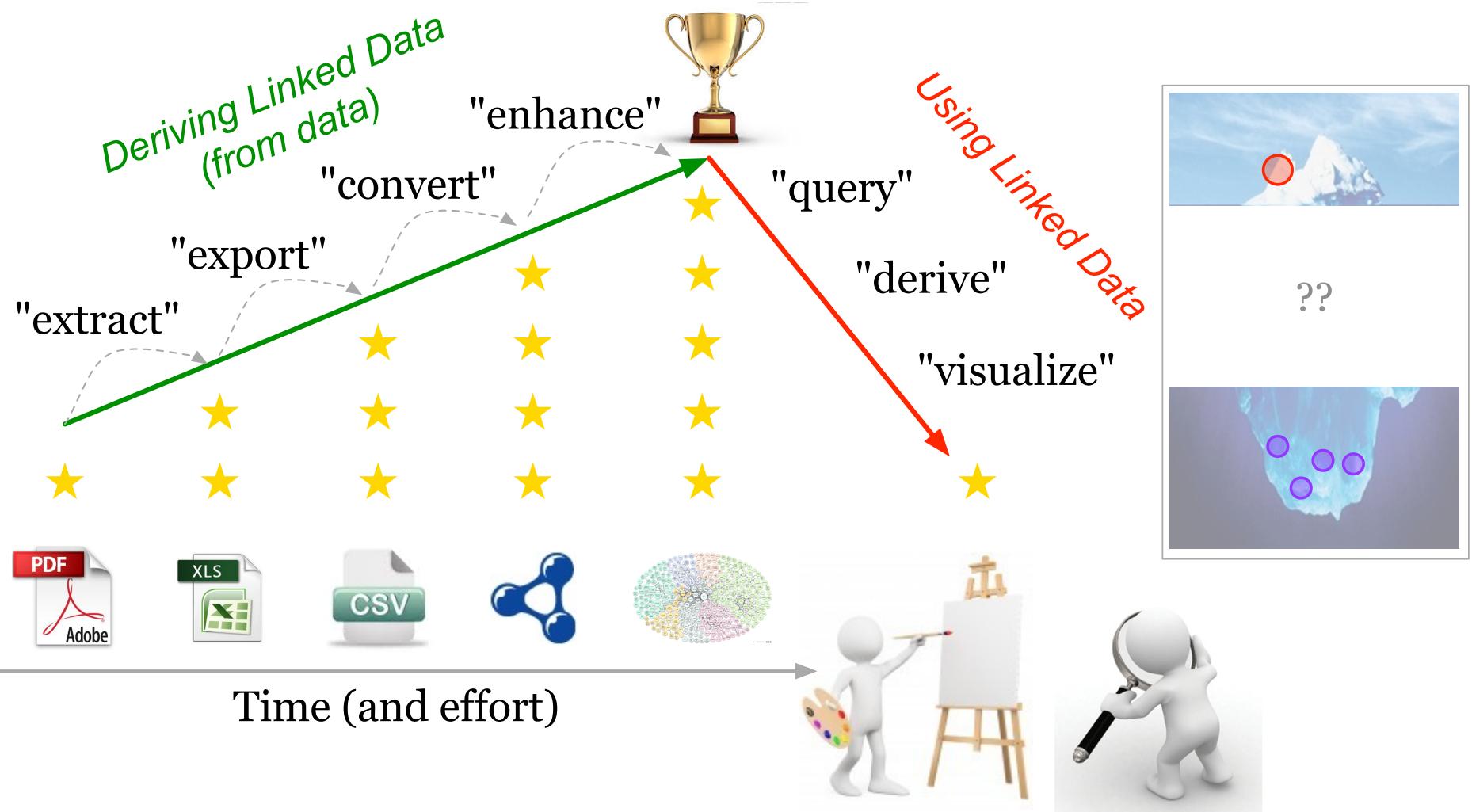


Bernard Vatant, Pierre-Yves
Vandenbussche <http://lov.okfn.org>

Richard Cyganiak,
Anja Jentzsch <http://lod-cloud.net>



The Irony of Using Linked Data





Why Linked Data Is Not Used

- Technical barriers **slow analysts** down in *creating, understanding, trusting, and re-purposing* results.
- Linked Data offers a huge **potential** for establishing explicit, understandable connections within and across data sources.
- Linked Data is **ready to use**, **but not yet useful**.
- Linked Data currently suffers from an **irony of use**.

