OULUN SEUDUN
AMMATTIKORKEAKOULU

Samuli Malinen

THE POSSIBILITIES FOR A STATISTICAL ANALYSIS OF STUDENT INFORMATION AT OULU UNIVERSITY OF APPLIED SCIENCES

THE POSSIBILITIES FOR A STATISTICAL ANALYSIS OF STUDENT INFORMATION AT OULU UNIVERSITY OF APPLIED SCIENCES

Samuli Malinen
Master's thesis
Spring 2013
Degree Programme in Information Technology
Oulu University of Applied Sciences

ABSTRACT

Oulu University of Applied Sciences
Master of Engineering in Information Technology

Author: Samuli Malinen
Title of Master's thesis: The possibilities for a statistical analysis of student information at Oulu University of Applied Sciences
Supervisor: Kari Laitinen
Term and year of completion: Spring 2013 Number of pages: 42

The number of graduating students will be an important factor in the future funding of Oulu University of Applied Sciences. There have already been some budget cuts made and the new funding model will be fully result based. This thesis was assigned by School of Engineering of OUAS. The initial aim was to introduce the statistical analysis software. As there occurred some problems in deploying originally planned software during the thesis work, the aim had to be changed. The new aim was to study different software options to improve the progressing of studies and their predictability.

A variety of databases and information systems used in OUAS are introduced in one chapter. An important national Raketti co-operation project and its numerous subprojects are also introduced. The importance of national co-operation will be more important in the future. The new, UASses independent actor state of universities of applied sciences will also give possibility to decide which information systems they use and develop either alone or together with other institutes of higher education.

The initial aim of this Master's thesis was to deploy software used for monitoring the studies at the University of Oulu. Because the lack of the features and compatibility issues prevented deploying of that particular software, new methods for the study progress surveillance were studied. Real time data from the current student information system can be a solution for certain user groups like teachers and the heads of certain degree programmes, but the need of study progress surveillance system with more accurate data, for example for directors, means that a more sophisticated user interface is needed. A possibility to use the Oracle Business Intelligence software was therefore studied.

As a result, a lot of experience of the study progress surveillance was received and some new kinds of study surveying systems were developed and deployed. This thesis includes recommendations how things should be done and organized in the future. The necessity for the study progress surveillance is obvious. The only way to make a perfect tool for it is in co-operation with users and developers.

Keywords

Oracle, database, data mining, OBIEE, Business Intelligence

TIIVISTELMÄ

Oulun seudun ammattikorkeakoulu
Master of Engineering in Information Technology

Tekijä: Samuli Malinen
Opinnäytetyön nimi: Tilastollisten analyysiohjelmistojen mahdollisuudet opiskelijatietojärjestelmissä Oamkssa
Työn ohjaaja(t): Kari Laitinen
Työn valmistumislukukausi ja -vuosi: kevät 2013 Sivumäärä: 42

Valmistuvien opiskelijoiden määrä tulee olemaan merkittävä tekijä tulevaisuudessa Oulun seudun ammattikorkeakoulussa. Valtakunnalliset säästöpaineet ovat aiheuttaneet jo jonkun verran budjettileikkaukset ja uusi rahoitusmalli on täysin tulosperusteinen. Opinnäytetyö tehtiin Oamkin Tekniikan yksikön toimeiksiantona. Alkuperäisenä tavoitteena oli ottaa käyttöön tilastollisen analyysin ohjelmisto. Tavoite muuttui alunperin suunnitellun ohjelmiston käyttöönotossa ilmenneiden ongelmien takia opinnäytetyön teon aikana eri opintojen etenemistä edistävien ja ennustettavuutta parantavien ohjelmistovaihtoehtojen tutkimukseksi.

Yhdessä luvussa esitellään Oamkin käytössä olevia tietokantoja ja tietojärjestelmiä. Valtakunnallisesti tärkeästä kansallisesta Raketti-yhteistyöhanke ja sen lukuisista osaprojekteista on kerrottu yhdesä luvussa. Kansallisen yhteistyön merkitys tulee olemaan entistä tärkeämpää tulevaisuudessa. Ammattikorkeakoulujen uusi itsenäinen asema antaa amkeille myös mahdollisuuden päättää, mitä tietojärjestelmiä he käyttävät ja kehittävät joko yksinään tai yhdessä muiden korkeakoulujen kanssa.

Alkuperäinen opinnäytetyön tavoite ottaa käyttöön ohjelmisto, jonka Oulun Yliopisto oli kehittänyt opintojen etenemisen seurantaan. Puutteelliset ominaisuudet ja useat yhteensopivuusongelmat estivät tämän ohjelmiston käyttöönoton. Tarpeet vastaavalle ohjelmistolle on kuitenkin olemassa, joten uusia menetelmiä opintojen edistymisen seurantaan tutkittiin. Reaaliaikainen seurantatieto nykyisestä opiskelijatietojärjestelmä voi olla ratkaisu tietyille käyttäjäryhmille, kuten opettajille ja koulutusohjelmien vetovastuussa oleva henkilöstölle. Yksikönjohtajilla sekä muulla ylemmällä johdolla mutta on tarve opintojen edistymisen seurantaan ennakoitavuuden parantamisen takia, tämä tarkoittaa sitä, että entistä kehittyneempiä käyttöliittymän tarvitaan. Tämän vuoksi tutkittiin mahdollisuutta käyttää Oracle Business Intelligence-ohjelmistoa tietojen seurantaan.

Tuloksena saatiin paljon kokemusta opintojen edistymisen seurannasta ja eräs versio siihen kehitettiin ja otettiin käyttöön työn aikana. Opinnäytetyö sisältää suosituksia siitä, miten asiat pitäisi tehdä ja järjestää tulevaisuudessa. Opintojen edistymisen seuranta on välttämättöntä. Ainoa tapa tehdä täydellinen työkalu siihen on käyttäjien ja kehittäjien yhteistyö.

Asiasanat:

Oracle, tietokanta, tiedon louhinta, OBIEE, Business Intelligence

Contents

TERMS AND ABBREVIATIONS

XDW	X - Data Warehouse, where X refers to a certain educational institute.
Data Warehouse	Integrated combination of several different databases.
OUAS	Oulu University of Applied Sciences
PHP	Hypertext Preprocessor, a programming language.
Asio	Student information system at OUAS.
BI	Business Intelligence allows an organization to gather, store, access and analyze data and use the results for making decisions.
OBIEE	Oracle Business Intelligence Enterprise Edition, Oracle's Business Intelligence software.
OLAP	OnLine Analytical Processing.
SOA	Service Oriented Architecture.

1 INTRODUCTION

One of the biggest challenges at Oulu University of Applied Sciences is that it takes too long for students to complete their studies. The writer of this thesis is no exception in this respect. There are also too many students who drop off and will not complete their studies at all. There can be many reasons for this but one reason might be that the student finds some courses too difficult and quits.

The research problem in this thesis can be defined as follows “How can we find the bottleneck courses by surveying the student data?” What kinds of tools can be used to carry out the necessary analysis of the bottleneck courses?

OUAS has several different databases in use at the moment. The importance of the databases varies, some of the databases are less important but it is easy to say that Asio student information system database is the most important of all the databases as it stores all the study related data of each student. The less important databases are relatively small and contain only few database tables and few thousand rows of data. For example, Asio student information system, which is the largest database, consists of approximately two hundred database tables and millions of rows of data. Asio was originally created in the early 90’s and it is approaching the end of its life cycle. Asio has many different applications for students and staff but does not support a sophisticated statistical analysis.

One chapter of this thesis also provides some information about a Raketti project, which will cause a lot of changes for different database / information systems of Finnish higher education in the near future.

In this thesis three alternative options for further development are introduced for analyzing the bottleneck courses. The purpose of this thesis was to find out how it would be possible to get some relevant information from the databases with the help of some statistical analysis software, and to use this information to improve the studying processes in order to improve the metrics.

The last part of this thesis suggests future plans for further developing of the Business Intelligence software for the use of OUAS.

2 THE DATABASES USED AT OULU UNIVERSITY OF APPLIED SCIENCES

Oulu University of Applied Sciences uses many different databases, as mentioned previously. The most important database is the student information system, which consists of many different software packages and a few dozen database tables. Some of the tables contain millions of rows of information (Figure 1.).

A lot of other databases and information systems are also used. Many of them are built by using Oracle database technology, some of them by using Microsoft technology and the rest of them by using MySQL and PHP.

