BUSINESS ANALYTICS AS A TOOL TO TRANSFORMING INFORMATION INTO AN INFORMING SYSTEM: THE CASE OF THE ON-LINE COURSE REGISTRATION SYSTEM

Dimitar Christozov	American University in Bulgaria, Blagoevgrad, Bulgaria	dgc@aubg.edu

ABSTRACT

Aim/Purpose	Sharing ideas generated in a Business Intelligence (BI) Applications class to upgrade an Information System in to an Informing System.
Background	Course Registration is the essential university's business process in a university that follows a liberal-arts education model. Almost all categories of users are involved, including students, individual faculties and departments, and administration. A typical Information System, designed to support this process, allows departments to schedule selected courses for a particular time slot and location, and allows students to choose courses to study for the semester.
Methodology	The course project is to design a BI application. Domain knowledge is essential for such projects and course registration was the natural choice for this class.
	The assignment includes (1) identifying the categories of stakeholders; (2) identifying the information needs of different categories; (3) identifying available information sources; (4) identifying how is possible to acquire the additional data; and (5) designing the Extract-Transform-Load (ETL) process and interface scenarios in a way to inform clients.
Contribution	Contributions are in two directions: (1) pedagogy - involving students in such a project motivates creativity, also enforcing students to think in cost-benefit framework may lead to creation of really effective and efficient solutions; (2) practice - implementation of some of the ideas could be with low cost, but with high impact.
Findings	Exploring BI techniques may increase the informing value of existing Information Systems.
Recommendations for Practitioners	Careful analysis of information needs and the way information is used, combined with deep domain knowledge and understanding the value provided by Data Mining techniques, is the way to initiate a process of transforming an Retrieval Information System to better inform clients.

Recommendation for Researchers	Combining pedagogy with practice allows one to overcome routine thinking and may lead to effective solutions. This needs further structuring and research on outcomes.
Impact on Society	Transforming Information towards Informing Systems has a significant impact by allowing users to make rational data driven decisions in an efficient way.
Future Research	The future of this project is implementation of developed ideas and assessment of the results.
Keywords	information systems, informing systems, business intelligence, analytics, course registration

"Information Systems are supposed to inform people,
but in practice they are designed to deliver physical stuff..."
(Buckland, 1991, p. 10)

INTRODUCTION

This paper shares the experiences of a class project, originally initiated for pure pedagogical reasons to demonstrate to students the potential of new technologies to provide improved decisions made by different stakeholders in a single business process. In designing a Business Analytics (BA) application, although the domain knowledge is essential, combining it with capabilities of applying available computer and communication technologies (CCT) is the major success factor. Additionally, one of the objectives of the course was to encourage students to consider cultural and psychological aspects that would influence stakeholders such as users' attitude toward decision making, including their own experience, mutual influence, and perceptions. In this respect, selection of the Course Registration Systems (Reg) was the natural choice for the case to explore in class.

Reg is a homemade system, developed as a student senior project about fifteen years ago. Its major achievement at that time was to allow on-line registration and to maintain all data used by different constituencies in a well designed data base, which allows quite flexible data retrieving.

Reg serves the three categories of users, namely, students, faculties, and administration. Students in the class (mostly seniors) were familiar with one side of the system, which allows them to register for courses for a given semester; the instructor with another side of it, the one designed to allow reviewing lists of students and to help decisions regarding which courses to offer for next semester, students advising, and assigning grades. The class lacks a bit of experience on how the system serves the third party, administration, which provides the opportunity to perform also requirements collection and analysis.

The major objectives of the course assignment were to specify (1) what functionality can be added to the existing system in a way to improve assistance to decision making to the above mentioned three categories of users by using existing data, and (2) what additional data is needed to allow different categories of users to make rational, informed, and data driven decisions in the course registration process.