Content-preserving graphics enables two new kinds of utility:

Interrogate visual results

Re-purpose visual results



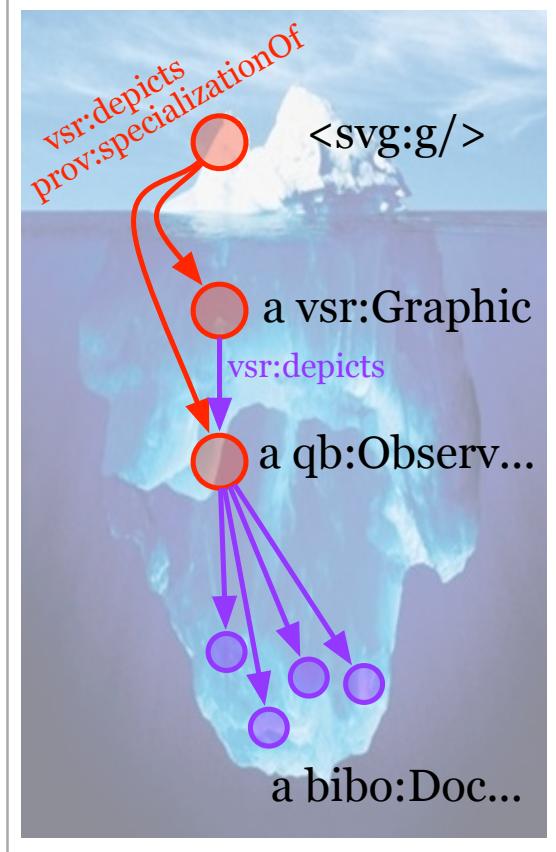
Content-Preserving Graphics

Useful Visual Handles on Useful Portions of LOD

Traditional Linked Data Applications