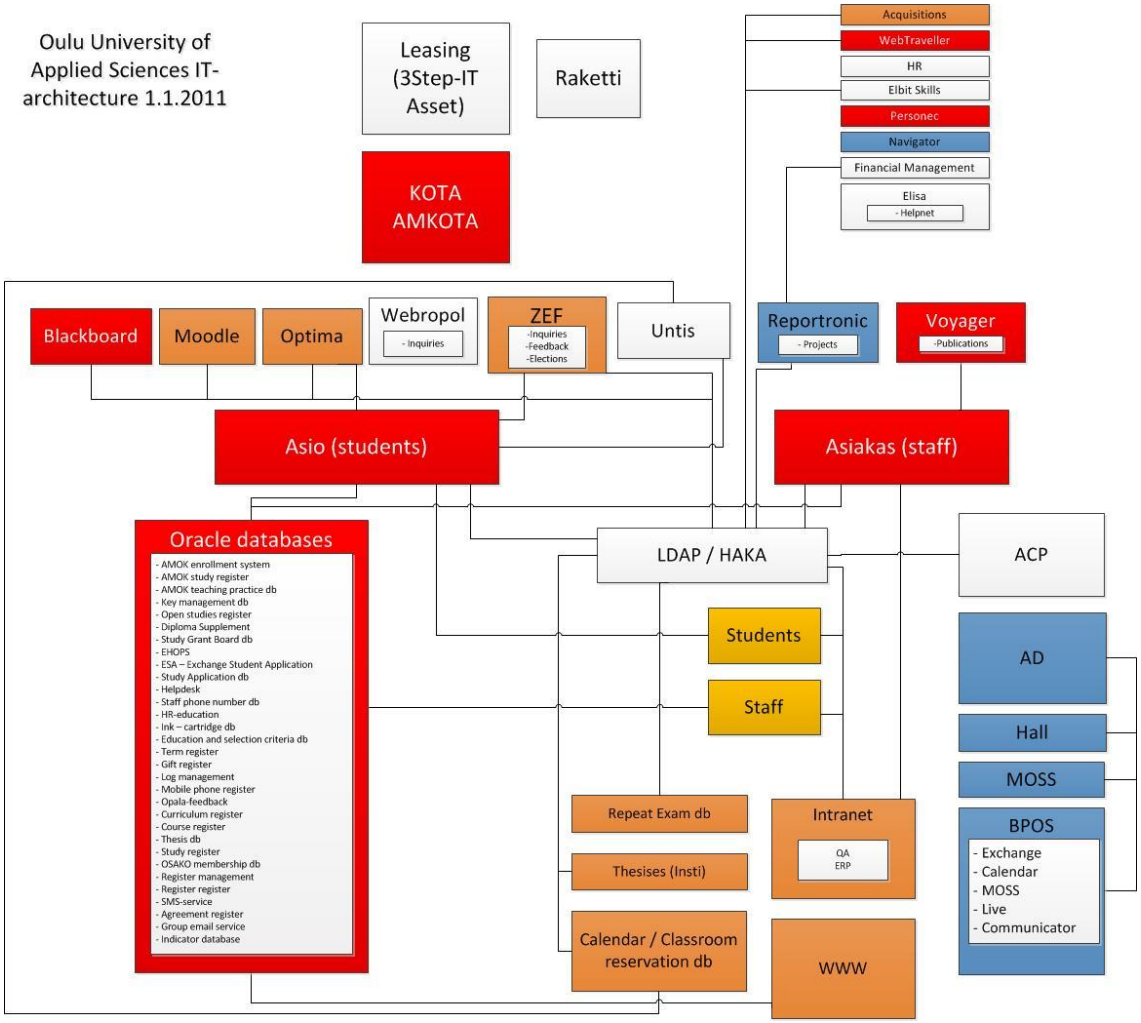


FIGURE 1 Oulu University of Applied Sciences IT-infrastructure 1.1.2011 (OUAS Intranet).

The majority of the databases have been programmed and maintained by the Department of Information Management. Almost all of the databases are still in use and are also needed on regular basis.

The Asio database include the most important data which Oulu University of Applied Sciences owns at the moment. Asio is the master database for all study related information. The OUAS was founded in 1995 due to the need of education. Every study record accomplished to present date in OUAS is stored in the Asio database. In the Asio database can also be found even small partial accomplishments which students have successfully passed. There are also all the courses which students have planned to study or at least supposed to study. The information from the Asio system is also forwarded to many other places like the Ministry of Education and Culture, Statistics of Finland or to the Social Insurance Institution of Finland. Also as an up keeper, The Oulu Region Joint Authority of Education (OSEKK) follows some key information of the Asio database. The student information system has been programmed by Asio-Data. The Department of Information Management (IM) has a direct access to all the data and software and it has therefore made many repairs to the software code. Normally, Asio-Data will also be informed about the bugs that IM has managed to found out from the code and if IM has already made a repair already, the repaired code is then sent to Asio-Data. This prevents the possibility that the same bug comes again with a new software package from Asio-Data. IM also makes many different mass updates to the databases and imports data from different sources directly to the database. A good example of importing data is all the applicant information from a nation-wide application server.

Along with the student information system, it is also necessary to have a staff information system. The staff information system is called “Asiakas/Heta” - “Customer” database, because the same database is used to manage and maintain the contact information of the partners and companions. This database is the master database for all employee related information. There is a direct link between the salary payment system (Personec) and Asiakas database and IM gets new staff member data directly from the Personec-database. A lot of statistical data is gathered via Asiakas database. The Asio student information system and the staff information system are used to gather all the data for example for the Statistics of Finland, which is the Official Statistics gathering authority in Finland. Asiakas and Asio databases are integrated to act together. When a new lecturer starts working at OUAS and his/her information is entered into the Personec or in the Asiakas register, the data also goes automatically to the Asio database. The data of one employee can be selected from both databases

with the help of one database query. This enables to manage all staff related data. As a partial database of the Asiakas database, there are a few tightly integrated systems. Key database consists of the OUAS key information combined with the person who has got for example the key of the locker room and the key of the employee's workroom. The same technique is also used for the staff phone number register and phone model register.

The Information Management department has also several databases for the needs of educational planning. One example is the course register, with all the data required to form a study course. This includes at least the unique course code, the course name both in Finnish and in English, the amount of the study points, the corresponding teacher information, the study materials for the course, the prerequisites, the language of instructions, the contents, the learning outcomes, the requirements and assessments. The course register is tightly connected to the study planning register where all degree programmes and their basic structures are planned and maintained. The study planning register makes the whole process much easier to use and maintain. Before the study planning register, all the recorded data was gathered with dozens of Excel files. The same applies also to the previously mentioned course register. Both of the registers act as the master database and they are integrated into different systems, for example into the electrical curriculum and into applicant's information web pages and also into the Asio student information system.

The electric personal study planning (eHOPS) software was originally specified and designed in IM. A lot of development has been made after the implementation of the first try-out version. Nowadays it has expanded to four universities of applied sciences (Jyväskylä University of Applied Sciences, Kajaani University of Applied Sciences and Mikkeli University of Applied Sciences) as a co-operation program with the national project funding. At the moment eHops consists of thousands of rows of PLSQL code and hundreds of thousands of rows of data in its own database. Luckily, it was originally designed to expand also to a larger scale.

OUAS was given a national notice when we had one of the first online study application forms. Every applicant had a possibility either to make an online application or send an old fashioned paper application. After 2-3, years a common national online application service was also introduced. After that only few applications for example for specializing study programs and Master level applications have been offered to applicants via IM provided application system. All the data from the application system

is transportable to the Asio database. The basic structure and the idea of this application system were also used in the next information system.

The planning officials of international affairs requested a software which could handle all incoming exchange student information from the application to the end of the exchange period including all statistical information. First, we tried to find out if there was some ready software available for this purpose. Also, three other UASes (the same ones mentioned earlier) were eager to participate in the design and specification definitions of the new software. Finally IM decided to program the whole Exchange Student Application (ESA) with them. Now it is used in four different UASes and users have been quite satisfied with ESA.

School of Vocational Teacher Education (SVTE) has a bit different needs than other schools of OUAS. The Department of Information Management has programmed a few special databases for SVTE. Their own enrollment software has been tailored strictly to meet their requirements and needs. They also have their own educational register and practical teacher training database.

Open study register has been designed in co-operation with the corresponding planning officials. The development continues because of the special form of the open studies. Tightly together with the open studies is also a virtual study database. They both are integrated to Asio, too.

The Social Insurance Institution of Finland (KELA) controls all social security benefits in Finland. One of the most important issues between Kela and OUAS is the student's Study Grant and Housing Supplement. The Department of Information Management has made a database of the Progress Surveillance for OUAS's the Study Grant board. It is a highly protected and secure database and can only be accessed by few persons. The statistical information on Progress Surveillance database is only summary based and it will never be part of the larger scale indicator database.

The indicator database was originally made to collect data for Oulu Joint Authority for Education (OSEKK). It was clear from the beginning that the information needed by OSEKK were less the one needed by OUAS's Rector's Office. So the indicator database was first created only for the needs of OSEKK. That included many other indicators that needed in Rector's Office. First, it was only a School Unit based but soon there was also a demand to expand it for the Degree Programme level.

OUAS's student union (OSAKO) has its own member database. This member register is also made by Asio-Data. It is integrated into the Asio student information system, so

that a student without an enrollment for the study period cannot get a membership card with benefits.

Calendar or perhaps it should be called as a classroom reservation system is one very large database programmed with a PHP code by Asio-Data. It is well integrated into Untis software which is optimized to fit four different dimensions: class, classroom, teacher and time. It can provide extremely interesting data for real estate keepers. It is often heard that there is not enough space for teaching, but you can find many empty classrooms at 8 A.M. OSEKK has launched a new work group which aims to find a classroom reservation system for the use of the whole OSEKK.

There are also some softwares which have been developed by eager-to-code teachers and they have been deployed to the whole OUAS. One of the examples is “INSTI” where practical training places and Bachelor’s/Master’s Degree Thesis work places are stored.

There still exist a few other databases, mainly for computers and email systems.

Integrating different technologies together, like Oracle and Microsoft at OUAS can sometimes be a little difficult. Fortunately, new integrating tools exist and mixing different technologies give better results for the customer’s needs. For integration issues, it is important to require open interfaces for data flows (in and out) from the information system provider.