There are numerous authors sharing experience in teaching BI. I can list here just a few: Piedade and Santos (2010) established a special infrastructure to support students learning; Barneveld, Arnold, and Campbell (2012) emphasize the need of establishing terminology standards; Baepler and Murdoch (2010) shared experience of teaching Data Mining by use of datasets generated via a course management system.

OBJECTIVES

The original objectives in selecting an approach, case, and the entire plan of course were purely pedagogical. A hidden objective was to direct students to acquire skills to investigate and understand dif-

ferences in the way different people use data, as well as the differences in setting priorities, and the differences in making decisions, and, last but not least, how to design appropriate visualization to help them.

The actual result of the class activities went well beyond these initial objectives, resulting in well defined scenarios of what may represent an informing service, what technology may apply, where and how to find and capture the needed data, and how the supporting information may migrate toward informing users' decisions.

The paper is organized in the following way. The next section introduces readers to the registration process in a liberal-arts institution. Those who are familiar with it may skip this section, but for the readers without relevant experience reading the section will allow them to understand the following explanation. The following two sections set in more details the objectives in Reg evolution and the framework of the class project. The section "From Analysis to Design" is dedicated to a scenario of what may represent an informing system. The discussion tries to answer the question of how to achieve the desired services, including how to organize evolution by following Gartner's Enterprise Information Management (EIM) maturity staircase.

IMPORTANCE AND CHALLENGES OF REGISTRATION PROCESS

In a university that follows the "liberal-arts" style of education, the registration process is the essential, major business process, relatively well defined and structured. Briefly, it includes the following steps:

1. Faculty departments suggest courses to be offered in the next semester, including the faculty to teach them, preferred time slots, and other details - whether the course has to be offered in a regular class room or in a computer laboratory, maximal number of seats, etc.

2. The Registrar's office reviews the proposals, summarizes the offers, resolves the conflicts in using facilities in cooperation with the departments, and develops the draft schedule.

3. The schedule is reviewed by the Dean of Faculty, and by students via Students Government.

4. Students submit requests for changes to the Dean, and the Dean, after consultations with the departments, finalizes the schedule.

5. The list of courses is then included into the registration system, which also includes all conditions to register as prerequisites, standing, etc.

6. A student has to decide which courses to Register for by considering the time the course is offered and also whether there will be place in the desired section of the course. Selection order depends on students' standing, more senior students choose earlier. When places of a given section are fully occupied, a waiting list is organized.

The three categories of users – students, faculty and administration – play different roles in this process and face different challenges in making decisions:

1. Students: One of the major characteristic of "liberal-arts" education is freedom of choosing one's own road-map in studying. The selection of courses for the very first semester paves the path and impacts decisions and progress till graduation. Over the semesters the uncertainty in making course selection decisions decreases, as shown in Figure 1. An additional challenge for students is the need to accomplish two categories of requirements – requirements for the selected major, including deciding which major to choose; and requirements for general education. To complete a major, a student has to take certain credits from two, relatively well defined, sets of courses – required and electives. Registration to the courses of the general education component is much more loosely defined compared to those of a given major.

2. Faculty: Faculties are responsible for course offering. In the liberal-arts framework, the course offering is driven by demand. The right number of students registered for a given course allows effective training and efficient application of the selected pedagogical approach – too many students create problems with grading and directing individual advancement, too few students create a risk that the course will be dropped, but also make some of the pedagogical techniques inapplicable.

3. Administration: The registration process is essential for organizational effectiveness. The number of students in class is important not only for pedagogical effectiveness, but whether the university resources – facilities and faculties – are used in an appropriate way and result in keeping the cost within the budget.