Content-Preserving Graphics



available by GRDDL

available by dereference

Concrete Visual Information

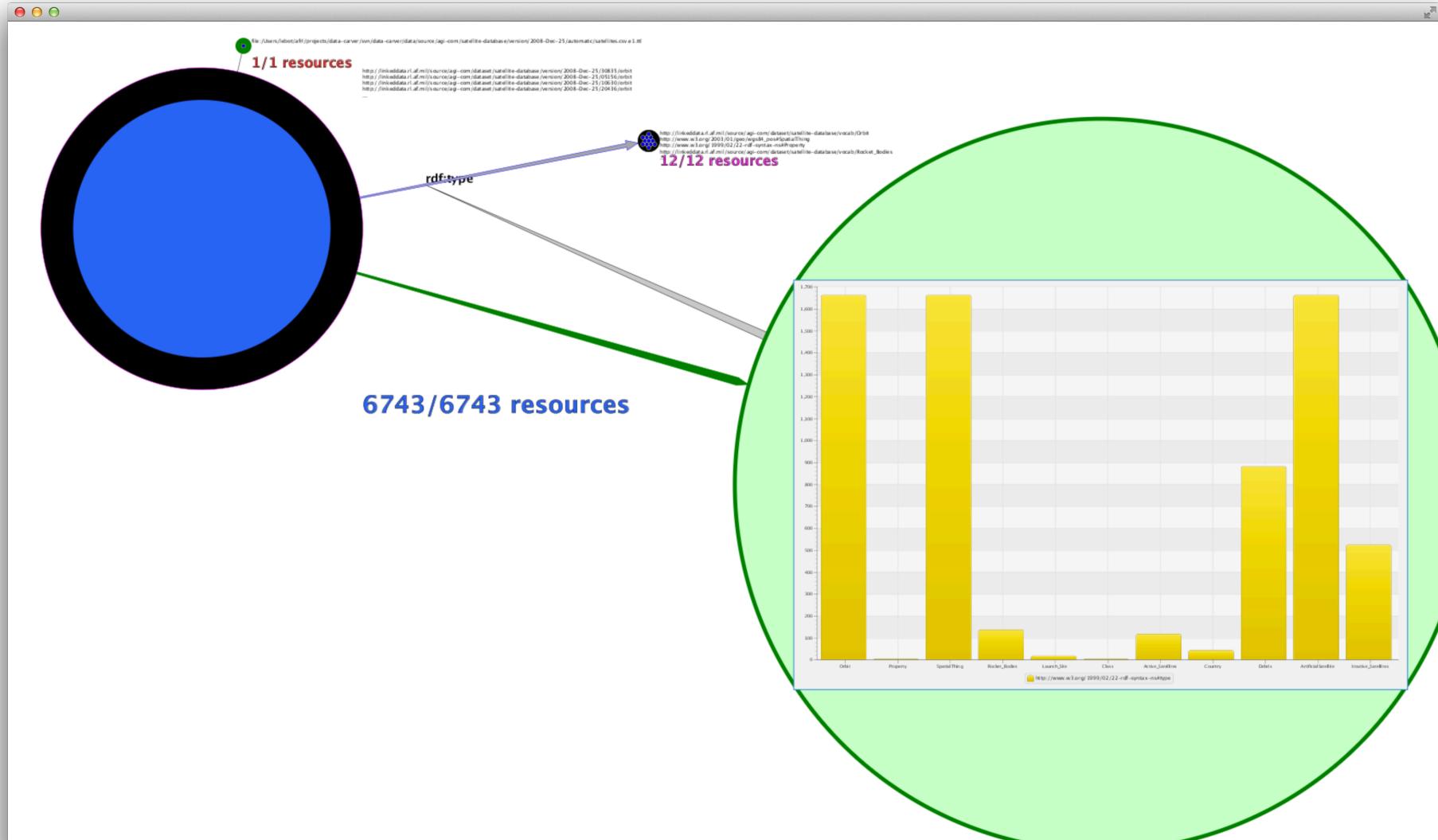
Abstract Visual Information

Derived Information

Existing Linked Data

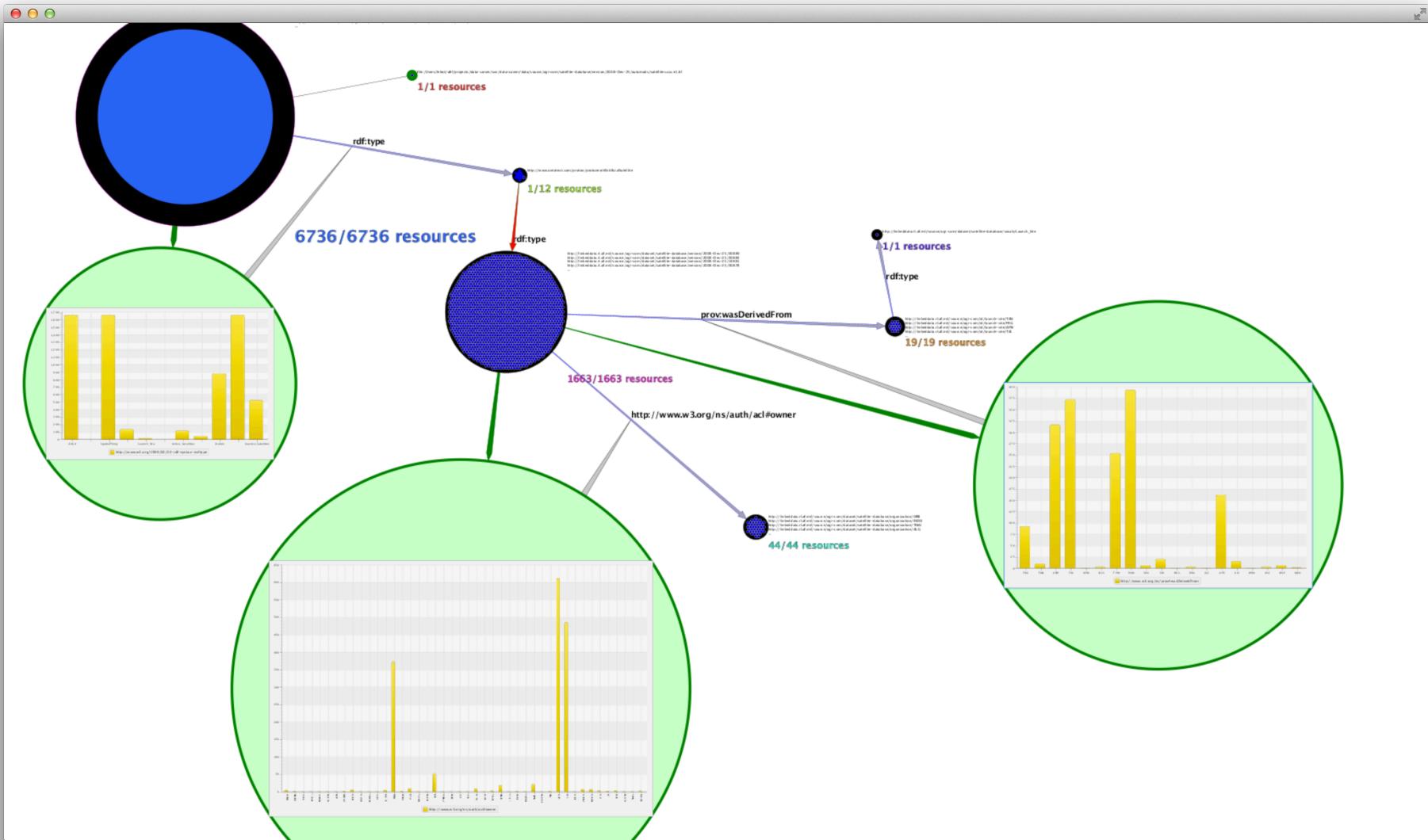


Demo!