3 THEORY OF DATA MINING

The amount of digital data is continuously growing and as a result of the improved database technology and increased hard disk capacity, huge databases already exist. All data from different human behaviors (for example shopping in the supermarkets, telephone calls, paying with a credit card or using a regular customer card) leave many marks in databases. The interest of tapping these data has grown at the same time. (Han & Kamber 2006, p.1)

The amount of data has also grown at OUAS. Every day hundreds of new entries in database are created.

Data can now be originated from many different databases and registers and eventually be stored with one data warehouse. Data warehouse technology includes data cleaning, data integration, and on-line analytical processing (OLAP). Although OLAP provides analysis techniques with functionalities such as summarization, consolidation and aggregation as well as the ability to view information from different angles, it still needs more tools to make required in-depth analysis. (Han & Kamber 2006, p.3).

The basic idea of data mining is to gather previously created data in a way that the knowledge can be retrieved and reused (Kankanhalli et. al. 2011, p.106).

OUAS has several different standard reports and there are a few hundred indicators from the data. It is probable that there is still much interesting information that has not been found yet.

Data mining of the analysis of observational data sets to find unsuspected relationships and to summarize the data in ways that are both understandable and useful to the owner of the data (Hand & Mannila & Smyth, 2001, 1). Data mining can be viewed as a result of the natural evolution of information technology.(Han & Kamber 2006, p.1).

Data mining refers to extracting (or “mining”) knowledge from large amounts of data. The term is a bit misleading if you think about mining gold from rocks is gold mining thus data mining should have been more appropriately named “knowledge mining from data”. (Han & Kamber 2006, p.5). Many other terms carry similar or only a slightly different meaning such as knowledge extraction, data/pattern analysis, data archeology and data dredging. (Han & Kamber 2006, p.5).

Many people treat data mining as a synonym for another popularly used term, Knowledge Discovery from Data or shortened as KDD. (Han & Kamber 2006, p.5)

Terminology is not very interesting but the idea behind it, i.e. how to get valuable knowledge from vast amount of the data.

Knowledge discovery as a process consists of the following steps:

1. Data cleaning (removal of inconsistent and noise data)
2. Data integration (integration of the multiple data sources)
3. Data selection (select only relevant data to the required analysis)
4. Data transformation (data transformation for appropriate forms for data mining)
5. Data mining (process where intelligent methods are applied in order to extract data patterns)
6. Pattern evaluation (to identify the truly interesting patterns)
7. Knowledge presentation (visualization and knowledge representation)

   (Hand & Mannila & Smyth, p.7).

Steps 1 to 4 are preparation for the data and they can be done by many different forms. Data cleaning routines work to clean the data by filling missed values, smoothing noisy data. Identifying or removing outliers, and resolving inconsistencies. (Hand & Mannila & Smyth, p.48). If there are missing values there are few possibilities to handle that situation. It is possible that you just ignore those missing values and then go on. You can also fill those missing values manually; use some global constraint to fill in the missing values, use a mean attribute or use the most probable value (Hand & Mannila & Smyth, p.63). If you suppose to include data from multiple sources in your analysis, and it would involve integrating multiple databases, data cubes of even files, it is data integration. (Hand & Mannila & Smyth, 49.) If it were be useful for the analysis to obtain aggregate information that is not part of any precomputed data in a database or data cube and you want to include it, it is called data transformation. (Hand & Mannila & Smyth, p 49). The data mining step is a combination of the user's knowledge about the data structure and gathered data. Understanding the business and knowing the source resources are a requirement for building a working and a truly helpful repository (e.g. Khan, 2012, p. 200 and Kuśkiewicz et. al. 2009, p. 78). The end of the process

produces new interesting knowledge about the raw data. (Hand & Mannila & Smyth, p.7).

A data warehouse is a repository of information collected from multiple sources, stored under a unified schema that usually resides at a single site. To facilitate decision making, the data in a data warehouse is organized around major subjects, such as customer, item, supplier and activity. A data warehouse is usually modeled by a multidimensional database structure, where each dimension corresponds to an attribute or a set of attributes in the schema. The actual physical structure of a data warehouse may be a multidimensional data cube. A data cube provides a multidimensional view of the data and allows the precomputation and fast accessing of the summarized data. (Han & Kamber 2006, p.12-13).

Data cube reflection model is used at business intelligence. Data cube is formed from the existing data. After that the data can be presented by different methods for different levels of information needs. For example, when a data cube is formed from OUAS database, the rector and vice rectors are interested in the information summarized from the whole OUAS but Director of the School of Engineering wants to get information, the most interesting part of the data cube is the part where is information about students or staff of the School of Engineering is located. This same drilling method can be used even to the level of one student, if that particular student information happens for example to be interesting to his/her tutor teacher.

Data mining should not be seen as a simple one-time exercise. Huge data collections may be analyzed in many different ways. The amount of the data is growing and new approaches can produce new interesting knowledge. (Hand & Mannila & Smyth, p.23).

Business Intelligence term was originally specified as gathering data from outside of the company, but at the beginning of 1990 software companies brought new software for companies to gather and analyze internal statistical data.

Business intelligence software has developed a lot during the last years and there are several companies who provide software. When business intelligence tools are used to gather data from several different sources, users can analyze data and get some useful information about it. Business intelligence solutions can be used at many levels of the organization and it also works well with the original idea behind this thesis work.

4. NATIONAL RAKETTI PROJECT

Ministry of Education and Culture launched a program called “Raketti” which comes from the Finnish words “RAkenteellisen KEhittämisen Tukena TIetohallinto” which is freely translated “Information Management Supporting Structural Development”. The Raketti project includes many different subprograms. The idea behind those programs is that by doing all the software and the processes in the same way, a big sum of money can easily be saved every year. Also, every software should have good interfaces and be fully integrated in the future. It is estimated that Universities and Universities of Applied Sciences spend over 150 million euros for ICT every year. The national benchmarking co-operation in Finnish Higher Education Information technology has confirmed that estimation (Benchmarking higher education IT, date of retrieval 20.11.2012). If co-operation saves even few percent every year, money can be used more effectively somewhere else. (CSC, date of retrieval 21.10.2012).

4.1 Raketti-OPI

The Raketti-OPI subprogram develops and supports the co-operability of information systems, notions and abbreviations in Finnish higher education. Co-operability and equal reporting require that all study related information is comprised in the same way in every institute. When all institutes of higher education have common understanding of notions and information, the effectiveness between institutes can be evaluated.

At the beginning of the OPI project the main aim was to define and create the basic structure of a new national student information system but this idea was dismissed as an old fashioned thinking and new possibilities for Service Oriented Architecture (SOA) were introduced. First, a SOA based co-operation project between Metropolia and Tampere University of Applied Sciences has produced software called Peppi and it has been planned that it is going to be in production use in February 2013. OUAS also participates in TIPTOP-project (’TIetoon Perustuvaa Tukea Opiskelijan Polulla’, freely translated as ’Knowledge based support for student’s study path’) where the new SOA-based architecture is used to create new nationwide services for students.

According to the Ministry of Education and Culture, they are not going to give commands or instructions on what information systems should be used in the institutes of higher education. It has to be decided in the institute of the higher education itself,

which information systems and technology they use and develop either by co-operating with some other institutes or by themselves.

The OPI subprogram facilitates the co-operation between actors and projects by providing regular co-operation meetings for technology and substance experts. This co-operation work group is named as ‘Synergia’, translated as ‘Synergy’, which refers to the benefits of the situation where experts of different background and different organizations find solutions to multidimensional problems.

The OPI subprogram also provides support for other sub-programs and development projects. Technological services and infrastructures are also available for those projects. The last but maybe the most important thing is that the OPI subprogram develops and maintains the reference architecture for study related information.

4.2 Raketti-XDW

XDW comes from X Data Warehouse, where the X represents both the universities and the universities of applied sciences. Raketti-XDW subprogram has been launched 1.5 years ago. As a result of this program a new data warehouse for all the data which is produced and needed for education will be developed. The main idea of the XDW is that it supports directors in their decision makings with the help of a reliable source of information.

Not only the student information data but also for example financial data, classroom reservation information and personnel information are imported to the XDW database. The data can be viewed with a user interface and there is also a possibility to export data from the XDW to a national data warehouse or some other instance which needs the information (for example, The Statistics of Finland). What is really interesting in the XDW is, when we are thinking about the possibilities to survey bottleneck course data, the fact that a same kind of database will exist in every university in Finland. The University of Oulu has started to plan a new bottleneck course finder tool based on the XDW information. The new software could increase the possible user base up to every university in Finland. What is unfortunate for the OUAS, the University of Oulu bases its development on a Microsoft SQL Server based database and OUAS will make its own XDW database with Oracle technology. OUAS will also send the data from the Oracle based XDW to the common XDW but it is not as real time as it would be in its own XDW. Nevertheless, the same design could be the basis for both of the technologies,

but the coding and maintaining will take twice as much resources compared to a single technology based solution.

When the data is in the same format, you have to code only one well designed software and the same software can be used in every institute of higher education. This would give some opportunities to enhance the co-operation. The new XDW system will greatly improve the information gathering. When all schools of higher education use the same data model and same specifications, data can be compared much easier. In the future there will also be a different model for distributing funds to education. Those institutes who achieve good results and belong to the top quarter receive more money. Currently only 70% of the funding of the universities of applied sciences is based on the number of the students starting their studies and 30% comes from graduated students. In the future the funding will be completely result based, for example there will be an indicator for students who get minimum of 55 credits per year and this can affect for 20% of the total funding.