The three parties' – students, faculties, and administration – decisions are marked with uncertainty, high risk, and needs of comprehensive information, on time and presented in an useful manner. The level of awareness regarding information needs among different members of these three parties was assessed as low by students in this class after completing the requirement collection and analysis. Intuition, sometimes supported by certain, not very well defined, "business rules", dominates in making decisions, and this was assumed as "appropriate" for a small university where "everything is known".
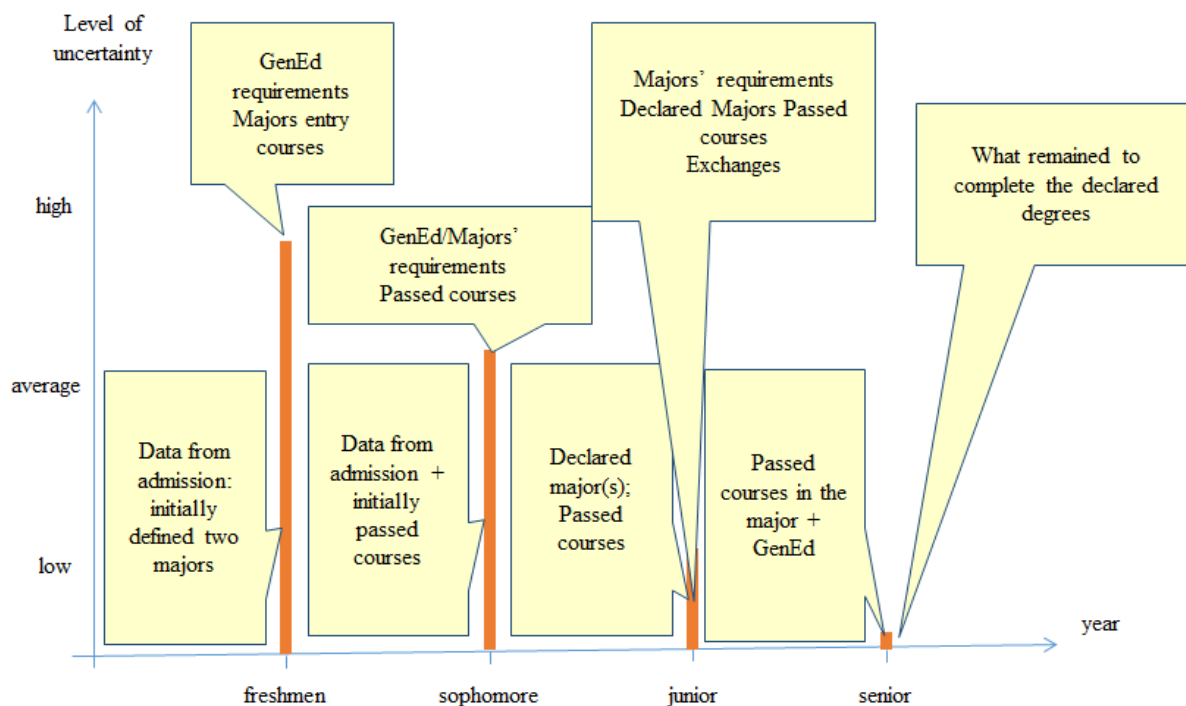

Figure 1. Level of uncertainty in course selection

TRANSFORMING REG INTO AN INFORMING SYSTEM

In principle, the objectives of the Reg system are to support the process of registration and to store data about courses, instructors, grades, completion of graduation conditions, etc. This is a typical information system at the operational level. Moreover, the accumulated historical data will assist user's decisions by providing them with the desired information in appropriate format, time, and completeness.

The mission of the Reg informing system can be defined as students' satisfaction. This includes offering customized support in making decisions in a way to ensure a smooth path toward graduation,

efficient use of resources, and creating effective training conditions. Management of expectations is also an important effect of a system to call it informing.

The transition from a system for registration, to a system to assist decision making is an ongoing process including solving several problems, such as:

1. How to use available data? Business Intelligence (BI) tools may support course offering and selecting. This support is described in the following sections.

2. What additional data are needed? Much deeper understanding regarding the decision making process as implemented on a group and individual level is needed in a way to assess their impact and to design tools to collect, record, and use these data.

3. How to compose consistent classes with students compatible in cultural and psychological aspects in a way to encourage learning?

The section Discussion and Conclusion are dedicated to problems two and three and how to implement these ideas.

