FUJITSU
shaping tomorrow with you

DERI

Empower Flexibility via Big Data: Unleashing Constraints in Financial Analytics

Pierre-Yves Vandenbussche – Fujitsu
pierre-yves.vandenbussche@ie.fujitsu.com

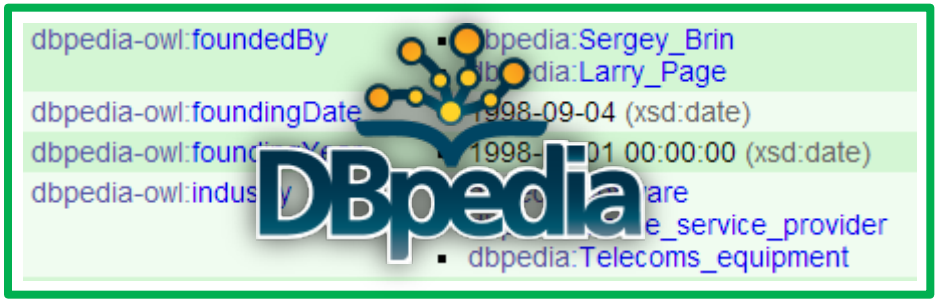
17 April 2013

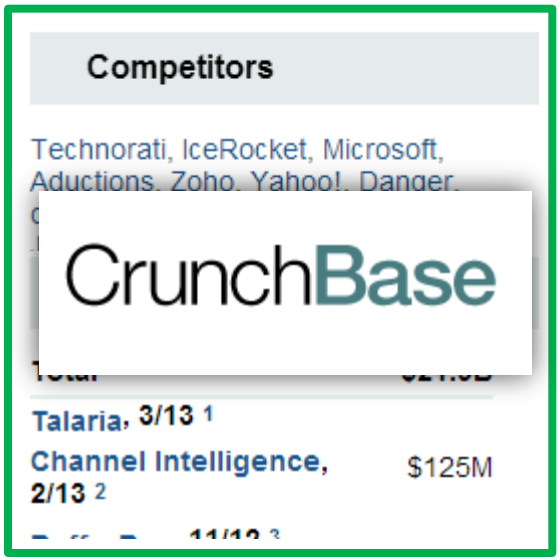
Project goals

- Take advantage of **Open Big Data** published on the Web
- Integrate Large Scale **Holistic data** to provide a more comprehensive information
- Perform **knowledge analytics** while preserving data provenance
- Enable **concept** (from diverse integrated sources) based **metric calculation**

Challenges

- **Interoperability** between Open Data
  - lack of inter data sources links
  - format heterogeneity (XML, schemas, API, WebSite crawling, CSV files, etc.)

