Would Cloud Computing Revolutionize Teaching Business Intelligence Courses?

Stevan Mrdalj
Eastern Michigan University, Ypsilanti, MI, USA

smrdalj@emich.edu

Abstract

In the last decade Business Intelligence has become one of the hottest areas in Information Technology. Meanwhile, teaching Business Intelligence courses for Masters of Business Administration programs faces a variety of challenges. One of the major hurdles in offering Business Intelligence courses is the complex and costly computing infrastructure necessary for the software used in such courses. One possible remedy might be to use cloud computing. Cloud computing offers flexible access the necessary software and hardware through the Internet. It is becoming increasingly popular because of its on-demand and pay-per-use business models. In this paper, we examine how cloud computing might be used to revolutionize the way of teaching business intelligence courses. In doing so, we present various possible infrastructure scenarios and the necessary tools. Finally, we discuss the perceived difficulties in adopting Business Intelligence as a service business model in academic settings. We also reflect on the economics of cloud computing.

Keywords: Cloud Computing, Business Intelligence, Software as a Service, Teaching BI, Pay-per-use.

Introduction

For the last several years, Masters of Business Administration (MBA) programs have been under mounting pressure to provide courses to prepare their students to become the next generation of Business Intelligence (BI) workers. In a recent survey of professors from more than 80 universities around the world (Betts & Kanaracus, 2010), Barbara Wixom discovered that 43% of the respondents reported that they can't adequately teach BI courses" because they lack access to the needed software, hardware and real-world business problems". She also indicates that "It's clear from the survey that instructors want to change the way students learn about business intelligence. Professors want to provide large data sets, contemporary software tools, and real-world content within their classrooms. But, factors like high technology costs, complex maintenance requirements, and steep learning curves present insurmountable obstacles." With the present economic crisis and its repercussions on the academic environments, one of the major hurdles in offering Business Intelligence courses is the complex and costly computing infrastructure necessary for the software used in such courses. On the other hand, unless Business Intelligence becomes a mainstream field at business schools, the shortage of business intelligence related professions will continue to rise.

The cloud computing, also known as software-as-a-service, delivery model has the potential to become a major

force in the future of offering BI courses. After a decade of growth and adoption, the commercial cloud computing offerings came into existence and products like Amazon EC2, Microsoft Azure and Salesforce.com helped to popularize it. The concept of on-demand BI promises to make it affordable for business schools to deliver quality BI instruction to MBA students.

Computer Science programs were among the first to develop experimental undergraduate data mining courses as reported by Lu and Bettine (2003). Banks, Dong, Liu and Mandvikar (2004) report on teaching data mining courses as "an exciting addition to the curriculum at the senior/graduate level" that provide the opportunity to apply computer related education to various domains and applications. Paper by Roiger (2005) presents a very good tutorial on data mining that can be used by educators as an introduction to business intelligence. The typical structure of a data warehousing and data mining course in the Computer Science curriculum is well described by Fang and Tuladhar (2006) and Musicant (2006). Information systems programs at business schools typically offer BI courses as elective courses (Watson, 2006). A good example of an applied BI course that is designed to appeal to the graduate students of the CIS programs as well as MBA students selecting the CIS concentration is presented by Mrdalj (2007). Specifically, the structure and components of this course focus on "business aspects" of data mining and data warehousing by having students learn how to answer today's BI questions. In one of the rare papers related to MBA programs, Mrdalj and Diallo (2010) present the structure and components of a BI module that is designed to appeal to students of the finance program as well as to MBA students selecting the finance concentration.

In the next section of this paper we stage our proposed solution with a brief presentation of the typical computing infrastructure architecture needed for teaching BI courses with a discussion of possible on-premises installations. It is followed by the description of cloud computing and its possibilities in teaching BI courses. We conclude with perceived difficulties in using cloud computing as a platform for teaching BI courses.