METHODOLOGY

A Business Intelligence course usually covers two aspects: theory, introducing the major data models, analytical techniques, and how to interpret results achieved in real-life circumstances; and practice, which is to apply some of the theoretical aspect on a case. In particular, the course project is to design a BI application. Because the domain knowledge is essential for such projects, especially for the ability to assess whether the achieved results are meaningful, the course registration was the natural choice for this class. Additionally, the registrar office provided a data set of 10 years historical data.

In particular, the assignment of class project includes (1) identifying the categories of stakeholders; (2) identifying the information needs of different categories; (3) identifying available information sources; (4) identifying how is possible to acquire the additional data; and (5) to design the Extract-Transform-Load (ETL) process and interface scenarios in a way to inform clients.

In a nut shell, in this particular course, the two major aspects covered:

1. Theory: Students are introduced to special data structures, such as OLAP cubes, defined for analytical purposes; selection of algorithms, such as apriori algorithms for identifying association rules, to perform data analysis; and informing vs. data supply in case of Big Data to train interpretation of results obtained by explored techniques.

2. Practice: To develop skills in applying the more or less abstract theories. Developing skills requires selecting a case for studying, which allows the use of different techniques and also allows to interpret results and to see the benefits of using analytical techniques. The selection of the course registration process is the obvious one. Students, mostly seniors, have sufficient expertise in course registration as a process, the long-term impact of decisions made, and the use of the Reg.

The procedure of performing class project requires students to investigate the case by following the phases:

1. Requirements collection and summarizing the findings. This results in identified problems faced by the different categories of users involved with Reg;

2. Defining what is needed to convert the existing system, as a system designed to register facts and allowing searchable access to stored data, to a system, which may effectively serve different categories of users, but also a particular individual, in making decisions, including data processing algorithms, customization, and visualization;

3. Identifying the limitations of the currently collected and stored data to serve effectively to inform its clients and especially to facilitate support to new students;

4. Defining which analytical techniques may serve to inform users in a better way, what is needed for them, and how to interpret results;

5. Defining ways to collect the needed additional data, to assess their efficiency and feasibility, and suggest tools to address data quality;

6. By following the evolutional sequence defined in Gartner's maturity model, to build a strategy or a multiyear plan to transform the business process toward proactive use of information to allow users to make informed, data driven, rational decisions.

In brief this is a plan for converting an information system, designed to store data and to allow data retrieving, designed by following the "classical" approach, into an informing system, emphasizing the way users explore information in making decisions, by addressing even sophisticated needs of a given individual.

THE FRAMEWORK OF THE CLASS PROJECT

Three categories of problems are usually addressed by a BI application:

1. Learning from the past/history for avoiding mistakes and preserving knowledge;

2. Adequate real time behavior – minimizing the effect of hazards in real-time behavior; and

3. Prediction – evaluating the exposure of potential events in a way to adopt measures to avoid negative development.

The history of developing applications to solve problems in those categories marked the evolution of BI which is in line with the growth of capabilities of available technology. Early applications emphasize "learning from the past", and even the definition of data warehouse offered by Inmon (1992) as "time variant" addressed this understanding. With advancement of technology, data driven support of decision making in real-time is become technically possible even in Big Data settings, and also the predictive analytics is becoming one of the leading areas for development.

Within the registration process these three aspects of BI for the three major categories of users were investigated by students. In brief, their findings can be illustrated in the following manner in a way to transform "old school" information system into an informing system.

FOR STUDENTS

During the registration process, a student has to choose five or six courses to study in the following semester. The information system shows all courses the student is eligible to register for, considering completed prerequisites. The system shows also the remaining seats for each course, as well as the time slot within the weekly schedule the course is offered. Any chosen course has impact on the progress of the given student toward graduation. The major risks of these decisions is to fail a course or to select courses where the student will fail and consequently be placed on probation. The informing system has to warn the student for such options, and to offer more safety scenarios.

