Modeling Analytical Streams for Social Business Intelligence

Indira Lanza-Cruz 1, Rafael Berlanga 1 and María José Aramburu 2

1 Dep. de Llenguatges i Sistemes Informàtics, Universitat Jaume I, Spain; lanza@uji.es; berlanga@uji.es

2 Dep. de'Enginyeria i Ciència dels Computadors, Universitat Jaume I, Spain; aramburu@uji.es

Motivation

Traditional Business Intelligence (BI): Decision making processes are limited to controlling internal business variables.

Social Business Intelligence (SBI) has to face the high dynamicity of both the social network contents and the company analytical requests, as well as the enormous amount of noisy data.

Objetives

- Provide a framework for the definition and processing of social indicators.
- Propose a streaming architecture specially aimed at SBI.
- Propose a new method to model analytical streams for SBI.

Background

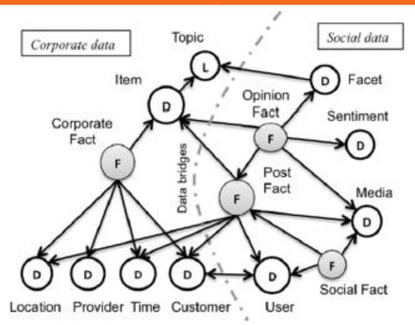

Fig. 1. Analytical Patterns for Social BI.

Modeling Social Indicators

Example Use Case

A Rental Car Company wants to study the popularity of different car brands by tracking the "user's Engagement" in the last period. It may need to identify:

- non-spam users.
- relevant domain users.
- User profiling.

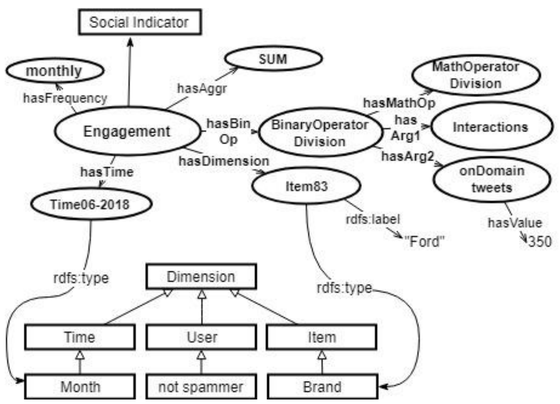

Fig. 2. Modeling Engagement Indicator.

Methodology

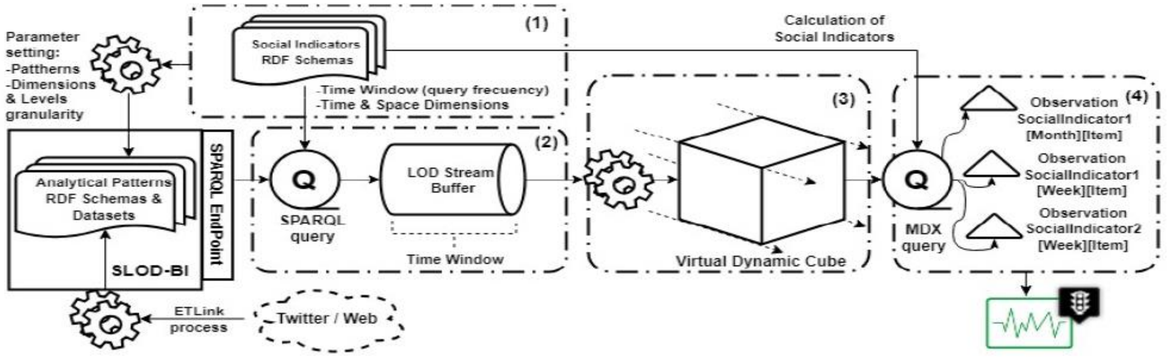

Fig. 3. Framework.

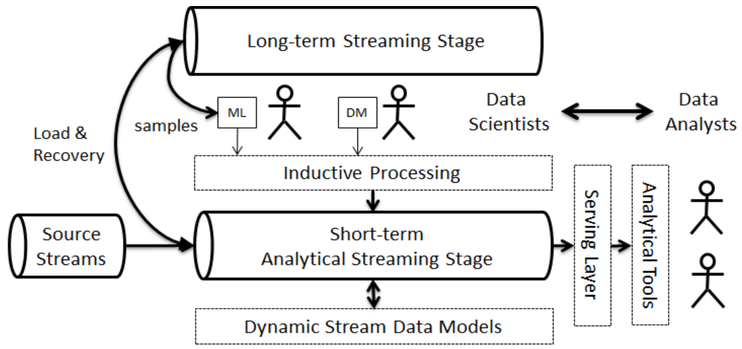

Fig. 4. Architecture for Dynamic Social Business Intelligence.

