Business Intelligence on non-conventional data

Enrico Gallinucci, PhD Student @ *DISI – University of Bologna (ITALY)*
enrico.gallinucci2 @ unibo.it

SOCIAL BI

SCOPE

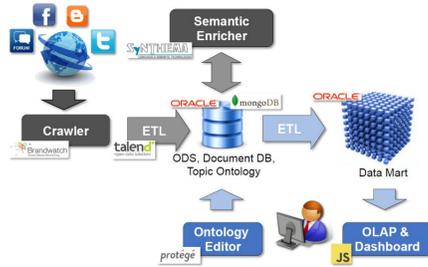**Analyse trends and moods** from the environment (social networks, blogs, forums, etc.) by combining semantically-enriched UGC with corporate data

RESEARCH ACTIVITY

A *hierarchy of topics* must be defined to perform effective analyses, but topics are

HETEROGENEOUS FUZZY EXPRESSIVE UNPREDICTABLE

META-STAR is an alternative to star-schema for modeling topic hierarchies, which combines meta-modeling, navigation tables, traditional DTs

- Irregular and schemaless hierarchies
- No impact on the intensional level
- OLAP queries' expressiveness extended
- Static enterprise dimensions integrated

RESULTS

- Full architecture deployed; formal definitions, execution plans and cost model provided
- Performances comparable with star schema
- SBI architecture and meta-stars have been successfully adopted on several projects

E. Gallinucci et al., Meta-stars: multidimensional modeling for Social Business Intelligence, DOLAP, 2013
E. Gallinucci et al., Advanced Topic Modeling for Social Business Intelligence, IS, 2015
E. Gallinucci et al., Meta-Stars: Dynamic, Schemaless, and Semantically-Rich Topic Hierarchies in Social BI, EDBT, 2015

BI on BIG DATA

SCOPE

Meet the requirements of data enthusiasts and enable the analysis of big data for users with no formal training in data science

RESEARCH ACTIVITY

Identify the homogeneous concepts or sub-schema that coexist within a NoSQL document database and extract a set of data according to relational patterns

- Flattening of every nested structure
- Application of clustering and frequent-itemset mining techniques to find frequent sub-schemata
- Application of conditional functional dependencies to further differentiate the derived schemata

PERVASIVE BI

SCOPE
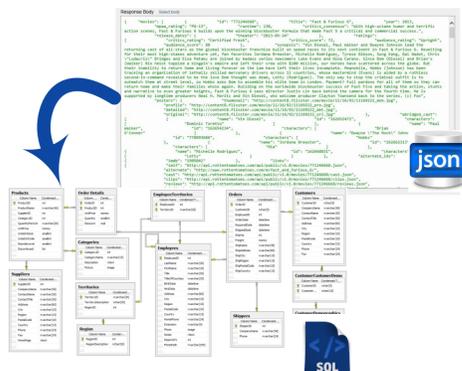
An innovative and forward-looking perspective for BI, characterized by timeliness (real-time queries), integration (BI anywhere, anytime), personalization

RESEARCH ACTIVITY

Traditional recommendation techniques are mainly focused on single OLAP queries

Proposal: a collaborative filtering approach focused on OLAP sessions

- ALIGNMENT: get the most similar OLAP session and align them to the current one
- RANKING: extract the common sub-sessions, rated on similarity and frequency in the log
- FITTING: adapt the top-ranked sub-sessions to the current one

RESULTS
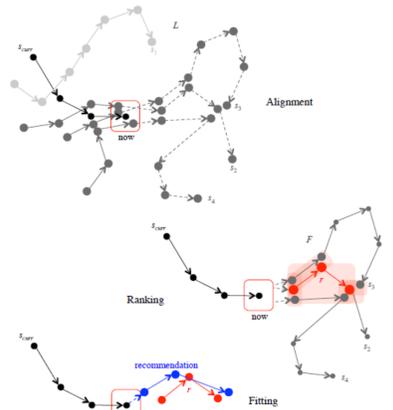
Recommendations meet many desirable properties
- NOVELTY (with respect to the log)
- RELEVANCE (frequency of execution)
- FORESIGHT (prevision of the intended direction)

Tests have been performed using CubeLoad, a parametric generator of realistic OLAP workloads

J. Aligon et al., A Collaborative Filtering Approach for Recommending OLAP Sessions, DSS, 2014
S. Rizzi et al., CubeLoad: a parametric generator of realistic OLAP workloads, CAiSE, 2014

SELF-SERVICE BI

SCOPE

Fusion Cubes extend local DW by integrating knowledge from external sources (linked data, open data, etc.) on a self-service basis

RESEARCH ACTIVITY
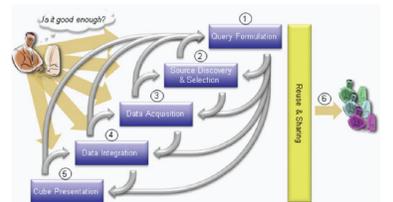
iMOLD (interactive multidimensional modeling of linked data) is an approach to semi-automatically derive multidimensional schemata from linked data

- Detection of roll-up relationships and levels on public ontologies by searching MD patterns
- Acquisition of knowledge and reusage by every user
- Adoption of QB4OLAP and SM4AM to model the internal knowledge
- Future work aimed at facts detection and deployment of schema and data

RESULTS
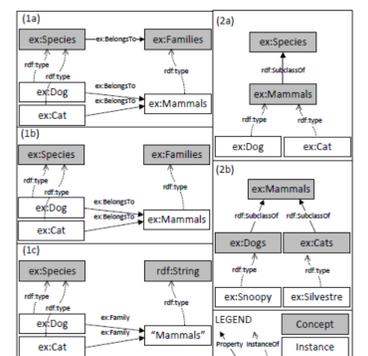
- Algorithms and UI implemented and tested
- Particularly useful in cross-domain ontologies, where relationships are copious and hard to grasp