1. Learning from the past. By applying a typical machine learning technique, a given student can be assigned to a cluster or category of students with similar profiles. In the class we applied a cluster analysis to the historical data accumulated in the Reg data. The class worked on a data set extracted from the Reg in a way to avoid breaking the confidentiality, including data about registration and grades at least ten years old. Also, all attributes which may allow identification of a student (names, id numbers, etc.) were removed. Belonging to given profile allows showing the student typical road maps, to assess which of them is more successful, in terms of graduation on time, overall grade, and avoidance of mistakes. Additionally,

when student selects a given course for the next semester, the system predicts the grade the student may expect.

2. Adequate real time behavior. If the above information is provided at the time of registration, the student may react by selecting an alternative course of action. Any such change may result in adjustment of the student's profile and the system may suggest how this given action can influence her or his road-map till the end of the study. Another option is to tackle the change of mind, e.g., change the initially specified major, switching interests, etc.

3. Prediction. The three major components of study can be predicted: (1) the expected time when student will be able to accomplish all requirements to graduate with the desired major, (2) the flow of grades over the semesters, in a way to avoid the risk of falling into probation, and (3) which course the student must take in any given semester and which course she or he may take or what are alternative courses to take.

In this way the system will allow a student to make an informed decision with clear understanding and projection of its impact to the entire process of study.

FOR FACULTIES

Faculties (individuals and the departments) are engaged in decisions regarding which courses to offer in the following semester and in which time slot on the weekly schedule. This is a pure demand and supply case. Risks are (1) the course is dropped because few students are registered or (2) the course is too popular with a long waited list that a new section is needed to open. Both cases are not desirable and force faculties to adjust original plans, addressing logistic and pedagogical issues. The first of these risks can be caused by either wrong prediction or overlapping of time-slots of courses desired by the same group of students. This kind of conflict is usually addressed within the departments, but not between the departments. The second hazard is simply a result of wrong prediction.

1. Learning from the past. The common-sense approach is to repeat the same offer as a successful one done in the past. This approach works for within departments decisions, but may result in time conflicts for cross-departmental demand. For example, a student may need to register for a given Computer Science course, required for the major, and for a History course required for general education component, but these two courses can be offered in the same time-slot. By applying apriori algorithm (see Han, Kamber, & Pei, 2011) to behavior of the students in the past, the system may help to identify the frequency of association "If student X takes course A he also takes course B" and to assess the impact of cross-departments time conflicts and to mitigate this risk.

2. Adequate real time behavior. In terms of the registration process, these are decisions to offer courses according to the demand of the current students' population. Maintaining students' profiles allows clustering course demand based on students' individual study road maps. This will allow offering courses in line with student's expectations.

3. Prediction. Projections of student's demand for the following semesters may help departments to renew curriculum on time, to design new courses, and to promote changes in a way to address evolution in the field. This is especially important for rapidly advancing fields like Computer Science and Information Systems. Another, equally important issues that may require prediction is the composition of the class. Psychological compatibility, existence or lack of friendship relations, cultural, ethical, and even linguistic similarity or differences may influence the successful pedagogical approach. Knowing in advance the structure of the class may help in reducing trials and errors and may allow faculty to plan in advance how to approach this particular class.

FOR ADMINISTRATION

The major objective of the administration is to ensure a smooth training process and also to identify direction to assign resources for development. The information generated from solving all of the above mentioned problems can be used.

FROM ANALYSIS TO DESIGN

The performed analysis was directed to identify two of the aspects of transforming the Reg information system into informing service: (1) what is needed to inform users, and (2) what data is needed to do this. The third aspect, architecture of the informing system, especially how to collect, organize, and process data, and how to visualize it to users in support decision making were not set among the original objectives of the course. Instead, students develop some scenarios to illustrative the benefits of applying BI techniques to support different stakeholders in the registration process.

