Cloud Intelligence – Challenges for Research and Industry (Panel)

Jérôme Darmont
Université de Lyon (Laboratoire ERIC)
France
jerome.darmont@univ-lyon2.fr

Torben Bach Pedersen
Aalborg University (daisy)
Denmark
tbp@cs.aau.dk

ABSTRACT

This panel paper aims at initiating discussion at the second International Workshop on Cloud Intelligence (Cloud-I 2013), about current research and industrial challenges in cloud business intelligence as a service.

Categories and Subject Descriptors

H. [**Information Systems**]

General Terms

Design, Economics, Experimentation, Human Factors, Languages, Legal Aspects, Management, Performance, Reliability, Security

Keywords

Elasticity, Data sources, User-centered, Pay-as-you-go, Collaborative, Privacy, Availability, Data analytics

1. FUSION CUBES – A FOUNDATION FOR COLLABORATIVE CLOUD INTELLIGENCE

Cloud Intelligence is characterized by being collaborative in nature. Many people, knowing each other or not, share data and results, and thus contribute, consciously or not, to achieve a goal. Additionally, the data "belonging" to the user with a certain (open-ended question) is typically not enough to answer the question. Thus, new data, typically available on the Internet, must be discovered, selected, acquired, integrated, and presented. However, current BI/analytics tools do not support this process, as they are designed for a closed-world setting, where data and questions are more or less known in advance.

The recent proposal of *Fusion Cubes* [1] aims to support exactly such scenarios. The concepts and ideas behind fusion cubes will be presented and it will be discussed how fusion cubes provide an ideal technical foundation for collaborative Cloud Intelligence.

2. PERSONAL INTELLIGENCE

Personal Intelligence aims to be a cloud decision-support service (in SaaS mode) that is accessible to individuals and very small companies, which implies constraints on usage easiness and low cost. Such an on-line application must also allow cross-analyzing private data (under privacy constraint) and external data (e.g., open data or data harvested on the Web), collaboratively; as well as sharing analysis results. To the best of our knowledge, few software tools currently explore this trend, though we can cite Google Fusion Tables [2] and the `illo.com` website, which propose a spreadsheet-like interface allowing to import data into the cloud and produce reports and graphs.

The idea behind Personal Intelligence is to push this concept one step further by following the design process of a datamart, from data integration with a simple Extract, Transform, Load (ETL) tool to On-Line Analysis Processing (OLAP) via a user-friendly interface, including the automatic multidimensional modeling of the datamart [3]. This approach is similar to that of Fusion Cubes (Section 1), but is less ambitious because it targets only small-scale decision-support applications. Data privacy and performance issues, which are fundamental in a Personal Intelligence context, will be discussed in this panel.

3. PANELIST BIOS

Prof. Dr. Jérôme Darmont is full professor of computer science at the University of Lyon 2, France, and the director of the ERIC laboratory. He received his Ph.D. in 1999 from the University of Clermont-Ferrand II, France, and then joined the University of Lyon 2 as an associate professor. He became full professor in 2008. His research interests mainly relate to database and data warehouse performance (performance optimization, auto-administration, benchmarking...) and cloud Business Intelligence (data security, query performance and cost, personal BI...). He has published more than 100 peer-reviewed papers. He is a member of several editorial boards and has served as a reviewer for numerous conferences and journals. Along with Torben Bach Pedersen, he initiated the Cloud Intelligence workshop series in 2012.

Prof. Dr. Torben Bach Pedersen is full professor of computer science at Aalborg University, Denmark. He

received his Ph.D. in 2000. His research interests span Business Intelligence topics such as data warehousing, OLAP, and data mining, with a focus on non-traditional and complex types of data, in large volumes – aka Big Data. He has published more than 130 peer-reviewed papers on these topics. He served as PC Chair for DaWaK 2009+10, DOLAP 2007, EnDm 2012+13, General Chair for SSTD 2009, and on numerous program committees, including SIGMOD, VLDB, ICDE, and EDBT.

He will serve as PC Chair for SSDBM in 2014. He has worked on Cloud Intelligence for several years and, along with Jérôme Darmont, he initiated the Cloud Intelligence workshop series in 2012.

4. REFERENCES

[1] A. Abello, J. Darmont, L. Etcheverry, M. Golfarelli, J.-N. Mazon, F. Naumann, T.-B. Pedersen, S. Rizzi, J. Trujillo, P. Vassiliadis, and G. Vossen. Fusion Cubes: Towards Self-Service Business Intelligence. *International Journal of Data Warehousing and Mining*, 9(2):66–88, April-June 2013.

[2] H. Gonzalez, A. Y. Halevy, C. S. Jensen, A. Langen, J. Madhavan, R. Shapley, W. Shen, and J. Goldberg-Kidon. Google Fusion Tables: Web-centered data management and collaboration. In *ACM SIGMOD International Conference on Management of Data (SIGMOD 2010), Indianapolis, USA*, pages 1061–1066, 2010.

[3] C. Phipps and K. C. Davis. Automating data warehouse conceptual schema design and evaluation. In *4th International Workshop on Design and Management of Data Warehouses (DMDW 2002), Toronto, Canada*, volume 58 of *CEUR Workshop Proceedings*, pages 23–32, 2002.