Computing Requirements for BI Courses

The typical computing infrastructure architecture needed for teaching BI courses is shown in Figure 1. It requires a database server to host source databases as well as the data warehouse.
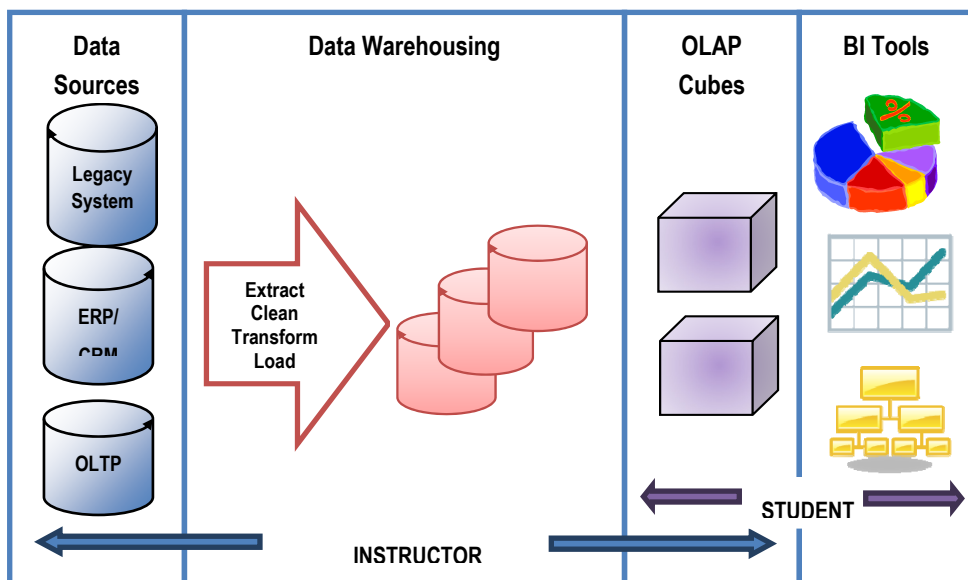

Figure 1: A Typical BI Architecture

Next, we need data warehousing development tools to extract data from the sources and to clean, transform and load data into the data warehouse. Although this part is not visible to the MBA students, it is required for the instructors in order to provide the students with "real-life" data sets so they can make sure that students learn from actual company data instead of textbook examples.

In addition, BI architecture requires technology such as Online Analytical Processing (OLAP) that provides a platform for end-user based utilization of data warehouse (Larson, 2009) and is a dominant paradigm in BI. This technology is essential for MBA students who are usually not familiar with SQL. OLAP tools give students flexible and fast access to multi-dimensional views (such as store, account, time, geography) and to measures (such as sales, payments, discounts), both summarized and detailed. Such access if further supported with standard OLAP operations that include drill-down, roll-up, dice and slice. Another benefit of using OLAP Cubes is that students can access them with tools such as the Microsoft Excel PivotTable (Harts, 2008; Mrdalj & Diallo, 2010) as illustrated in Figure 2.
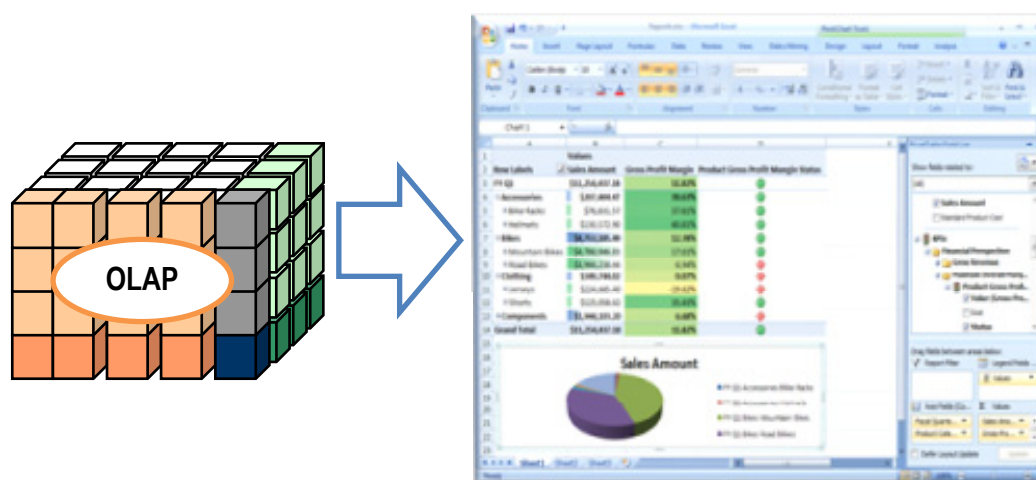
Figure 2: End-user BI Access

Tools like the Microsoft SQL Server Data Mining Add-ins (MacLennan, Tang & Crivat, 2009) and SAS® Enterprise Miner (Cerrito, 2006) enable Microsoft Excel to become the analytical client to the OLAP analytical servers. This enables MBA students to access multi-dimensional data sources and to perform reporting, ad-hoc analysis and data mining within the very familiar environment of Microsoft Excel (Mrdalj & Diallo, 2010). Consequently, Microsoft Excel, perhaps the oldest BI tool, has regained its role in MBA education as an affordable BI front end tool.