The major reason to limit the scope was original findings regarding the level of maturity of the organization. Students investigated the readiness and awareness of the three categories of stakeholders by following Gartner's EIM maturity model evolution (see Newman & Logan, 2008) and concluded that the university is on the "Unaware" level and, according to Gartner's recommendations, the most efficient way to advance in introduction of intelligence into information services is via "informal" training. The developed scenarios were communicated and this enriched awareness, especially among students, about the benefit of informing services.

SCENARIOS

In designing scenarios for transforming the current Reg into an informing system, we discuss the following three topics mostly from the point of a view of a student:

1. What services a system has to provide in a way to call it "informing"?
2. What can be done with currently collected and recorded data?
3. What additional data is needed and how to find and collect those data?

Figures 2 summarizes the informing services provided by the system at the time of registration, including the information sources for them. It is obvious that not all of these services can be implemented with data stored by the Reg System.

In a nut shell, to become an informing system, the user interface of the New Reg has to allow the following:

1. When the student enters that system, it provides a list of all courses offered for the given semester, open for this student, and colored for urgency in the following way: red - courses the student must take this semester; green - courses that is better to take this semester, not to postpone; yellow - courses not in the main stream of courses to graduate, but which are in the category of "secondary" interest. The New Reg has to use student's profile with study road map, and the history of the student's achievements so far. The data available in the Reg includes the history of courses taken so far, and the usual schedule of offering (e.g., the course is offered every semester, once a year, once in two years).

2. When the student pointed to a course, the New Reg has to provide the following:

    - Time conflicts: (1) with other courses already registered, and (2) with personal habits;

    - Expected grade - "students like you get grades around B" - this requires to cluster students, by profiles and history. Additionally, the system may predict the accumulated GPA.

- Average efforts needed to complete the course - "students like you need in average 5 hours a week to accomplish assignments and to prepare for tests". Additionally, based on the course syllabus, the system provides information about the periods when the most efforts has to be dedicated (for example "in 3, 7, and 10th week the student has to allocate not less than 10 hours to complete the required work");

- Time balance - by accumulating efforts for all, already registered courses, in average and by weeks;

- Classmates - whether there are friends (white list), among already registered for the course;

- Information about the instructor and his/her specific requirements, including comments, but only provided by "students like you".

3. When a student registers for a course - the system updates the profile.

The two phases of this scenario, the front-end as described above, and the back-end to update the profile of the student, are illustrated on Figures 2 and 3.
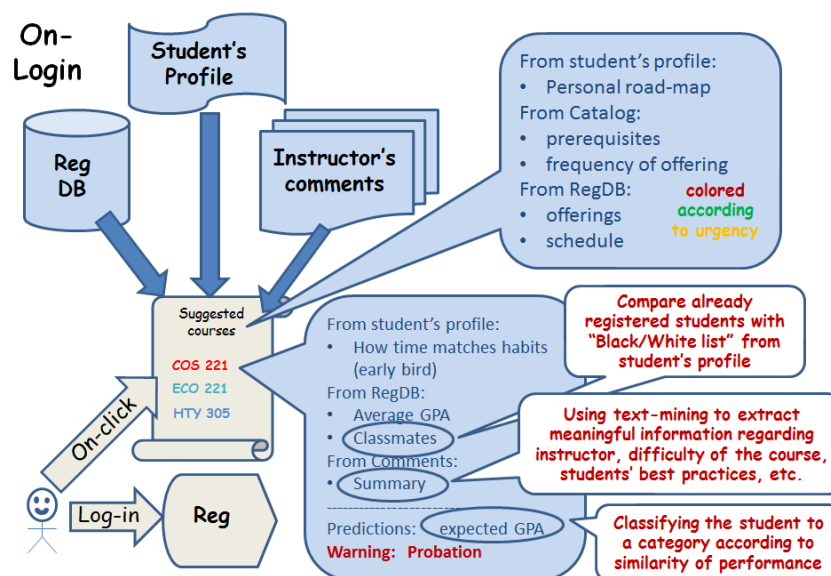