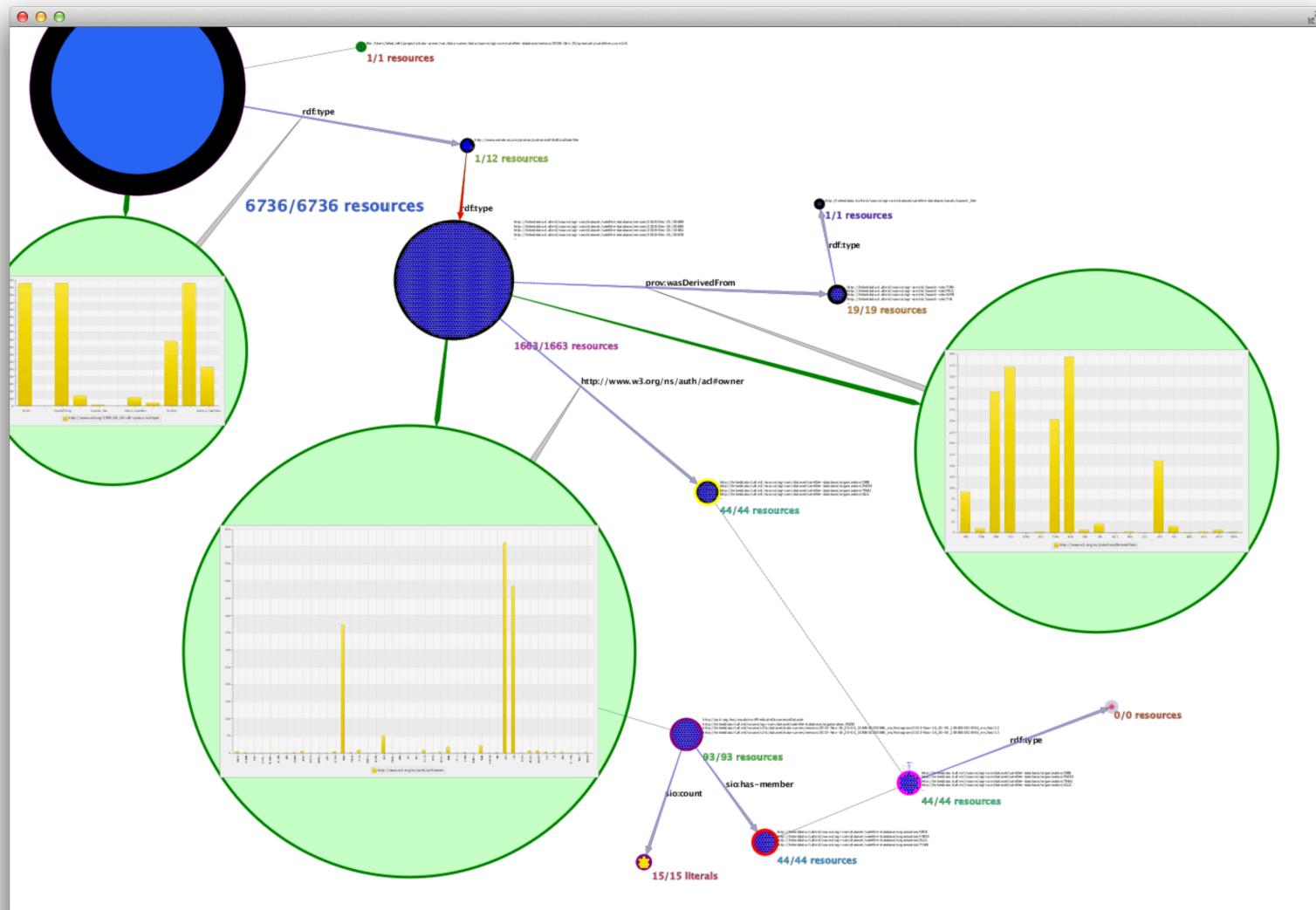


Demo!





Demo!





Conclusions

- **Objective:** Reducing the time required for analysts to *create, understand, trust, and re-purpose* any existing result.
- **Hypothesis:** Large factor is dearth of explicit, meaningful connections within and among analytical results is a significant factor for the challenges currently facing practitioners of data analytics.
- **Approach:** Maintains machine-understandable connections between the view of data and the underlying data itself.

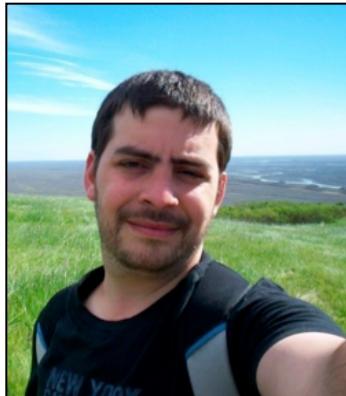


Future Work

- Evaluation!
 - Use cases from Visual Analytics community
 - Use cases from Linked Data community



Thanks!



- Tim Lebo
- Alvaro Graves



Deborah McGuinness