4.3 Raketti-KOKOA

Raketti-KOKOA is part of the Raketti program which aims to save money by defining Enterprise architecture principles in every public facility. New law is named as Act on Information Management Governance in Public Administration (634/2011).

The purpose of that Act is to improve the efficiency of activities in public administration and to improve public services and their availability by laying down provisions on information management governance in public administration and on promoting and ensuring the interoperability of information systems (Information Management Act, date of retrieval 30.5.2012).

That Act will have a very important role in the future when new systems are acquired or old systems are developed. The Act gives only 6 months of time to UASes to make plan on how they are going to make their own Enterprise architecture.

This makes the original dilemma very interesting. If some actor in Finnish higher education is able to define and even program some fancy software which has good interfaces for common use, they might benefit from it by selling licenses to other actors. Even now, OUAS has several software which are used elsewhere. Software license selling is not a primary object, as of course the teaching and educating together with

the development is the most important, but if OUAS wants to be a forerunner, developing their own processes will be the key to a success.

4.4 Raketti-TUTKI subproject

The Raketti TUTKI subproject focuses on the research and research administration of higher education. TUTKI subprogram guidance was given by the experts of higher education institutes in research, development and innovation along with other key actors in Finland (like Academy of Finland).

At the beginning of the Raketti-TUTKI subprogram (2009-2011) the main focus was on the creation of research and research administration operational architecture. Also, Jure project was part of the TUTKI subprogram which aimed to deploy Release Registry service.

The second phase of the Raketti-TUTKI subproject (2012-2014) focuses on creating common definitions of the terminology and information related to research, development and innovation. Also requirements of the authorities information needs are going to be observed.

Research, development and innovation activities in the universities of applied sciences focus on adopting the results of basic research to a practice level. (Oamk – Oamk tutkii, kehittää ja innovoi, date of retrieval 25.12.2012). Therefore it is difficult to compare the amount of scientific articles to universities where the focus is on more scientific basic research. The renewal of the national information gathering has given good information about the research activities in higher education.

4.5 Raketti-Tietohallinto subproject

The Raketti – Tietohallinto subproject (translated as Raketti Information Management) promotes and develops co-operation between the departments of the Information Management in the institutes of higher education and the Ministry of Education and Culture. The project also supports the practical level compatibility between information systems in the institutes of higher education. Raketti-Tietohallinto subproject also steers the Raketti-VIRTA subproject.

Raketti-Tietohallinto subproject also anticipates and influences all Information Management related requirements and brings the perspective of the Ministry of Education and Culture to the preparation work of the national co-operative planning and designing.

4.6 Raketti-VIRTA subproject

Raketti-Virta subproject was launched in spring 2012. The basic idea of the project is to create a national data resource and to coordinate authorities' data flow.

The work has been scheduled so that in January 2014 the Statistics of Finland should get all the required data/information from the national data resource. Raketti –Virta subproject has released specifications of the required data. The used data model to be used will naturally be the XDW – format and data is exported by generating XML files from the institute of higher education's own data warehouse. When the data transfer is automated, there is no longer a need to make a separate data transfer on certain data gathering days like for example 20.9 or 20.1.

All of the relevant information will be transferred in the new data resource. This mainly includes basic information about the students (address etc. information will be gathered from Population Register Centre) (Väestörekisterikeskus, date of retrieval 27.12.2012) information about their studying rights in a certain degree programme, enrollment information etc.

One of the first new generation's services, 'KSHJ=Korkeakoulujen Sähköinen Hakujärjestelmä' – freely translated as 'Higher Education Electrical Study Application', will use the new data source for the use of the applicants. Even accreditation of previous studies will be easier in the future.

Raketti-Virta subproject has a quite tight time schedule. Data transfers from the institutes of higher education should start in autumn 2013 and every institute of higher education should be able to send its own data. After the first evaluation, OUAS got green light for their own project plan. OUAS also provides a technical solution for three other universities of applied sciences.

5 ETANA SOFTWARE

The original research problem of this Master's thesis was introduced through a co-operation between OUAS and the University of Oulu. After the demonstration of Etana software at the University of Oulu, the same software was planned to be deployed in OUAS too.

5.1 What is Etana?

School of Engineering is a participant in the OTE project, which aims to support students in different ways during their studies. To support the project, the University of Oulu has given the source code of the software it has developed for statistical analysis. The software is called Etana.

Etana software provides tools for some sort of data mining. Etana software makes it possible to find out bottleneck courses that are hard for student to pass. University of Oulu has got good results by using the Etana software. For example they have changed the teacher of the course and after that the percentage of passed students in that course has risen.

Using Etana software requires an integration of several different technologies to work together. First, all the required data is gathered from the database in a certain format. Then a part of the data is handled by statistical analysis software called SAS. After that all information has to be inserted into a MySQL database and from the database the data can be displayed by using a PHP programming on a web page. In the MySQL server there are several database tables in use for the Etana software.

The University of Oulu provides Etana software with an open general public license. This means that all possible changes to the software code also belongs to the University of Oulu. Also the original license owner has to be notified after changes have been made.

School of Engineering in Oulu University of Applied Sciences would like to have similar software too. If it is not possible to deploy the Etana system directly at Oulu University of Applied Sciences, other methods have been taken under consideration. School of Engineering will be the Case Study but the possibility of deploying the chosen method

in the whole OUAS has to be taken under evaluation. There has also been some interest in Etana software in other institutions of Finnish higher education. For example, The University of Vaasa has installed Etana and has made improvements on it.

One of the biggest issues will probably be the data extracting from different student information databases. OUAS Oracle database provides different possibilities to gather and present data, but it is possible that Etana uses some data that it is not found in OUAS student information system databases. If this will happen, other ways to extract the required data will be found out in this master thesis. It is possible to make some replica of the data needed artificially or another possibility is that some kind of small program is made which makes it possible to add the data for the use of the Etana software. Normally, Oracles own PLSQL language is used in a similar situation but some other methods are also possible.

One of the concerns is the fact that SAS is not used as the statistical analysis software of OUAS. It is also possible that the statistical software is not needed at all, if those functions what are used to calculate statistical calculations can be made directly in the database or by a PHP code. However, in this thesis also other methods to get the required results have to be inspected. This Master's thesis should also clearly recommend which software should be taken into use in OUAS.

The PHP code of the software has also to be altered to fit OUAS needs. This may also cause some problems because the author of this Master's thesis is not a very experienced PHP programmer but it is possible to get some external help from the people who are more experienced PHP programmers.

Etana software was developed in the University of Oulu. It was a student project where the aim was to find a tool for searching bottleneck courses and to predict the amount of students who will graduate after a certain time period. Specifications were made strictly for the Faculty of Technology. (Figure 2)

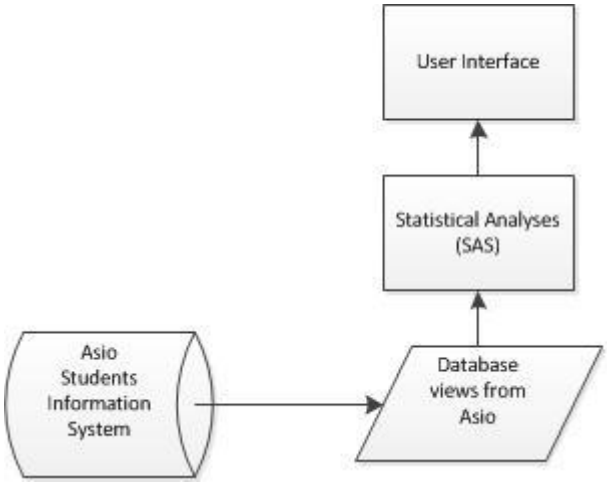


FIGURE 2. Etana data flow made by author.

5.2 Etana installation

All the basic components of the Etana Software were sent to OUAS by the OTE project team leader from the University of Oulu. The packed file contained all of the software code and all the required database scripts to generate the database for the Etana software along with some example data. The materials were compressed to a single zipped file; all that had to be done was to unzip it to one folder. There were some configuration files which were easy to use to make all the required changes. Then all the code and files were installed into one server, the main purpose of witch was to serve as a test server for all kinds of testing and deploying new software. This developmental server is called as “Hiekkis” – freely translated as a sandbox.

There were a lot of different codes and software which were used to try out the new software or to deploy some new software. Also all the database tables and the example data for those tables were easily deployed. Only the PHP administrative tools were needed for importing example data into the database.

The amount of the example data was quite small, just a few rows of example data per table. It was difficult to comprehend how the Etana software actually works and what it needs to be fully operational. The amount of example data was so small that it was not possible to try out graphic examples. From the administrative perspective everything seemed to be fine for testing the Etana software.

When the type of the data which was needed from the OUAS's own student information system was known, the necessary SQL queries from the Asio database were made. Those queries were quite similar when compared with some other previously made queries and only after some minor modifications to the existent SQL queries, all the required data was finally extracted.

Asio student information system uses Oracle database. It could have been possible to make a database view in the Oracle and use that directly from the Etana PHP code, but there had to be some statistical processing with the SAS software before the data was ready to be inserted to the Etana database.