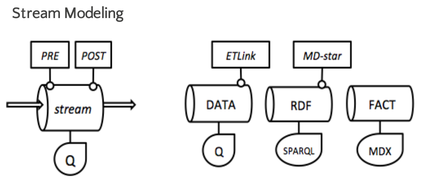

Fig. 5. Stream graphical models

Experimental Study

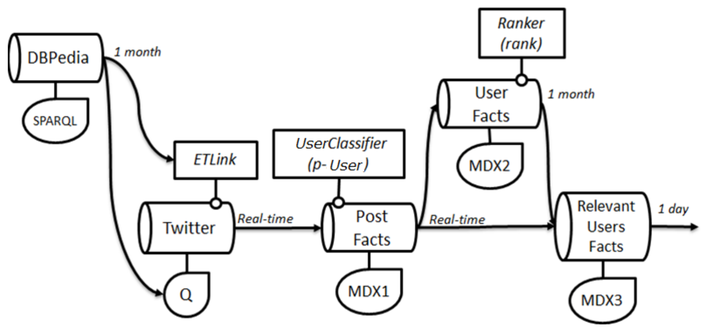

Fig. 6. Example workflow

Prototype :

- Implemented with Python.
- Source streams from Twitter API
- Classifier using Anaconda framework (Pandas and ScikitLearn packages)

Cars Dataset 2,625.186 tweets 11/2014 - 02/2017

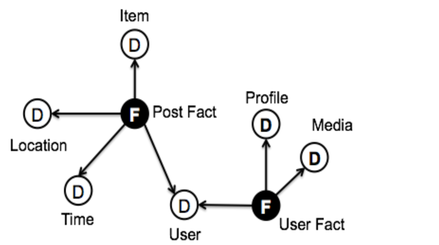

Fig. 7.Graphs obtained after applying ETLink.

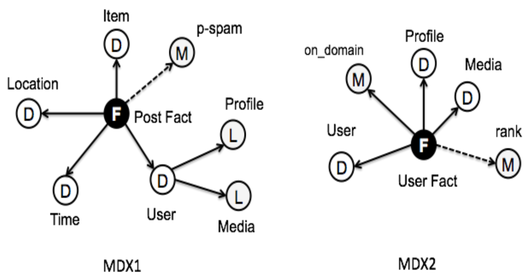

Fig. 8. Output schemas after applying MDX1 and MDX2.

Results

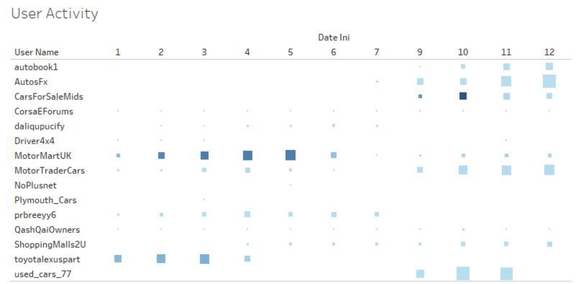

Fig. 9. Relevant Users per activity.

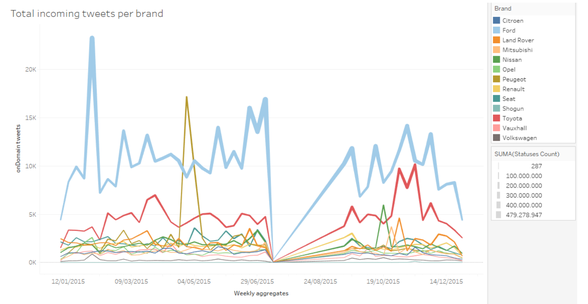

Fig. 10. Total incoming post facts

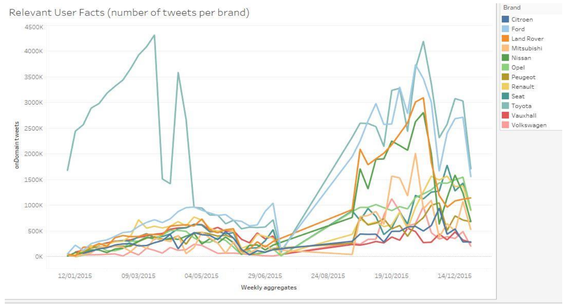

Fig. 11. Post facts per brand for top 15 relevant users.

Conclusions

- A novel framework for definition and monitoring Social Indicators.
- We propose a new architecture that aims at covering all these requirements as well as at integrating Data Science and Data Analysis tasks in the same working area.
- We adopt a Kappa-like streaming architecture to cover the requirements of both kinds of actors.
- The architecture relies on both Linked Data and multidimensional modeling. The adoption of semantics also facilitates the validation and follow-up of the developed workflows.

Future work

- Automatic creation of descriptions and queries associated with social indicators.
- Full implementation of the architecture in Python and its integration with a highly scalable streaming framework like Kafka.
- Author profiling methodology for SBI.

- Lanza-Cruz, Indira; Berlanga, Rafael.; Aramburu, María José. Modeling Analytical Streams for Social Business Intelligence. *Informatics* **2018**, *5*, 33.
- Lanza-Cruz,Indira; Berlanga, Rafael. Defining Dynamic Indicators for Social Network Analysis: A Case Study in the Automotive Domain using Twitter. 10th International Conference on Knowledge Discovery and Information Retrieval 2018.

Acknowledgements: Supervised by Professor Rafael Berlanga. This work has been financed by the Ministry of Economy and Trade with the project of the National R&D Plan with contract number TIN2017-88805-R.We also have the support of the Universitat Jaume I pre-doctoral scholarship programme (PREDOC/2017/28).