Figure 2. Support of student's decision at the time of registration

The adjustment of the student profile includes adjustment of the road map - on which semester which courses that the student has to take, expected graduation time, expected GPA. This information can be provided to student in a form of warning, for example, "Next semester you have to Register six courses instead of five and you have to plan at least 20 hours in average to study for these courses weekly", or "Next semester the course XXX is not offered usually and to complete the requirements for the major you have to study one additional semester".
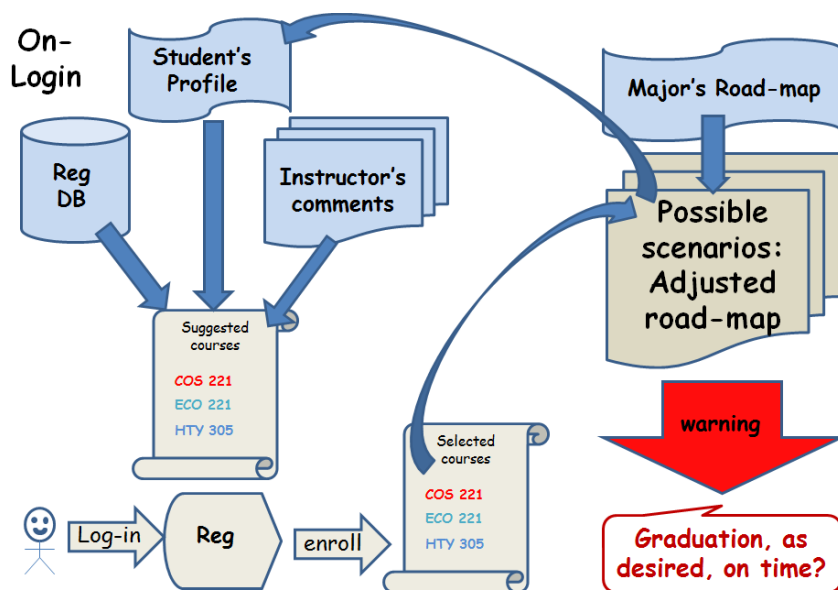
Figure 3. Data processing after confirmation of registration

DISCUSSION: HOW TO ACHIEVE DESIRED PERFORMANCE?

SERVICES AVAILABLE BY USING REG DATA

Limited informing services are possible by using only the data currently stored in searchable way in Reg as one can see in Figures 2 and 3. There are two additional sources of information, which are collected, stored, but not integrated into Reg system - data collected during admission, which includes initially chosen majors; and collection of syllabi for all courses, including grading schemas and description of assignments, which may allow to judge about the efforts needed to complete the course.

By using Reg data, the road-map of a student for a selected major can be defined, the student can be classified to a particular cluster, which will allow the student to assess and manage his or her expectations. In this way, by using the Reg data only, the service for junior and/or senior students, who have already defined majors, may be much richer than this for the younger students. From other side, just the freshmen and sophomore students need better service.

EXTENDING THE DATA SOURCES

To integrate admission data and syllabi with Reg requires decisions and investments on an institutional level. To do this the institution need to pass from the "unaware" to "reactive" stage on Gartner's maturity staircase and, to reach the level allowing implementing all of the identified services, to go further to the "proactive" level.

- **From unaware to aware.** For this step, Gartner recommends informal training. The developed scenarios are designed to accomplish this particular task. Students start understanding the benefits of informing services. Regi Shehi (2014) developed a personal student's scheduler, maintaining a student's developed road-map, with build in functionality to integrate it to Reg system. This is a step towards exploring crowdsourcing (see Howe, 2006) in launching informing services.
- **From aware to reactive**. For this stage Gartner recommends formal training. As a result, investments in integrating data sources for establishing informing services have to be allocated. This requires substantial support by the administration and activities to overcome resistance.

- **From reactive to proactive.** Identifying needs of additional data to allow more comprehensive support and finding ways to collect and to explore those data requires combined efforts by administration and IT staff.

