

Global Journal of Engineering Education

Computer-aided business intelligence for non-business solutions

Sławomir Wiak, Dominik Jeske, Maciej Krasuski & Rafał Stryjek

Technical University of Łódź
Łódź, Poland

ABSTRACT: In this article, the authors endeavour to demonstrate how Business Intelligence (BI), when used for non-business purposes, can improve the process of education, and examinations in particular, and prevent certain problems from arising. A distance examination system called *E-matura* that made use of a Microsoft SQL Server 2008 was created, to build analytical data cubes and a data warehouse. Conducting school-leaver examinations using the Internet made it possible to collect data to create a platform to support statistical analyses and draw conclusions concerning the educational process. For the past three years, students have been using this system to sit for a matriculation examination in mathematics. Moreover, the authors demonstrate how BI techniques and data analysis could improve the mathematical analysis of graduates. The data collected during the examinations are used to provide a deep analysis of the graduates' knowledge and to draw conclusions concerning the educational process that took place. The analysis takes place at the chosen level of granularity. Furthermore, using this system, one can analyse the results at the level of student, school, region or country. Some critical issues concerning the development and application of such a platform are also presented and discussed.

Keywords: Business Intelligence, automated examination systems, computer-aided mathematical analysis, IT secure systems, distance examination, e-examinations, *E-matura*

INTRODUCTION

How do particular aspects of a sale look to you? Which regions have the best sales? When is it worthwhile to introduce a new product with the greatest chance of success? Managers look for answers to these and similar questions from Business Intelligence (BI) systems. However, data warehouses also can be used to improve the quality of education.

Why is BI so popular? Because it is a data acquisition and analysis system that supports business decisions [1]. Conclusions are based on historical data and can, for example, determine the financial standing of a company, and enable projection ahead. At present, BI systems' origins date from the 1970s. Companies by then had started to aggregate transaction data, which could be used for business decision support by seeking trends and relationships between the data [2].

BUSINESS BENEFITS FROM BUSINESS INTELLIGENCE SOLUTIONS

Managers often analyse business processes, such as sales, based on dimensions of time, region, branch or product. This type of analysis can be done using data cubes. Data processing, based on analytical cubes, enables answers to be derived for questions such as *How do sales look in a given region?* or *In which regions were the best sales made over the past quarter?*

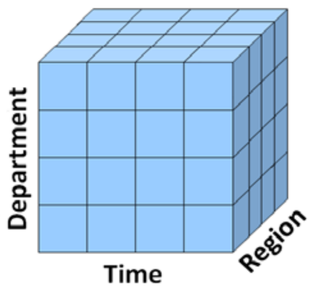


Figure 1: The analytical cube.

Why not, then, in education, search for answers to such questions as: *What types of school produced the best results? Which batches of teaching material gave the worst results?* or *How does the size of a town influence results?* It is possible to derive answers to these questions based on the analytical cubes shown in Figure 1 and Figure 2.

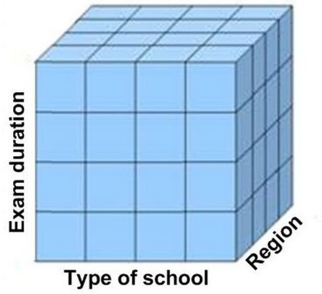


Figure 2: Example of an analytical cube in non-business usage.

NON-BUSINESS APPLICATIONS OF BUSINESS INTELLIGENCE

The *E-matura* Examination System

The *E-matura* examination platform was created under the direction of Professor Sławomir Wiak at the Technical University of Łódź, Łódź, Poland, and the development was carried out under the auspices of the Ministry of Education. Information technology corporations such as IBM and Microsoft have committed themselves to the project, supporting the programming team with professional servers and the required software.

The examinations were held in October 2009, May 2010 and April 2011. Several thousand students from all over Poland took a matriculation examination in mathematics simultaneously via the Internet. It was heavily publicised and attracted positive coverage by the Polish media. The project is under continuous development. On completion of the project, the final product will be ready for use on a larger scale. The goal is not just to create a reliable system to conduct examinations at a distance, but also to produce a system for computer-aided analysis of graduate attributes.