OUAS does not have a SAS license and the cost of one license was inquired from the manufacturer's Finnish subsidiary corporation. Unfortunately, the price was too high for this one tryout project. Fortunately, the OTE-team from the University of Oulu provided a possibility to use their SAS software as it is not reasonable to buy the software before we know for sure that it will be used in OUAS in the future.

When the data from the SQL queries was extracted and sent to the team leader of the Ote project in the University of Oulu, it was also analyzed in the University of Oulu. Some modifications to the queries had to been made because there were some restrictions on field sizes etc. All too large fields were managed to be reduced and fitted to the required size. After a couple of weeks the statistically analyzed data back was sent back to the OUAS. First, some altering to the data files had to be done but then it was possible to import all the data to the Etana database.

5.3 Technical problems in Etana usage at OUAS.

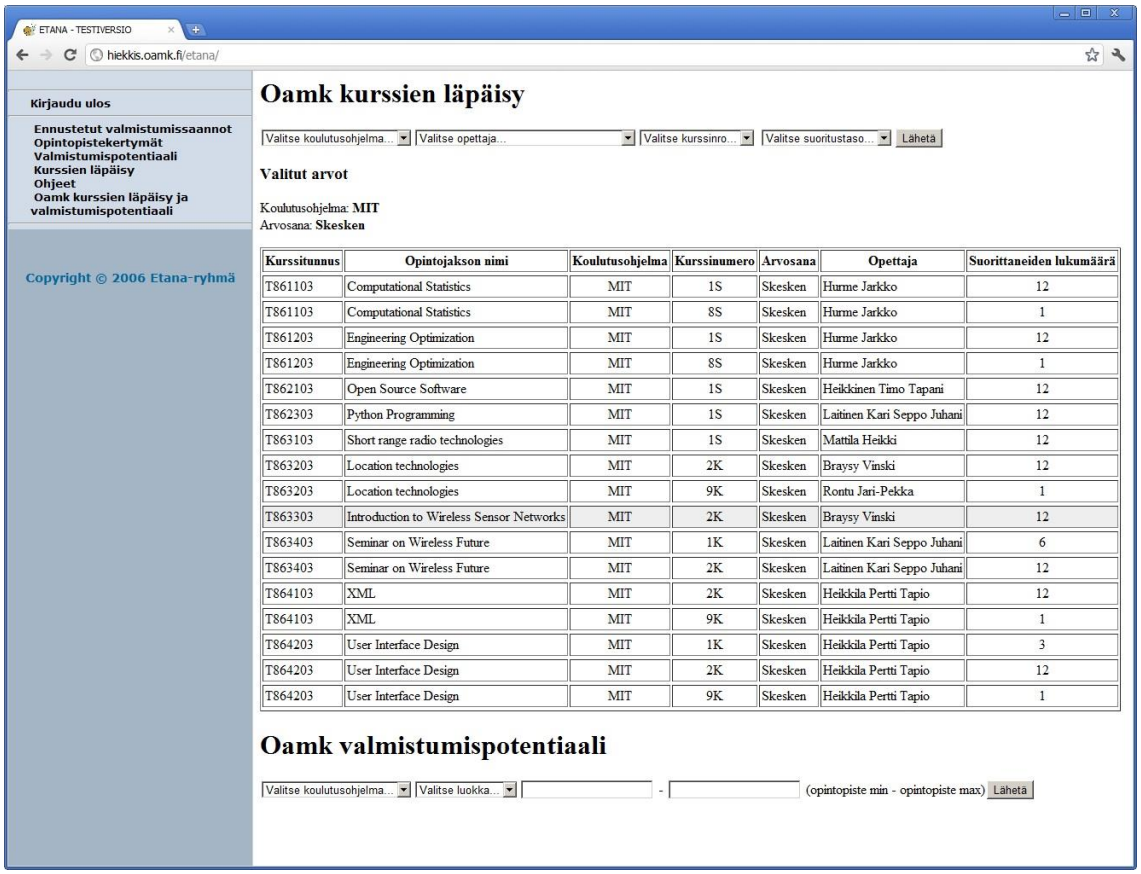
When the importing of data to the Etana database began, it was clear that the Etana data format cannot be deployed directly to the OUAS Etana database. The amount of faculties and degree programmes is different between the Faculty of Technology in the University of Oulu and the School of Engineering in the Oulu University of Applied Sciences. Also, the names of the Degree Programmes were different in the Etana database. The developers of Etana had used shortened Degree Programme names as the database field names. The reason for this was the fact that originally Etana software has been strictly developed for the needs of the Faculty of Technology in the University of Oulu. Not only all the database field names were static, but also many variables had static names and there were a lot of hard-coding used in the PHP-code. Basically it meant that it was almost impossible to adopt the original Etana-code directly for the needs of OUAS. After new database fields in the database were made, the data imports could be done, but there was still a major problem – the PHP code was not able to use the data because of the static coding. The amount of the database table fields increased too much because the old shortened field names for the University of Oulu's use still existed and dozens of new fields were needed for the use of OUAS. Also the original plan that Etana software could someday be widely used in the whole OUAS seemed to be unrealistic. If the whole OUAS would use Etana, the amount of degree programmes would rise up to 30, which would mean that the amount of all database fields would be multiplied by 3. Lastly, there were some issues with the amount of the table fields; they were already in the upper limit.

When the whole Etana system was inspected more closely, one thing became very clear: it is not reasonable to try to use the Etana software and its codes directly in OUAS. There are many good aspects and considered use cases on Etana software and they could be used as an example for the development of some future tool to find bottleneck courses. Needs are still the same and it was clear that it would be much easier to develop a whole new tool to accomplish the required specifications.

5.4 Conclusions about Etana in OUAS

It was discovered that Etana software has to be specified from the beginning and this meant that new coding and implementing should be done in order to get it into use. Meanwhile the University of Oulu also found out the same issue when they were trying to expand Etana usage to other faculties. The University of Oulu made the same conclusion as OUAS: they decided to make new software, Etana II.

During Etana software tryouts and inspections, the directors of OUAS started to require new faster and more agile means for surveying data. This meant that they wanted to have data from the databases in real time, not just one or two reports per year. First tryouts were made with the same user interface as is used in Etana. (Figure 3)



Kurssitunnus	Opintojakson nimi	Koulutusohjelma	Kurssinumero	Arvosana	Opettaja	Suorittaneiden lukumäärä
T861103	Computational Statistics	MIT	1S	Skesken	Hurme Jarkko	12
T861103	Computational Statistics	MIT	8S	Skesken	Hurme Jarkko	1
T861203	Engineering Optimization	MIT	1S	Skesken	Hurme Jarkko	12
T861203	Engineering Optimization	MIT	8S	Skesken	Hurme Jarkko	1
T862103	Open Source Software	MIT	1S	Skesken	Heikkinen Timo Tapani	12
T862303	Python Programming	MIT	1S	Skesken	Laitinen Kari Seppo Juhani	12
T863103	Short range radio technologies	MIT	1S	Skesken	Mattila Heikki	12
T863203	Location technologies	MIT	2K	Skesken	Braysy Vinski	12
T863203	Location technologies	MIT	9K	Skesken	Rontu Jari-Pekka	1
T863303	Introduction to Wireless Sensor Networks	MIT	2K	Skesken	Braysy Vinski	12
T863403	Seminar on Wireless Future	MIT	1K	Skesken	Laitinen Kari Seppo Juhani	6
T863403	Seminar on Wireless Future	MIT	2K	Skesken	Laitinen Kari Seppo Juhani	12
T864103	XML	MIT	2K	Skesken	Heikkila Pertti Tapio	12
T864103	XML	MIT	9K	Skesken	Heikkila Pertti Tapio	1
T864203	User Interface Design	MIT	1K	Skesken	Heikkila Pertti Tapio	3
T864203	User Interface Design	MIT	2K	Skesken	Heikkila Pertti Tapio	12
T864203	User Interface Design	MIT	9K	Skesken	Heikkila Pertti Tapio	1

FIGURE 3. An example view of courses which are completed in the Master Degree Programme of Information Technology (Hiekkis, IM's test server).

5.5 Etana re-designed

The University of Oulu has started to plan a new Version of Etana. During this thesis work some suggestions for the new version were introduced.

The original Etana was designed and coded completely for the use of the Faculty of Technology in the University of Oulu as mentioned previously; the common data model (XDW) for Finnish higher education was not introduced. The XDW – model is a much better solution for a source information system; it expands the possible user base up to every institute of higher education in Finland. Therefore, new redesigned Etana should definitely use the XDW-model as a source data system.

To prevent previous problems with hard coded degree programme names or database table column names, new Etana version should provide flexible and dynamic structures for the amount of degree programmes and faculties/schools. Also, all the column names had to be generic and use for example some kind of table variables to handle source data of different size.

Using open source software would increase the amount of interest because it would lower the total cost of ownership to the user institute. The finance of the institutes of higher education in Finland has suffered big budget cuts and there are still many similar years to come.

One really interesting idea would be the possibility to use the data received from the electrical application system. When a person wants to study in some institute of higher education he/she needs to fill in an application form in the joint application service. There are many questions about the person's previous study success like grade in mathematics or grade in science in a high school. The new version of Etana could find certain correspondences between these kinds of information. This could facilitate setting new requirements for the new applicant's previous study success and thereby help the institute of higher education to get better effectiveness results.