CONCLUSION

The paper shares ideas of how an Informing System can be designed on top of an existing Information System via use of Business Intelligence or Business Analytics tools to enrich the provided services. This is essential for making informed, data driven decisions. These ideas were created by discussions that took place in a class of Business Intelligence Application. The course registration was selected as the most appropriate case to train students in designing information services nowadays.

There are at least two benefits in selecting the Reg system as a case to use in the course. The first one is pedagogical – elaborating a business case so familiar to students encourages engagement, but also provokes creative thinking. Some of identified problems were technically extremely easy to solve, but with clearly visible benefit, as in the case of exploring apriori algorithms in solving the time conflicts' problem. Students clearly understood the critical role of human factor for the success or failure of such a project. The meeting of students with top administrators - president and provost - clearly illustrated the real complexity in convincing others to support the given idea. Finally, students learn how to explore the assessment of a maturity level in proposing a realistic approach to launch informing services. All these will play significant role in their future professional career.

The second one was developing a business case, scenarios, and experiments with historical data that proved the potential and real benefits of using BI tools in university practice. The materials developed are only the very first step toward transforming an existing information system, build on "old school" principles into an informing system, supporting rational, based on data behavior. But this step is essential, because it initiates the process of learning and building "awareness" among the stakeholders regarding the potential benefits of implementing analytic tools.

FURTHER RESEARCH

Adoption of developed ideas is the next step in the process towards establishing informing systems for the next cohorts of students. This step includes more development than research activities. But, from pedagogical point of view, it is to direct students in acquiring skills to identify and address potential resistance. The approach I am planning is to develop several, relatively simple, modules within the existing Reg systems to help administration, the Dean of Faculty in particular, and to provide a bit more comprehensive information services to students. In this way, step by step, the level of awareness of university community may increase.

REFERENCES

Baepler, P., & Murdoch, J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning, 4*(2), Article 17. Retrieved on May 5, 2017, from https://doi.org/10.20429/ijsotl.2010.040217

Barneveld A., Arnold K., & Campbell J. (2012). Analytics in higher education: Establishing a common language Retrieved on March 2, 2017, from https://library.educause.edu/~/media/files/library/2012/1/eli3026-pdf.pdf

Buckland, M. K. (1991). *Information and information systems*. Westport, CT, USA: Praeger.

Han J., Kamber M., & Pei J. (2011). *Data mining: Concepts and techniques*. Morgan Kaufmann.

Howe J. (2006). *The rise of crowdsourcing*. Retrieved from https://www.wired.com/2006/06/crowds/

Inmon W. (1992). *Building the data warehouse* (1st ed.). Wiley and Sons.

Newman, D., & Logan, D. (2008). Gartner introduces the EIM Maturity Model. *Gartner Research Publication, ID Number: G00160425*

Piedade, M., & Santos M. (2010). *Business intelligence in higher education: Enhancing the teaching-learning process with a SRM system*. 5th Iberian Conference on Information Systems and Technologies. Retrieved on February 15, 2017, from http://ieeexplore.ieee.org/abstract/document/5556702/

Shehi R. (2014). *Mobile student scheduler*. Diploma Thesis, American University in Bulgaria

BIOGRAPHY

Dimitar Christozov is a Professor of Computer Science at the American University in Bulgaria, since 1993. He has more than 35 years of research and education experience in areas as computer science, applied statistics, information systems. His recent interests are in the field of business analytics and data science. He graduated Mathematics from Sofia University "St. Kliment Ohridski" in 1979. He completed his Ph.D. thesis "Computer Aided Evaluation of Machine Reliability" in 1986 and D.Sc. thesis "Quantitative measures of the quality of informing" in 2009.

Professor Christozov has more than 100 publications as separate volumes, journal papers and papers in refereed proceedings. He is a founding member and fellow of Informing Science Institute and chair of Bulgarian Informing Science Society; and founding member of the Bulgarian Statistical Society. Currently, Prof. Christozov was elected a member of Executive Board of Bulgarian National Science Fund.