The examination used both multiple choice and open question tests. Students undertaking the examination valued the new technologies, animations in the questions and the speed of achieving results. Teachers involved with the examination received the results of their students. The nature of the project enables statistical, comprehensive and precise analysis of the students’ answers to be derived, hence, improving the quality of education.

Problem Definition

Teachers, examiners and authorities are trying to improve the learning process. Without additional information the process is limited - many of the factors contributing to the learning results are not examined. Analysis only at the level of a particular school or for the whole country does not show all the factors. Hence, *E-matura* offers not only a modern examination, but also great opportunities to study the process of teaching and learning. Sample analysis enabled by *E-matura* includes:

1. Which questions were the most difficult?
2. How does the sequence of questions and answers affect the results?
3. How does the city, region or type of school affect the results?
4. How do factors not directly related to examination factors affect examination results? For example, whether students received tutoring, had a computer with Internet access, etc.

Computer-aided analysis of mathematics graduates can be divided into groups of end users:

1. Students - the result against the class, school, city and country. In which parts of the programme should greater emphasis be placed;
2. Teachers - which parts of the programme should be repeated with students. How do the results of one teacher’s students relate to the background of the school and country;
3. Authorities - which types of school recorded the weakest results? How does the size of the city or the region affect the result of the examination? How do other factors affect the results?

End users use different interfaces. For teachers and students, the results of analysis are presented in the system; for the authorities they can be presented in an easy-to-use spreadsheet file.

Rich User Experience

Today's systems for conducting examinations focus primarily on the so-called test questions known as closed questions, which consist of the question itself and a set of possible answers. Due to the availability of modern technology, such as

Silverlight, the *E-matura* system could go an extra step, that is, to include open-ended questions. This technology allows the construction of questions that include elements such as:

- Interactive animations;
- Checkboxes;
- Interactive areas where the user can drag the selected items;
- Interactive areas where the user can highlight a range of numerical values.

These few examples show that the possibilities of a modern examination go beyond the traditional forms of the examination. Further, using animation and appropriate graphics, an examination becomes more user-friendly to people who often become stressed during tests. This is especially important for young people who prefer this form of examination. For them, it is much more interesting to undertake this type of testing rather than the traditional examination methods. Figure 3 shows a picture that highlights a range of numerical values.

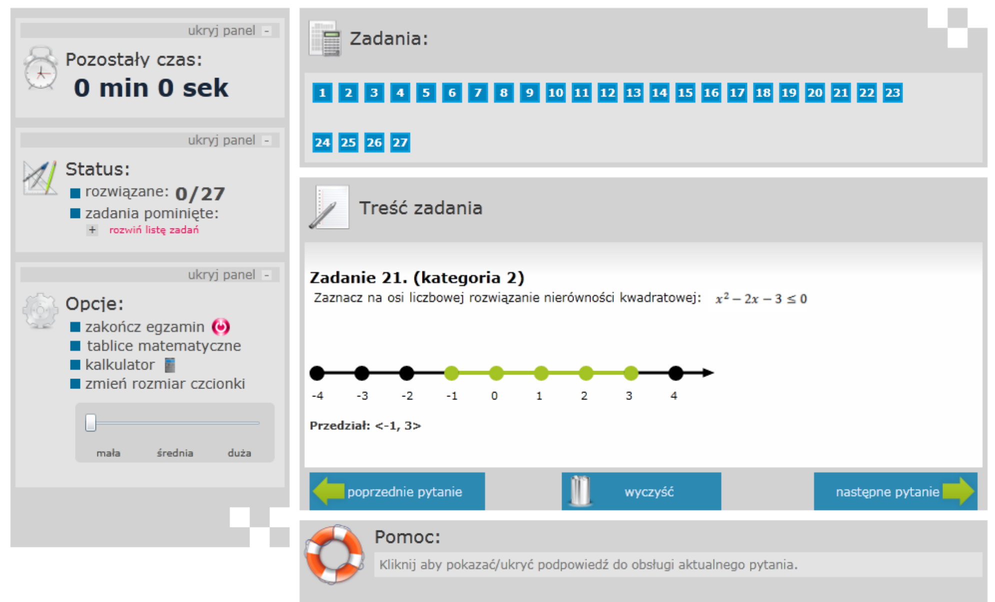


Figure 3: Highlighting a range of numerical values.