6 REAL TIME DATA FROM ASIO STUDENT INFORMATION SYSTEM

After the idea of the original Etana software was found not to be useful as a bottleneck search engine, some new ideas emerged how the required functionalities could be done. Many of the ideas of the Etana software were quite brilliant, and there is no reason not to use similar queries in the further development. One very important thing which affects a lot of the real time queries is the question: how large queries will be used? If there are only two or three tables joined together, one query will not take too much time to execute. If the query is designed to extract data from four or more database tables, the time for executing the query will increase remarkably. Too heavy queries will also cause some other software, which is used in same databases, to slow down and this is not what is desired (Figure 4).

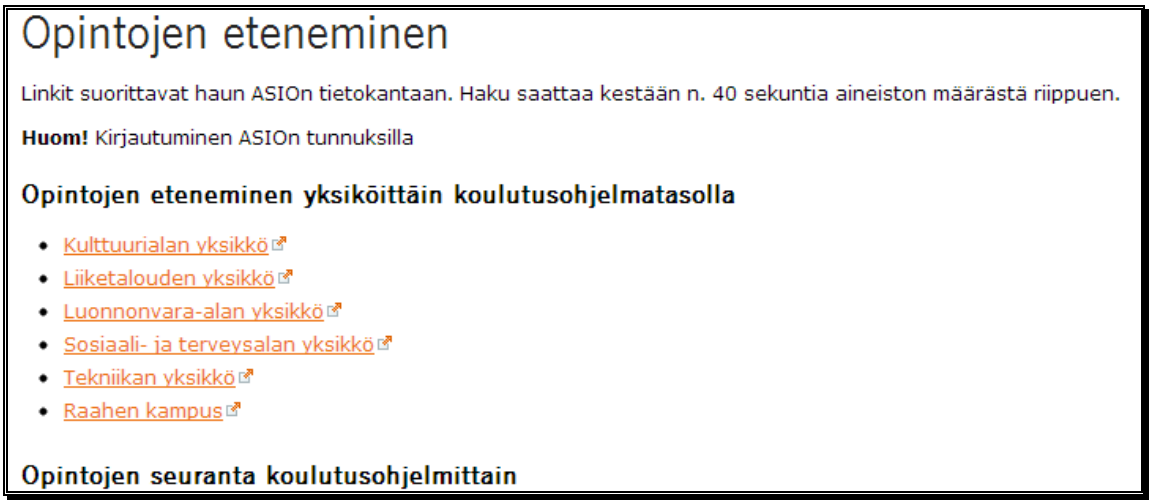


FIGURE 4. The main index of the Stugy Progress subpages in OUAS intranet. There is an indication that the search may take some time. (OUAS Intranet)

Another important issue is also the quality and righteousness of the data. When queries are made directly to the production database, data is not necessarily valid (Figure 5.). There is no ‘human intelligence’ to check out the numbers or the data. Errors in the database data are not only Asio student information system’s problem. Errors can be found in every information system, even if some kind of controlled input interfaces are used. The officials in study affairs office can forget some important data field blank or another error can be some extra enrolment marking from a mass insert or even the student him/herself can change some of their personal data and cause some troubles to the exact queries. To ensure the quality of the data, drilling paths should exist for checking purposes.

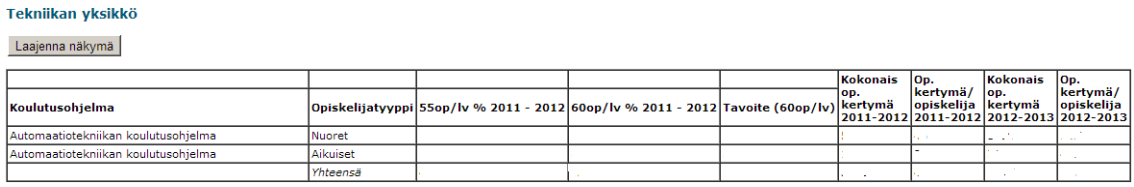
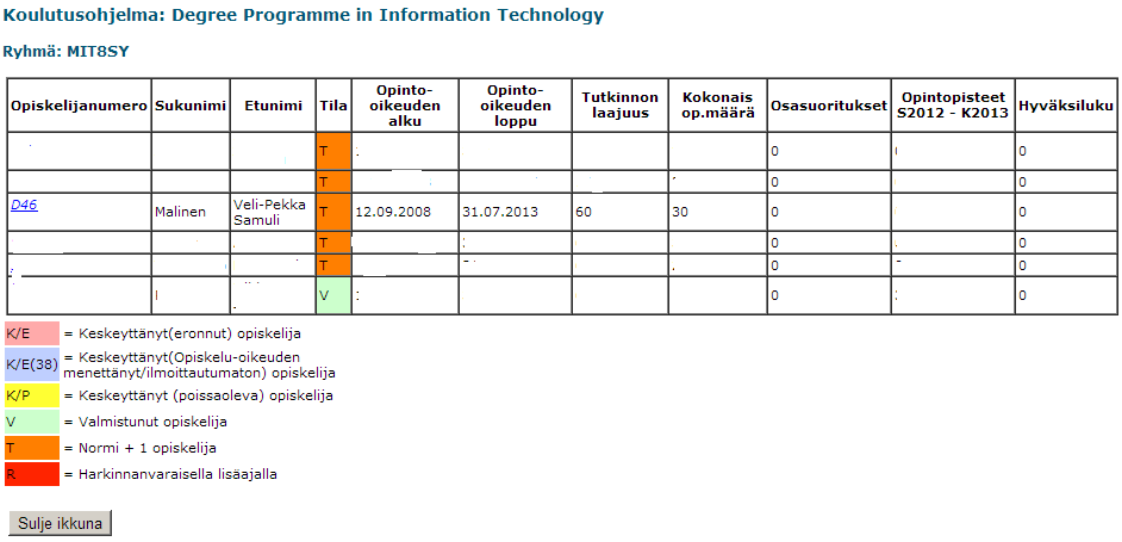


FIGURE 5. An example view from OUAS intranet: The study progress of Automation technology degree programme at School of Engineering (actual data hidden). Actual data comes directly from the database. (OUAS Intranet).



Koulutusohjelma: Degree Programme in Information Technology

Ryhmä: MIT8SY

Opiskelijanumero	Sukunimi	Etunimi	Tila	Opinto-oikeuden alku	Opinto-oikeuden loppu	Tutkinnon laajuus	Kokonais op.määrä	Osasuoritukset	Opintopisteet S2012 - K2013	Hyväksiluku
			T					0		0
			T					0		0
D46	Malinen	Veli-Pekka Samuli	T	12.09.2008	31.07.2013	60	30	0		0
			T					0		0
			T					0		0
			V					0		0

K/E = Keskeyttänyt(eronnut) opiskelija
K/E(38) = Keskeyttänyt(Opiskelu-oikeuden menettänyt/ilmoittautumaton) opiskelija
K/P = Keskeyttänyt (poissaoleva) opiskelija
V = Valmistunut opiskelija
T = Normi + 1 opiskelija
R = Harkinnanvaraisella lisäajalla

Sulje ikkuna

FIGURE 6. An example of a drilled view from OUAS intranet: The study progress of a degree programme in Information Technology (actual data hidden from any other person expect the author). An option to continue the drilling down to the personal study credicts also exists (OUAS Intranet).

There were two possible approaches for implementing Etana's ideas for OUAS use. Meanwhile there was a new intranet coming to every student and staff member daily use. The new intranet acts simultaneously both as a ERP (Enterprise Resource Planning) and as a QA (Quality Assurance) software. The indicator database is integrated very tightly to the intranet, but it provides only data collected in certain days only a couple of times a year. The directors of the OUAS want to have real time data that they can evaluate how the numbers develop before the nation-wide statistics days.

The first task was to find the best database queries which give the exact information on the wanted information. Old queries can always be used, but each query should be checked by running it a couple of times with different variables so that you can estimate how good the query actually is. If it is possible to get some collegial help, it is always useful to check the most important queries together. When you have the best query, which gives you the needed data from the database, there are two different options to choose from to publish the data.

You can use the query as a basic of the database view and create a view which can be displayed for example via PHP in the intranet. This gives a lot of benefits compared to the Etana's method. The data can be so large that it is reasonable to make a software code that gives parameters for the database view and speeds up the SQL query a lot.

Those parameters could be for example the organizational unit or a degree programme. There is also the option to give some value space where for example the user can determine that he/she wants to have all the students who have 30-60 credits of passed courses. If there are many similar queries, all information can still be in the same place so that the user gets all the real time data from one place.

If the queries are still too slow with the database views, one option is to create well designed database tables and use the previously created SQL queries with the help of Oracle's timed services. When all the data gained from the queries, which use several different tables, are inserted into one summarized table, all further queries based only on that one table are much faster. This single table could be easily used as a some kind of a report portal with some personated user interface. This single table could also be referred to as a data cube.

It is quite easy to make data cubes after you have decided what information you want to get in the cube. And when the data is in one cube, it is easy to make bigger scale summaries based on the cubes information. Directors are normally interested in the whole OUAS's information, but at the school level they want to get information on their own degree programmes or even smaller units, like certain classes. The possibility to create drilling paths give many options for a designer, but the designer should also know what results are expected from the user.

Real time SQL inquiries mean that enough calculating power is needed from the servers. Other possibility is to make timed SQL-queries, which servers execute during night-time, but this also means that the data is not directly real time data, only data from the previous night.