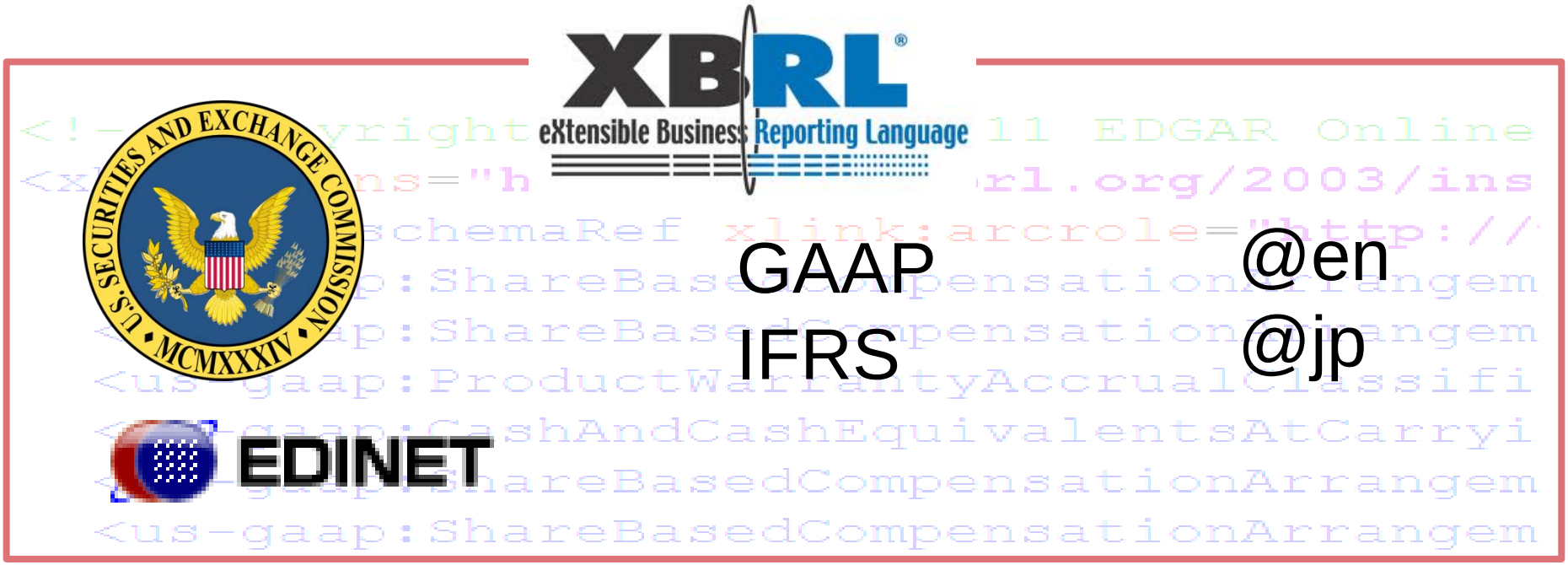




Challenges

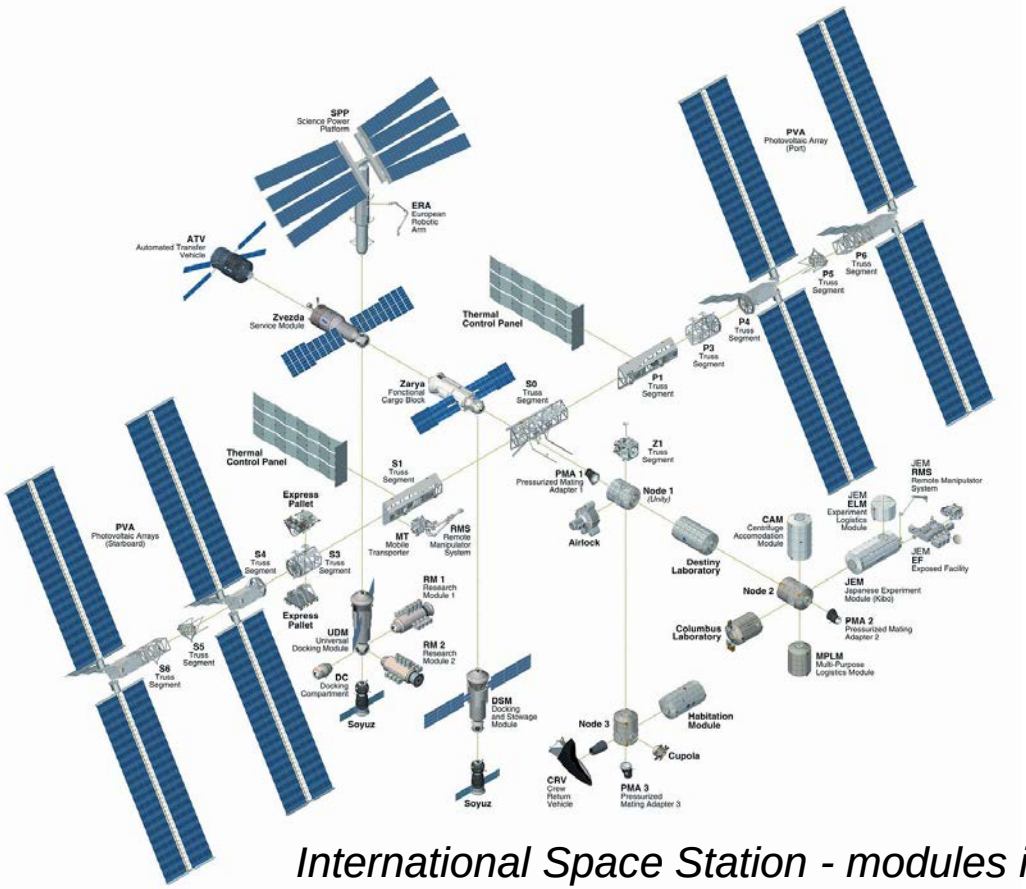
- Lack of a unique **worldwide identifier** for a company

- XBRL **Taxonomies**
  - updated by regulation in a frequent manner. Backward compatibility?
  - taxonomy alignment (XBRL concepts <=> Banking Regulations)



Challenges

- Integrate and represent all Data in a **standard model** to
  - process knowledge analysis algorithms over all data
  - allow concepts calculation regardless of the provenance (XBRL, New York Times, DBPedia, Stock Price, etc.)



International Space Station - modules integration

Solutions we propose

- Use Linked Data as
  - a technology to **bridge** Open Big Data set silos and facilitate **interoperability**
  - a **graph model** to represent all the data in a **unique format** and provide a **standard access**