When used in the education process, a rich interface also makes learning enjoyable as demonstrated in another project undertaken by the authors, called IMSI e-platform. The use of a modern interface gives new opportunities in e-learning. The authors have provided many features to improve the efficiency of learning. Students can use features, such as the *virtual board*. Teachers can draw something that will appear in the student's application. Moreover, they can write notes quickly and easily. There are two ways of adding student notes to the lecture:

- Notes can be added directly to the lecture file - students are able to highlight some part of text and add a text note;
- In the student application, there is a special tab with notes. The notes tab includes virtual sheets of paper where students can draw or write thoughts.

Other features, which improve effectiveness, are:

- Audio/video streaming - students are able to see and hear the teacher during the lecture.
- Chat, which enables asking questions directly to the teacher.

The use of these features is possible through a modern interface.

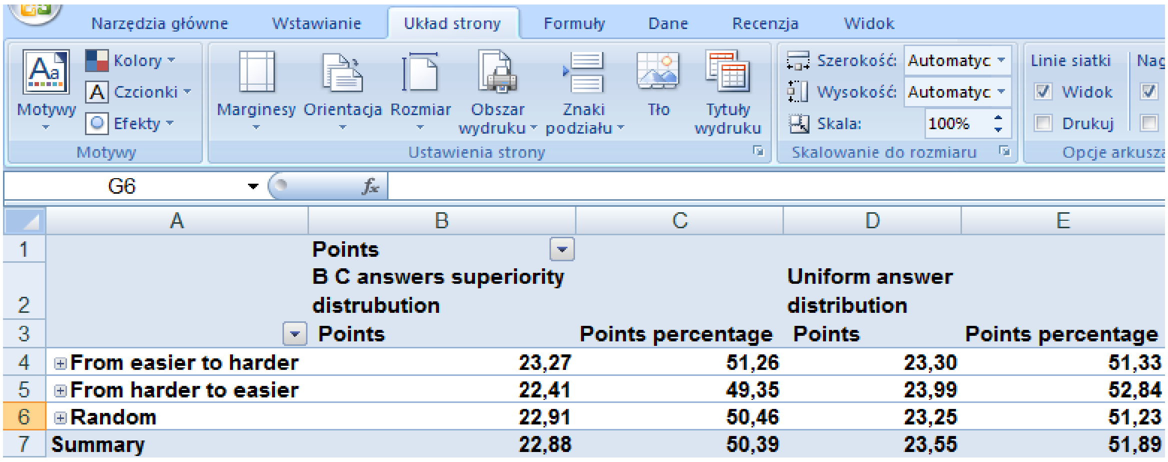
SOLUTIONS

- Which questions were the hardest?

During each examination, hundreds of thousands of responses were collected. If the students answered any given question several times changing the given reply, the system recorded all their steps. The time of giving every answer and how many times they went back to question were also registered. Based on this, one can analyse the results of the questions by student, school or country.

- How does the sequence of questions and answers affect the results?

Three sets of questions were implemented - from simplest to the most difficult, from the most difficult to simplest and randomly. Therefore, it is possible to analyse how the problem of initial questions affects the examination. For research purposes, two answer versions were created: first for 70% questions the correct answer is hitting on B or C (however, 30% is A or B), and in the second version, this order is completely random (Figure 4).