The above described BI architecture can be implemented on-premises in several different ways. We will discuss the following three most common approaches.

Stand Alone Infrastructure

The simplest and most restrictive is the stand-alone approach when the entire BI architecture (database server, analytical server and BI tools) is installed on each individual workstation in the lab or computer classroom. It requires early setup of all the software and data sets which can be easily cloned to all workstations. To accommodate the continuity in students' work, such setup would require that students use the very same workstation throughout the entire semester and to keep their own backups. Each change in software and data sets during a semester would require a manual installation of such changes on all workstations since re-cloning would diminish the previous student's work. This setup would also require that students would need to do their assign-

ments and projects on very same workstations that they use during classes or labs. Such a requirement poses serious constraints on the type and quantity of assignments that can be given to students. Otherwise, students would need to have the complete BI architecture installed on their laptops/desktops.

Setting up a proper BI environment on students' laptops is a multi-step process (for illustration purposes we will describe a scenario with Microsoft SQL Server). First, it requires installing the SQL Server. The next step is to install sample data warehouses including the corresponding OLAP Cubes. Unfortunately, deployment of OLAP Cubes requires some familiarity with the Microsoft SQL Server Management Studio or the Microsoft Business Intelligence Development Studio project. The next step is to download and install Microsoft SQL Server Data Mining Add-ins. The last task of this process is to connect Excel to the SQL Server Analysis Service and to configure an analysis database to be used by Excel to perform data mining tasks. The above described multi-step installation process would certainly provide a major challenge to MBA students as well as the necessity to acquire all the needed software. Based on our experience in using this approach, we consider this to be prohibitive for the majority of MBA students and, therefore, we consider it a most restrictive approach for teaching BI courses.

Local Area Network Infrastructure

The step-up approach would have a local area network where the database server hosting the data warehouse and the analysis server hosting the OLAP cubes are installed on the network server and all the client workstations would need only Excel with BI tool add-ins as shown in Figure 3.
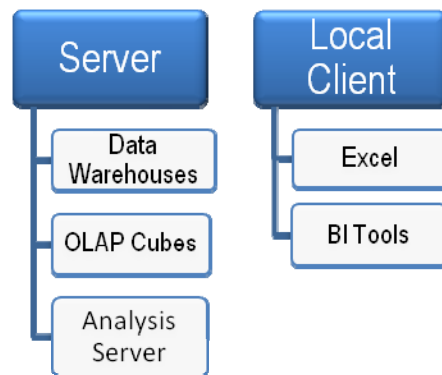

Figure 3: Network Configuration

This approach certainly allows any mid-semester changes in the software and data sets. It requires moderate DB administration since all clients would need proper authentication to access OLAP cubes. The major problem with this approach is that those students wanting to do their assignment on their laptops would need to be permanently connected to the network since they would need access to the analysis server to perform BI tasks. Students who would like to do their work outside of the lab environment would need to have the stand-alone installation described above.

Web Enabled Infrastructure

To overcome the problems with the previous two approaches, business schools would need to invest in the web enabled solution depicted in Figure 4. The advantage of this approach is that students would only need to install Excel add-ins and to have available Internet access. On the other hand, business schools would need considerable hardware and software resources to im-

plement it. In addition, this approach would require a high level of expertise to configure database and web servers as well as to make them operational non-stop. The cost of designing, building, installing and running the web enabled infrastructure continues to be the most noticeable limiting factor for business schools. Frequently such prohibitive costs have been the main limitation for proliferation of BI courses at MBA programs.
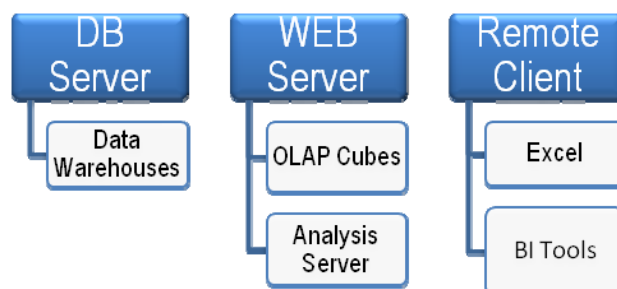
Figure 4: Web Enabled Configuration

Another variation of this approach is to use remote access tools for the remote clients. Such an approach would even eliminate the need for students to have Excel as well as BI tools installed on their laptops. A comprehensive discussion of the remote access tools (Scher, 2010) is beyond the scope of the paper, but we list some of them as examples: GoToMyPC, LogMeIn Pro2, Microsoft Remote Desktop Connection, Remote PC and pcAnywhere.