If the ultimate solution for data gathering and providing is the primary object, all directors have to give some guidelines and steering and it is up to the information management to provide that information in the way they want to have it. It is easier to make right decisions when you have all the required information. That sentence refers to both, the directors and the IM.

7 POSSIBILITIES OF BUSINESS INTELLIGENCE TOOLS

There have been some sorts of Business Intelligence (BI) tools for a few decades. BI tools are used on a daily basis almost everywhere but the public sector has been quite slow to adopt new BI possibilities to survey data with proper tools.

Oulu University of Applied Sciences needed some solution for the User Interface (UI) for its own Data Warehouse. After experiences which the Department of Information Management got from 'Mittarit' database and after a careful study of what other institutes of higher education were doing, a new Oracle Business Intelligent Enterprise Edition (OBIEE) software was purchased. The amount of licenses is quite smal because at the beginning there is only a quite small amount of users but because of the scalability it is possible to make new purchase for a bigger user amount later.

During autumn 2011 and spring 2012 the Department of Information Management had two separate projects which focused on building a reference deployment of the Data Warehouse. The first project (autumn 2011) basically made a standard ETL process from four different source systems, the second project's (spring 2012) main goal was to create a User Interface with OBIEE software for the Data Warehouse by using the data which the earlier project had successfully loaded.

Because of the interesting nature and the possibility to use the same techniques for searching bottleneck courses, the knowledge gained from the spring project was adopted in order to try to find the ultimate solution for surveying student's study success.

7.1 Preparations for Student Success Survey by using OBIEE

By using OBIEE a user can get data from many different sources (Oracle database, Data Warehouse, flat file, Microsofts SQL server etc.) and view it in a graphical format. Orginally, OBIEE was purchased for the use of the Data Warehouse but OBIEE can also handle a data retrieval from the Asio Student information system. According to the original plan of this thesis work School of Engineering was defined to be the target for every action, however the possibility to expand the chosen method to other OUAS schools also had to be noticed.

Three individual summative tables were created, one table was full of study transcipts, another table included student's personal information and last included teacher data. It would have been possible to bring all real database tables from the production server but it would only result in extra slowness with massive amounts of study records in a certain database table.

7.2 Creation of a repository

A repository is the 'heart' of Business Intelligence software, so the basic idea of a repository is quite the same in the whole Business Intelligence. Data is gathered somewhere and retrieved and used in the form of knowledge. One repository can consist of many different data sources and it also includes the relationship between the different sources. The quality of a repository affects the quality of the knowledge which it produces. Our experiments pointed out that it is very important to create the repository very neatly, and there should be a standardized process to maintain and develop a repository (Figure 7.). That is a reason why creating the repository of the study surveillance system is described in a more accurate way.

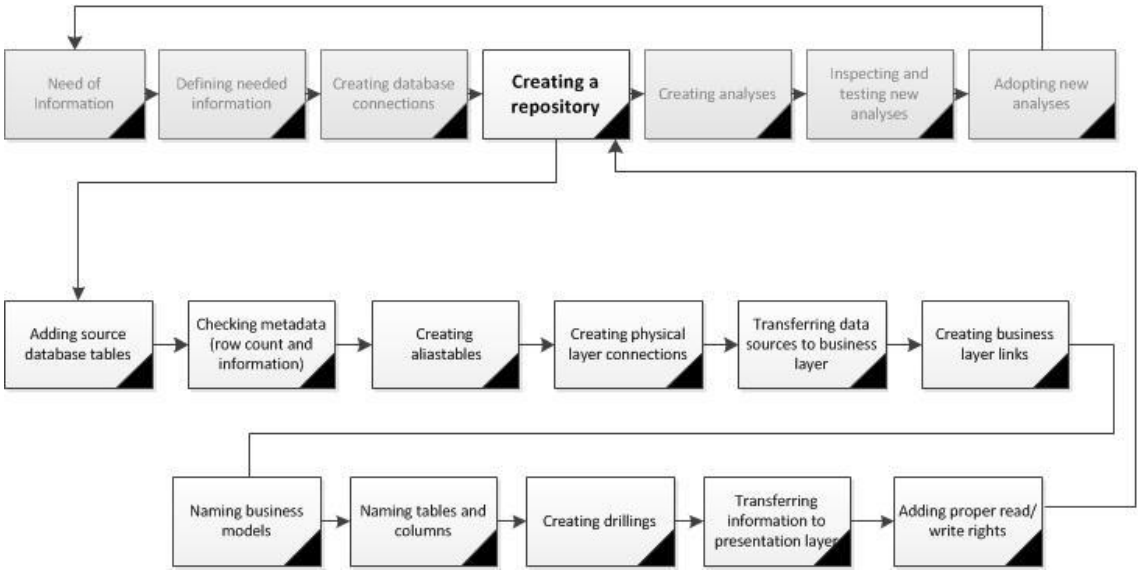


FIGURE 7. Process developed during the deployment of OBIEE for creating a working repository (Final report of spring project).

Successfully built reporting database repository should facilitate fast reporting, fast data loading and accurate reporting. It should also minimize the data volume. According to Khan et. al (2012, p. 202-206) there are eight general rules to be used when creating a good repository:

1. All dimensions of the database should be complete
2. The database tables should be generic
3. Partitioning of the bigger tables to smaller ones.
4. Prudent indexing between tables should be used.
5. Use aggregations in all data columns where aggregation is possible.
6. Close monitoring of both speed of reports and accuracy of data is a helpful way to create a good repository.
7. When designing improvements for the repository can most used reports can be as an example to new ones.
8. If a single record is stored into the database, it should not be broken down into multiple records and allocating the facts across those records.

OBIEEs repository is made by using three different layers: the physical layer that handles all connections between the real actual data sources, the business model and mapping layer where all business logic is made and the presentation layer where all the terminology is transferred to a language that the users can comprehend.

The physical layer does not contain any data. There are only references constructed from different source systems. For study surveillance system there were three source tables from Asio student information system that were connected to the repository and the connections between the tables were created (Figure 8.). Also alias tables were created in order to prevent the problems with the renaming of the columns. All primary and foreign keys were also defined. The source dependencies are encapsulated on the physical layer and it makes the information portable and possible to federate. (Oracle BI Administration Tool Help, 2011) OUAS Study Surveillance system could also use other sources in addition to the Asio like for example students previous study success. There still existed all source files where applicants from the electrical application

system were transferred as new students to Asio. Those source files could be used to make deeper research on the study success.

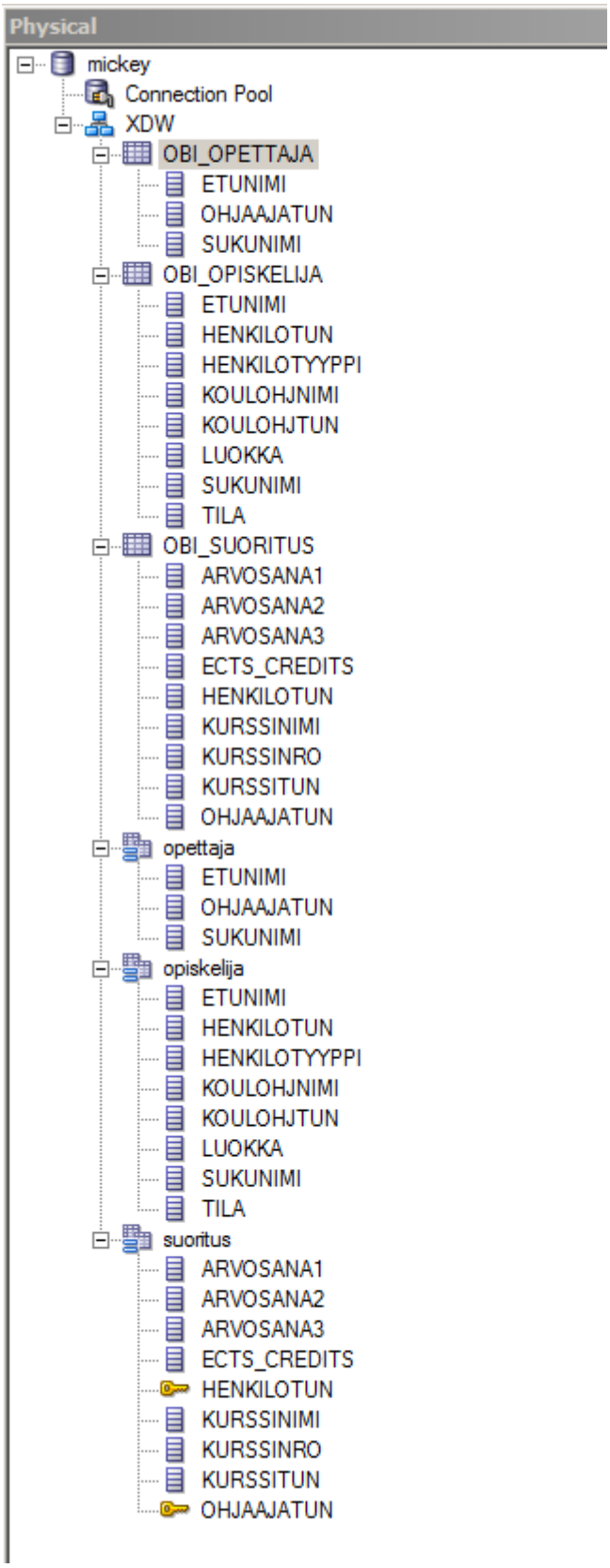


FIGURE 8. The physical layer of the Oracle BI Administration tool. There exists three database tables and their alias tables (Oracle Business Intelligence Enterprise Edition Software).