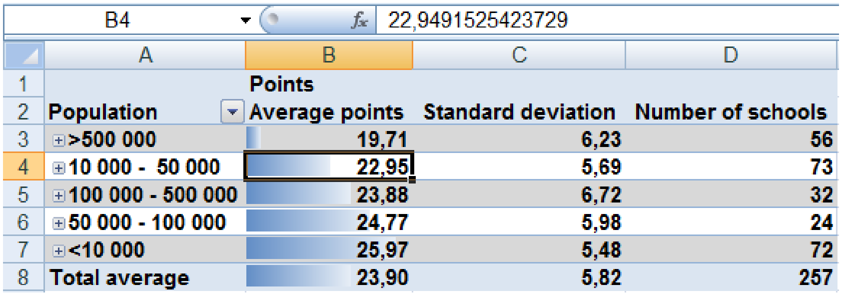


	A	B	C	D	E
1		Points			
2		B C answers superiority distrubution		Uniform answer distribution	
3		Points	Points percentage	Points	Points percentage
4	From easier to harder	23,27	51,26	23,30	51,33
5	From harder to easier	22,41	49,35	23,99	52,84
6	Random	22,91	50,46	23,25	51,23
7	Summary	22,88	50,39	23,55	51,89

Figure 4: A report on how initial questions and correct answers order affect the examination results.

- How do the city region and type of school affect the results?

Because of the capacity to connect the correctness of answers to given questions with the region, city or type of school, it is possible to draw conclusions concerning the improvement of the quality of the education process. Figure 5 shows a report on how the size of the city affects the examination results.



	A	B	C	D
1		Points		
2	Population	Average points	Standard deviation	Number of schools
3	>500 000	19,71	6,23	56
4	10 000 - 50 000	22,95	5,69	73
5	100 000 - 500 000	23,88	6,72	32
6	50 000 - 100 000	24,77	5,98	24
7	<10 000	25,97	5,48	72
8	Total average	23,90	5,82	257

Figure 5: Report on how the size of the city affects the examination results.

One can analyse how the size of the city affected the results. Unfortunately, not enough data have been gathered to draw strong conclusions, but the statistical selection of schools in the future will enable conclusions to be drawn.

- How do factors not directly related to the examination affect examination results? For example, those who received tutoring had a computer with Internet access.

After the examination, students respond to a questionnaire. They respond to questions, such as whether they had received tutoring, who the authority is, how much time per day they spent at the computer using the Internet, if both parents work, and whether parents helped in learning. Answers to these questions are combined with the students' results and the partial results of the questions. Further, it is possible to analyse these data at the level of school, city, region or nation. The responses allow for an examination of correlations between the results of students and factors contained in the questions that are included in the survey.

- One student's result against the class, school, or country. In which parts of the programme should greater emphasis be placed?

The *E-matura* is available also for students to check their knowledge before the examination. Currently, only the individual result is displayed. Students can test their knowledge in a test similar to the matriculation test. Researchers are currently working on displaying a full analysis of a student's knowledge: what areas need additional study, and the topics within these areas that should particularly be repeated.

The current plan is to provide students with additional tasks concerning the topics diagnosed by the system as needing to be repeated. Tasks will be drawn from a pool of tasks from previous years or will be generated automatically by the system - in the tasks for which it is possible to do so.

- Which parts of a programme should be repeated by students?

The teacher will see the results for the whole class, including a comparison of the average result from the school, city and country. It will be possible to refer students' results to the results of students attending the same type of school. Moreover, each student will receive detailed information on areas requiring additional repetition.

DISCUSION OF THE RESULTS

It should be mentioned that 600,000 answers were collected during the examinations. If students answered a given question several times, changing their reply, the system recorded all the steps. The time taken to produce each answer was also registered. Furthermore, by determining the correctness of answers to a given question by region, city, and type of school, it was possible to draw conclusions concerning the quality of the education process. Three sets of questions were implemented for the examination - from the simplest to the most difficult or from the most difficult to the simplest - to be accessed randomly. Further, it was also possible to analyse how the choice of initial question affected the examination.