If the universities do not have adequate resources or if they do not want to acquire and manage technical resources or if they lack the necessary technical experience and skills, then they should consider using cloud computing.

Using Cloud Computing for Teaching Business Intelligence

Cloud computing is one of the most attractive technology areas today due to its cost efficiency and flexibility. There are many definitions and interpretations of cloud computing, but for the purpose of this paper we will use the following definition created by the National Institute of Standards (Mell & Grance, 2009):

"Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

In the cloud computing paradigm, the basic architecture needed for teaching BI courses based on the three layered NIST service model (Mell & Grance, 2009) and its customizations for BI (Chandra & Iyer, 2010; Reyes, 2010) is presented in Figure 5. At the lowest level, the processing power refers to the hardware needed for running database servers and networking. Normally, the customer does not control or manage this layer of the cloud infrastructure. The cloud's Platform as a Service (PaaS) provides instructors with the capability to deploy large data warehouses from real businesses. The Software as a Service (SaaS) layer of the cloud provides students and instructors with applications running on a cloud infrastructure. These layers may be considered as services to the layer above them.

In relationship to the BI courses, the data warehousing tools would allow instructors to deploy data warehouses and OLAP cubes as well as to maintain them over the course of the semester. Students would gain access to the OLAP cubes and reporting, data mining, and business performance management applications. The OLAP cubes and applications are accessible from remote clients' devices trough a thin client interface such as a web browser (Mell & Grance, 2009) or Excel add-ins (Predixion, 2010). In short, cloud computing shifts the physical construction of the data center, creating network connections between students and the center, and providing all the necessary software to the trusted service provider (Thomson, 2009).

Figure 5: Basic BI Architecture in the Cloud

Thanks to recent advances in technology in recent years, the commercial cloud computing offerings may enable business schools to implement BI courses in a fraction of the time and avoid the capital expenditure required by traditional installations (Thomson, 2009). Frequently mentioned advantages of cloud computing are lower cost and non-stop operation. Even more important features for the academic environments are summarized below (adapted from Katzan, 2009):

- The responsibility for hardware, application software, storage facilities, and professional services resides with the provider.
- Systems software is available from a trusted vendor for supporting cloud services.
- Data centers are available for sustainable and reliable operation.

In other words, cloud computing would provide virtual and scalable infrastructure as a service where the complexity of resource management is hidden from the client while enabling the client to exploit supercomputing power on-demand without investing in huge infrastructure and management costs (Ananthanarayanan et al., 2009). A broad discussion of the available Cloud BI tools is beyond the scope of this paper and it can be found in (Rayes, 2010; Tsai, 2009). Instead, we list some of them as examples: Predixion Insight, CloudOLAP, IBM Smart Analytics Cloud, Informatica 9, and Rockspace Cloud.

This brings the "thin client" architecture back to life where both instructors and students would utilize the needed services from the cloud provided as shown in Figure 6. The biggest advantage that cloud computing brings to instructors is that they can obtain the proof-of-concept for the development of their courses in a matter of days. Instructors would be able to upload their example

data warehouses and OLAP Cubes and make them available to students wherever Internet access is available. Students would gain easy and location independent access without the need to acquire complex and expensive software that needs to be installed on their workstations in order to do their assignments.
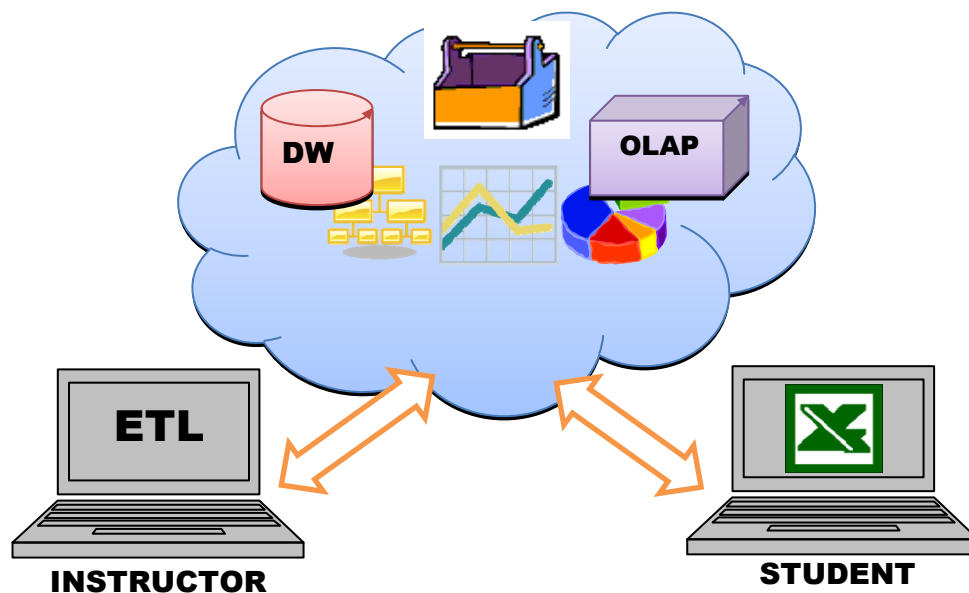