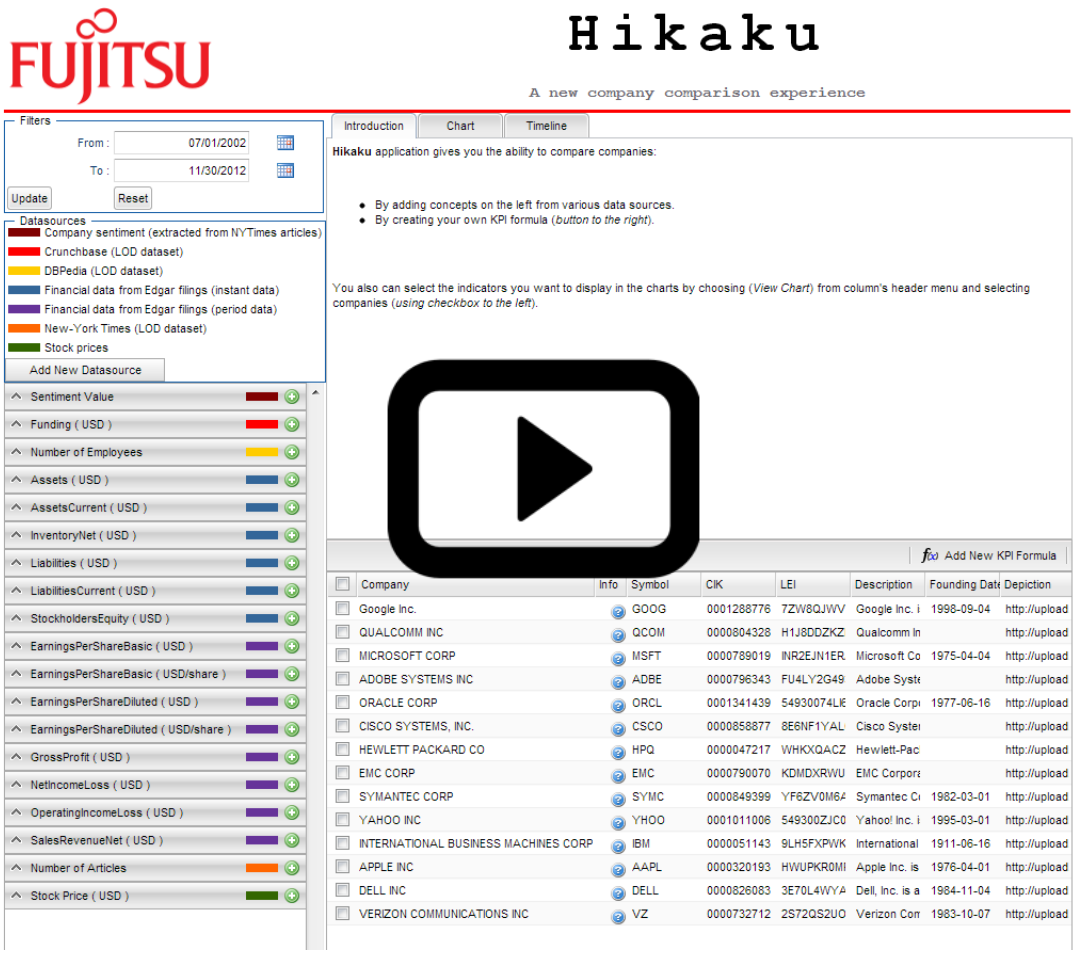
Existing Problem	Solution we offer
Entity identification and relations	Use of the **LEI** (legal Entity Identifier) to uniquely identify companies. Create some **mapping** between taxonomies
Disparate data formats	Represent all the data in a **unique format (holistic data integration)**
Knowledge analytics	Compute **company sentiment** based on New-York Times press articles

Solutions we propose

- Customisable interactive data

Existing Problem	Solution we offer
Combining concepts coming from disparate data sources	Develop a KPI/formula editor for a user to build its own comparison metric regardless of the data provenance (e.g. XBRL and DBPedia)
Few company selection criteria	Use XBRL data and Open Linked Data information to filter companies to compare

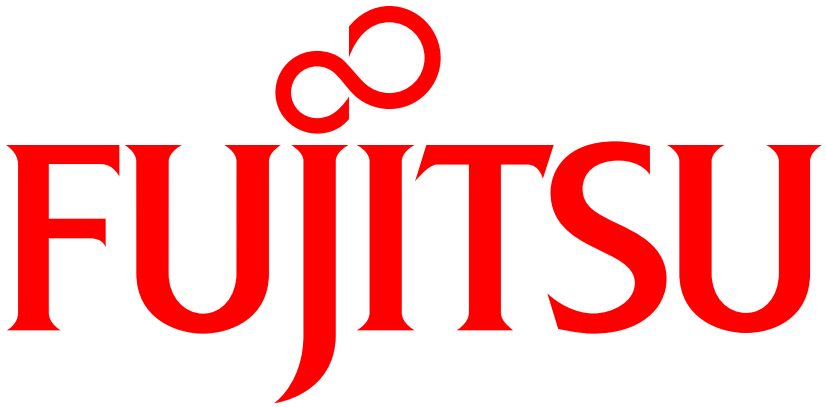
Demonstration



Conclusion and Future work

- Using Linked Data technology to tackle Open Data interoperability issue
  - Use this technology to encompass more data sources (EDINET, etc.)
- Experiment innovative Sentiment Value from text
  - Use this algorithm to detect facts from XBRL free text description and other Open data sources
- Enhance financial analysts experience by allowing KPI calculation using concepts regardless of their source
  - Enable dynamic addition of data sources

Thank you for your attention.
Also demonstrating at Fujitsu booth!



shaping tomorrow with you

Acknowledgement

Fujitsu Team: Goto Masatomo, Gofran Shukair, Vivian Qian Li, Jurgen Umbrich, Yutaka Mitsuishi, Aisha Naseer
DERI Team: Sean O'Riain, Richard Cyganiak, Vit Novacek

Components Architecture