A15 | lubuskie

Province	Avarage	Maximum	Minimum	Average standard deviation	Number of schools
⊞ **dolnośląskie**	**25,44**	**41**	**8**	**5,60**	**10**
⊞ **kujawsko-pomorskie**	**21,60**	**42**	**7**	**6,62**	**16**
⊟ **lubelskie**	**24,22**	**43**	**7**	**6,66**	**15**
BIAŁA PODLASKA	22,00	30	15	5,75	1
BIŁGORAJ	20,00	25	12	6,27	1
JÓZEFÓW	30,00	38	26	3,65	1
KRASNYSTAW	19,00	29	14	6,78	1
LUBLIN	25,00	42	7	8,72	6
PUŁAWY	25,00	28	21	3,61	1
WISZNICE	16,00	37	9	10,08	1
ZAMOSC	32,00	43	16	10,93	2
ŻÓŁKIEWKA	29,00	34	26	4,16	1
⊞ **lubuskie**	**21,50**	**34**	**13**	**4,55**	**3**
⊞ **łódzkie**	**26,94**	**45**	**6**	**5,82**	**43**
⊞ **małopolskie**	**22,60**	**39**	**9**	**6,20**	**6**
⊞ **mazowieckie**	**26,18**	**41**	**7**	**4,54**	**22**
⊞ **opolskie**	**27,33**	**39**	**15**	**5,27**	**3**
⊞ **podkarpackie**	**21,75**	**39**	**10**	**5,19**	**26**
⊞ **podlaskie**	**23,00**	**43**	**9**	**4,45**	**6**
⊞ **pomorskie**	**23,00**	**40**	**0**	**7,14**	**16**
⊞ **śląskie**	**22,07**	**42**	**6**	**6,68**	**20**
⊞ **świętokrzyskie**	**28,00**	**43**	**13**	**5,30**	**7**
⊞ **warmińsko-mazurskie**	**22,20**	**36**	**13**	**5,43**	**5**
⊞ **wielkopolskie**	**22,56**	**42**	**6**	**5,67**	**21**
⊞ **zachodniopomorskie**	**23,33**	**40**	**14**	**5,75**	**4**

Figure 6: Report on how region and city affected examination results.

The data warehouse in Microsoft SQL Server Analysis enabled processing based on analytical cubes and in-depth analyses were carried out. It is possible to process aggregated data in the popular spreadsheet Microsoft Excel 2007 by using Microsoft SQL Server Analysis Services. In addition, it is convenient for the user to apply familiar tools. When examining the reports, it is possible to find information on the types of school and what part of the programme caused problems and, additionally, how a specific region affected the results. Figure 6 shows a report on how region and city affected examination results. By applying *Drill down* and *Drill up* functions, it is possible to go from the whole to the detail and vice versa. It is possible to conduct data analysis with the differing levels of accuracy. A report concerning this matter is shown in Figure 7.

F3

	A	B	C	D
1		**Points**		
2	**Population**	**Average points**	**Standard deviation**	**Number of schools**
3	⊟ **>500 000**	**19,71**	**6,23**	**56**
4	GDAŃSK	17,00	7,50	5
5	GDYNIA	27,00	6,98	3
6	ŁÓDŹ	19,00	6,71	34
7	POZNAŃ	11,00	2,14	2
8	WARSZAWA	20,00	6,93	11
9	Wrocław	25,00	6,63	1
10	⊞ **10 000 - 50 000**	**22,95**	**5,69**	**73**
11	⊞ **100 000 - 500 000**	**23,88**	**6,72**	**32**
12	⊞ **50 000 - 100 000**	**24,77**	**5,98**	**24**
13	⊞ **<10 000**	**25,97**	**5,48**	**72**
14	**Total average**	**23,90**	**5,82**	**257**

Figure 7: A report showing the *drill down* technique.