Figure 6: Cloud BI

Perceived Difficulties in Using Cloud Computing

Despite many advantages and benefits that cloud computing brings to business schools there are several perceived difficulties as well. In this paper, we will only concentrate on those that are relevant to using cloud computing for teaching purposes. Many other challenges related to cloud computing used for commercial purposes such as scalability, security, large data volume, speed of access, and reliability can be found in (Reyes, 2010; Thomson, 2009).

Customization is one of the very important aspects of cloud computing used for teaching BI courses. With non-configurable SaaS (Katzan, 2009), the cloud provider delivers a unique set of applications to clients and limits instructors to tailor them to specific classwork needs. A similar concern is for instructors teaching multiple sections when a multi-tenant SaaS would be required to host common applications and unique data sets for different sections.

Another important consideration of adapting cloud computing for teaching BI courses is that its pricing structure is variable and complex. There are three categories that we might observe as valid options for the academic environments (adapted from Katzan, 2009):

- *Perpetual license* refers to an "up front" payment for service, and unlimited access for an unlimited time.
- *Subscription*, as a form of cloud service monetization, can be conceptualized as a time-based perpetual license, often applied to multiple users.
- *Transaction based or variable pricing model,* also known as pay-as-you-go plan, is based on the data transfer volume and number of transactions.

The variable pricing model may look quite attractive for business schools since data volume and number of transactions may not be large, but it would bring an uncertain cost associated with a given course. Many would argue that the subscription model is more economical and that it pro-

vides budgetary predictability and a lower cost over a longer period of time (Thomson, 2009). What may drive up the cost for either pricing model are configurability and isolation. From a pricing stand point non-configurable and single-tenant SaaS have a lower cost, but are quite prohibitive in customization.

Both pricing models are something that business schools may not be accustomed to have associated with offering courses. Traditionally, universities have budgets for computer refresh programs and laboratory maintenance and not for cloud subscription fees associated with an individual course. Although either of the pricing models mentioned above might cost much less compared to the costs of on-premises BI architectures, it would require a paradigm shift in university budgeting models.

Conclusions

We examined the requirements for implementing the BI architecture necessary for teaching BI courses. Based on them, we concluded that cloud computing is an attractive solution for business schools wanting to implement cost-effective, rapid and dynamic environments for their BI courses. The promise of cloud computing is already proven at many companies and it might be prudent for business schools to adopt this new approach instead of investing in ever more demanding and expensive computer labs and classrooms.

By outsourcing the physical construction of the data center, creating network connections between students and the center, and providing all the necessary software to students, business schools can concentrate on rapid development of BI courses and their integration into MBA programs. With cloud computing, business schools may significantly reduce their costs while gaining guaranteed performance and high student satisfaction.

We also hope that cloud-based BI service providers like Predixion, CloudOLAP and especially Microsoft Azure (with the offering of the analysis server in the near future) would be willing to provide academic discounts for their services to those universities teaching BI courses.