Based on the results, one can easily analyse how the size of a city influences the results. The researchers have not gathered enough data to draw strong conclusions, but the statistical selection of schools in the future can help to draw conclusions. Work is continuing on the development of a system able to provide aggregated data for each group of users. Computer-aided diagnosis of mathematical graduates can be divided into groups of end users to achieve the following:

- Students: Results are compared with the class, school, country, including an investigation into the parts of the programme on which greater emphasis should be placed. The *E-matura* system is also available for students to check their knowledge before the examination;
- Teachers: They will see results for the whole class, with a comparison to the average result from the school, country and city. They learn which parts of the programme should be repeated with students, and how the results achieved by their students relate to the background of the school and country. Also, they will be able to compare students' results with the results of students attending the same type of school;
- Authorities: They will be able to ascertain the type of schools that recorded the weakest results, how examination results were affected by the size of the city or the region and how other factors affect the results.

It should be emphasised that end users may use different interfaces. For teachers and students, the result of the analysis will be presented in the system, for the authorities it is easier to use a spreadsheet file.

ENSURING HIGH AVAILABILITY

The classic examination system, running in the on-line environment is activated on a single Web server, which handles all the traffic generated by this and other applications that are installed on the computer. This solution works in most cases because the average numbers of people who use the Web server at any one time are not able to overload server resources. In a situation where tens of thousands of people are referred to a single server at the same time, overload is always a potential problem. It is because any connection to the server requires a certain allocation of memory and CPU time.

To meet the requirement of high availability of the examination, the solution based on so-called *load balancing* was applied. The solution is to build a cluster of servers in which one can distinguish two main parts. The first part is a computer constituting the access point (AP) to the examination. All connections are directed to this computer, but they are not directly supported, but only transmitted to computers in a cluster. Based on the selected load management algorithm, this server redirects traffic to the least loaded server in the cluster.

This solution is highly scalable and allows for virtually unlimited expansion of the cluster. The only limitation is the Internet bandwidth at which the communication takes place. An additional feature that may satisfy this server is decoding an encrypted secure sockets layer (SSL) message and passing it along as decrypted message. This causes a reduction of the burden on the target servers; however, with a large number of connections, it may cause an overload of the server balancing traffic.

In this project, when using IMSI e-platforms, one confronts similar problems, which needed to ensure high availability of the system for the final users. On the e-Learning platform, the solution architecture consists of client computers connected to a central server and teacher computers, which also connect to a central server. The teacher's computer has a very important role in this case, because from this computer the audio-video signal and synchronisation signal are transmitted to the server and, then, are sent to the connected students' computers. Thanks to such infrastructure, the teacher's computer is not overloaded with appeals from users' computers. Experience gained from the above IMSI e-platforms were used in the design of the *E-matura* project.

As mentioned earlier, after the examination, students also complete a survey questionnaire. They answer several questions, covering such issues as whether they were attending tutorials, who was the authority, how much time per day they spent at the computer using the Internet, whether both parents work and whether parents help them in their learning. Answers to these questions are combined with the students' results, and the partial results of the questions. Further, it is possible to analyse these data at the school level, city, region or nationwide. The responses allow examiners to look for a correlation between the results of students and factors contained in the questions included in the survey.

CONCLUSIONS

Business Intelligence is a technology that can be used for more than business purposes. In using the *E-matura* platform, it is possible to analyse trends to improve education and diagnose problematic issues that require additional attention. The use of Business Intelligence can support the education process at every level. Students can check their knowledge, teachers can find out what needs repetition, and the authorities can examine how different factors affect students' knowledge to prevent problems.

Anyone can analyse results based on aggregated data analysis cubes in an easy-to-use spreadsheet. The collected data can be used to analyse the different levels of detail. Information is collected at the level of student responses to individual questions, but it is also possible to analyse the results at the national level. Surveys completed after the examination allow for correlations between various factors and examination results to be sought.

In conclusion, *E-matura* is a modern system for conducting examinations at a distance, but it is also a computer-aided mathematical tool for analysis of graduate outcomes. The collected data can be analysed at different levels of detail. It allows for the collection of information at the level of student responses to individual questions, and there is no problem in comparing results with those at the national level, for example. Survey questionnaires completed after the examination allow examiners to look for correlations between various factors and examination results.