References

Ananthanarayanan, R., Gupta, K., Pandey, P., Pucha, H., Sarkar, H., Shah, M., & Tewari, R. (2009). Cloud analytics: Do we really need to reinvent the storage stack? *Proceedings of the 2009 Conference on Hot Topics in Cloud Computing*. Retrieved September 12, 2010 from http://www.usenix.org/event/hotcloud09/tech/full_papers/ananthanarayanan.pdf

Banks, D. L., Dong, G., Liu, H., & Mandvikar, A. (2004). Teaching undergraduates data mining engineering programs. *Proceedings of the 34th ASEE/IEEE Frontiers in Education Conference*, October 20 – 23, 2004, Savannah, GA, T1A-1 - T1A-6. Retrieved September 22, 2010 from http://www.public.asu.edu/~huanliu/papers/fie04.pdf

Betts, M., & Kanaracus, C., (2010, April 5). Professors cite challenges in teaching BI in college. *Computerworld*. Retrieved November 10, 2010 from http://www.computerworld.com/s/article/347880/Professors_Struggling_to_Teach_BI

Chandra, B., & Iyer, M. (2010). BI in a cloud: Defining the architecture for quick wins. *SETLab Briefings*, *8*(1), 39-44. Retrieved October 10, 2010 from http://www.infosys.com/research/publications/Documents/BI-in-a-cloud.pdf

Cerrito, P. B. (2006). *Introduction to data mining using SAS Enterprise Miner*. Cary, NC: SAS Institute.

Fang, R., & Tuladhar, S. (2006). Teaching data warehousing and data mining in a graduate program of information technology. *Journal of Computing Sciences in Colleges, 21*(5), 137-144.

Harts, D. (2008). *Microsoft Office 2007 business intelligence: Reporting, analysis, and measurement from the desktop*. McGraw-Hill/Osborne.

Katzan, H. (2009). Cloud software service: Concepts, technology, economics. *Service Science, 1*(4), 256-269.

Larson, B. (2009). *Delivering business intelligence with Microsoft SQL Server 2008*. McGraw-Hill/Osborne.

Lu, Y., & Bettine, J, (2003). Data mining: An experimental undergraduate course. *Journal of Computing Sciences in Colleges, 18*(3), 81-86.

MacLennan, J., Tang, Z. H. & Crivat, B. (2009). *Data mining with SQL Server 2008*. Wiley Publishing.

Mell, P., & Grance, T. (2009). *The NIST definition of cloud computing*. National Institute of Standards and Technology, Version 15, 10-7-09. Retrieved October 10, 2010 from http://csrc.nist.gov/groups/SNS/cloud-computing/

Mrdalj, S (2007). Teaching an applied business intelligence course**.** *Issues in Information Systems, 8*(1), 134-138.

Mrdalj, S., & Diallo, A., (2010). Bringing business intelligence into finance curriculum**.** *Issues in Information Systems, 11*(2), 168-175.

Musicant, D. R. (2006). A data mining course for computer science. *ACM SIGCSE Bulletin*, *38*(1), 538-542.

Predixion Insight. (2010). *Predixion software*. Retrieved November 15, 2010 from https://www.predixionsoftware.com/predixion/Portals/0/pdf/Predixion_Insight_DS_FINAL_9_4_10.pdf

Reyes, E.P. (2010). *A systems thinking approach to business intelligence solutions based on cloud computing*. Boston, MA: Massachusetts Institute of Technology. Retrieved November 2, 2010 from http://hdl.handle.net/1721.1/59267

Roiger, R. J. (2005). Teaching an introductory course in data mining. *ACM SIGCSE Bulletin*, *37*(3), 415-421.

Scher, R., (2010, October). Remote access tools. *PC Today*, 20-23.

Thompson, J. (2009). Business intelligence in a SaaS environment. *Business Intelligence Journal*, *14*(4), 50-55. Retrieved October 5, 2010 from http://tdwi.org/research/2009/12/business-intelligence-journal-vol-14-no-4.aspx?tc=page0

Tsai, J. (2009, August). Intelligence in the cloud. *Customer Relationship Management Magazine*, 38-43, Retrieved October 15, 2010 from http://www.destinationcrm.com/Articles/Editorial/Magazine-Features/Intelligence-in-the-Cloud-55488.aspx

Watson, H. J. (2006). BI and data warehousing in universities. *Business Intelligence Journal*, *11*(3), 4-6. Retrieved October 5, 2010 from http://tdwi.org/research/2006/09/business-intelligence-journal-vol-11-no-3.aspx?sc_lang=en

Biography

Dr. Stevan Mrdalj is a Professor of Computer Information Systems at Eastern Michigan University and teaches a range of information systems courses. His research interests include business intelligence, data mining, data warehousing and systems analysis and design. He has over fifty articles published in journals and conference proceedings. He has been a session chair at numerous conferences. He has been a reviewer for seventeen books and he has published supplements for one textbook. He is a member of the editorial board and a reviewer for the Computer Science and Information Systems Journal.