REFERENCES

1. Garnter Reasearch, Business Intelligence Tools: Perspective (2003).
2. Surma, J., *Business Intelligence Systemy Wspomagania Decyzji Biznesowych.* Warszawa: Wydawnictwo Naukowe PWN (2009) (in Polish).

BIOGRAPHIES

Professor Sławomir Wiak holds MEng, PhD and DSc degrees, and is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the International Compumag Society (ICS). He is presently the Dean of the Faculty of Electrical, Electronic, Computer and Control Engineering at the Technical University of Łódź, Poland. He was a Visiting Professor at the University of Pavia, Italy, and University of d'Artois, Bethune, France. His area of specialisation includes computer science and electrical engineering (Computer Aided-Design, Computer Modelling and Simulation, Data Base and Expert Systems, Mechatronics and Engineering Knowledge). He has published extensively in the areas of his expertise and his publications include: monographs (8), books (3), chapter in books (7), editorship of special issues (15), editorship of conference proceedings (14). He is the author and co-author of textbooks (10), contributions to international conferences (149), contributions to national conferences (28), journal articles (94), projects (34), invited lectures (29) and referee reports (432). He delivered invited lectures to his peers at the University of Southampton; the National Technical University of Athens; the University of West Bohemia; the University of Prague, the Czech Academy of Sciences, the University of d'Artois, the University of Maribor, the University of Pavia and the University of Vigo. He served as a member of 23 International Steering Committees, and is Chairman of the International Symposium on Electromagnetic Fields (ISEF) that concerns Mechatronics, Electrical and Electronic Engineering. He is also a member of the World Institute for Engineering and Technology Education (WIETE), based in Melbourne, Australia, where he holds the position of a Vice-President of the WIETE International Academic Advisory Committee.

Rafał Stryjek is a PhD student at the Technical University of Łódź in Poland. He works under the supervision of Prof. Sławomir Wiak and carries out research on the use of Business Intelligence techniques to improve the learning process through e-examination. Rafał is involved in the *E-matura* project, a modern system for e-examination, through which thousands of students from all over Poland take their matriculation examination in mathematics at the same time via the Internet. In this project, he is responsible for undertaking computer-aided mathematical analysis of graduates, database designing and developing business logic at the database level. The results of research will be implemented in the new inference project module and e-tutoring sub-system. Rafał is two-time finalist of the ImagineCup competition in Poland. He has gained professional experience while working with companies such as IBM and Comarch, and his achievements have been confirmed by obtaining professional certificates such as the Oracle Database SQL Certified Expert and the Oracle Advanced PL/SQL Developer Certified Professional. He is a trainer, who supports Polish largest companies and institutions through advanced training and consultation on the application of databases.

Maciej Krasuski is a PhD student at the Technical University of Łódź in Poland, where he conducts research under the guidance of Prof. Sławomir Wiak. His research concerns geographically dispersed database clusters used to provide high availability for e-examination systems. Maciej is involved in the *E-matura* project, which is a system to conduct matriculation examination via the Internet designed to examine in the same time hundreds of thousands of graduates from across Poland. In this project, he is responsible for the design and implementation of database infrastructure that ensures high availability combined with high performance. Maciej is a two-time finalist in the Polish edition of the ImagineCup competition. He has gained professional experience by working with many IT companies in the field of database administration. The experience gained in the IT industry has helped him to conduct research and development projects in this new area of academic endeavour.

Dominik Jeske is a PhD student at the Technical University of Łódź in Poland, where he conducts research under the guidance and supervision of Prof. Sławomir Wiak. His research concentrates on motion detection systems for assisting education purposes. Dominik is involved in the *E-matura* project that consists a system to conduct matriculation examination simultaneously via the Internet for hundreds of thousands of graduates from across Poland. In this project, he is responsible for designing and implementing Internet information server applications that ensure high availability combined with high performance. Dominik is a two-time finalist in the Polish edition of ImagineCup competition. He has acquired professional experience by working with companies he helped to develop logistics systems for mobile devices. Furthermore, Dominik cooperates with Microsoft in their student and academic programmes designed to help students to access the latest technologies on the IT market.