The Business model and mapping layer are more abstract than the physical layer. They use simplified models from the physical layers objects and it is not obligatory to take all of the columns and tables to the business model and mapping layer from the physical layer. The business model and mapping layer also combine data from different data sources and users do not even necessarily know that multiple data sources are being used. The final business model should represent the user organization's business functions logically and hide all the complexity of the source systems. In OBIEE several business models can exist in one repository. If a study surveillance system will be used in OUAS wide by using the OBIEE software, all business logics should be done on the Business model and mapping layer. For example, if all three grade columns are being imported from the Asio student information system to the repository the highest grade can be deduced to the final grade column. All interesting information like an average of all courses by a certain teacher, an average of all courses, a teacher average per course can also be aggregated on the business model and mapping layer. The possibility to enhance information on the business model and mapping layer was not used in the reference deployment of the repository.

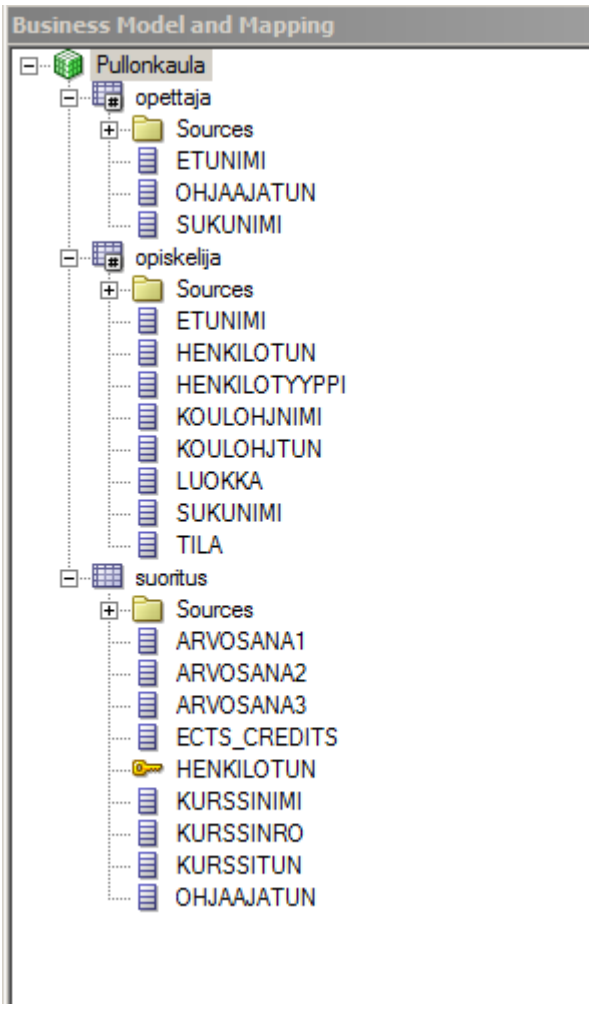


FIGURE 9. The Business Model and Mapping layer of the Oracle BI Administration tool. The business model has been named as ‘Pullonkaula’.

The presentation layer lets the developers of the repository to choose how different objects are displayed to the end user. User customization, security and access rights (roles) are defined on the presentation layer. The final naming of the columns and tables can be done in the presentation layer (Figure 10.) It is even recommended that it should be done on the business model and mapping layer, only security issues were checked during Study Surveillance’s reference deployment.

After all the three layers were successfully done and the global consistency check run successfully, the created new repository was installed on the OBIEE server.

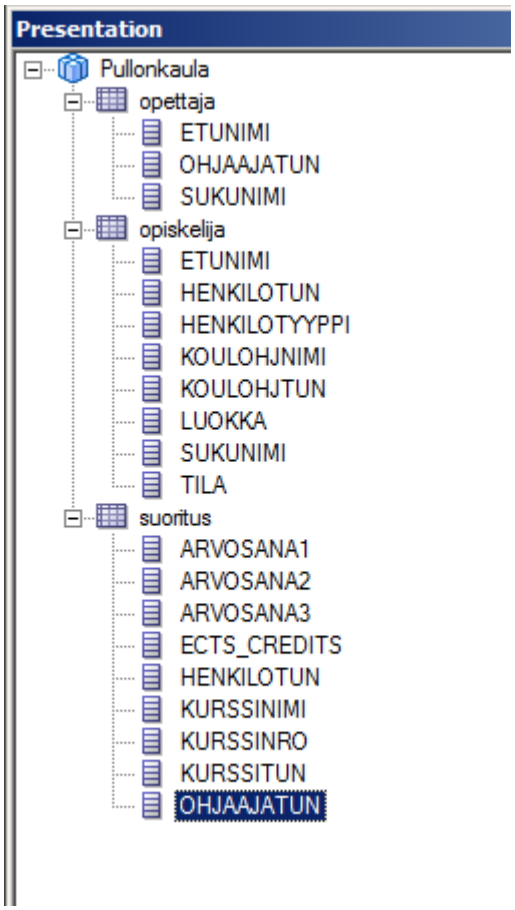


FIGURE 10. The presentation layer of the Oracle BI Administration tool. Only the necessary parts of the Business model have been transferred.

7.3 Using business analysis for Study Surveillance

Unfortunately, even though the consistency of the repository was checked multiple times by three persons, the OBIEE server could not launch the repository properly. The OBIEE server did not even restart properly with the old version of the repository. After some maintenance work the old version of the repository was successfully loaded and the server restarted but that short brake time for the experimental repository testing was already gone. The main purpose of OBIEE is still to be the user interface for the data warehouse but in the same repository, a study surveillance system could also exist as its own business model as mentioned previously. The possibility to create drilling paths from a graphical image is very useful to visualize plain numbers. One example result of the previous project is displayed in Figure 11.

FIGURE 11. Example of the OBI analyse.

8 CONCLUSIONS & FURTHER DEVELOPMENT POSSIBILITIES

It is clear that Oulu University of Applied Sciences has to improve the percentage of completed studies and graduated students. The methods which were tried to create in this thesis work might give some results but more efforts are needed until a suitable and efficient method is found.

It was a bit disappointing that initially so promising Etana software was not the ultimate solution. Still the basic idea behind Etana is very close to the perfect solution. The future will bring a lot of economic difficulties and challenges so solving this dilemma will bring more financing.

Achieving the main objective assigned at the beginning of this work was not very successful. The main reason why the Etana software is not in an operational use is the same reason why the University of Oulu also unsuccessfully spread the use of the Etana in their organization.

The time scale of this thesis was too long for many different reasons. During this time period the University of Oulu also started to deploy Etana software to the whole university. They found out the same difficulties with scalability and hard-coding with the original Etana software. Real time reports from the Asio Student Information system were deployed and it might have been one of many issues which have already improved the metrics. Those real time reports can be found in OUAS intranet. Last decade steering was given mostly by metrics of the ‘calculation’ dates (20.1. and 20.9) and compared to those days. Real-time reporting gives better possibilities to affect things before it is too late. The quality and exactness of the data is not on the same level in real time reports when compared to the data got from data the warehouse.

For the quality and exactness of the data the best solution can be found ín the user interface made for the data warehouse. The data in the data warehouse is only maximum of one day old and it has been checked by an automated verification routine. The automated verification routine also informs the office of study affairs about the possible flaws in the recently input data.

It is highly dependent on the user role which is the best method to survey student’s study success. The best user interface for a class level surveying could be some real time system reporting directly from the database. However, for school level surveying this method can be too inaccurate.

Oulu University of Applied Sciences has many good premises for the challenges of the future. Being one of the largest Universities of Applied Sciences gives a chance to make most of the mandatory service costs very efficient. The amount of young adults who need education is one of the highest in Finland. New independent position of decision-making gives more possibilities to enhance activities and improves the planning of focus areas on research. Development and innovation also give a driving edge for the educational process. All activities should be surveyed by some business intelligence software which can help OUAS in making sensible decisions.

REFERENCES

Bencheit – Benchmarking higher education IT, http://www.bencheit.info/, date of retrieval 20.11.2012.

CSC, http://raketti.csc.fi/en, date of retrieval 21.10.2012.

Hand, D. & Mannila & H. & Smyth P., 2001. Principles of Data Mining. The MIT Press.

Han, J. & Kamber, M. 2006. Data Mining Concepts and techniques. Morgan Kaufman Publishers.

Kankanhalli, A., One-Ki, L. & Kai, H. L., 2011 Knowledge reuse through electronic repositories: A study in the context of customer service support. Information & Management, Volume 48, Issues 2–3, March 2011, Pages 106-113.

Khan, H. , Screen, C. & Ward, A. 2012. Oracle Business Intelligence Enterprise Edition 11g: A Hands-On Tutorial. Birmingham, UK. Packt Publishing.

Kuśkiewicz, L., Małecki, K. & Szarafińska, M. 2009. Information Systems and Knowledge Management in Businesses. Polish Association for Knowledge Management. Series: Studies & Proceedings No. 20, 2009.

Oamk – Oamk tutkii, kehittää ja innovoi, http://www.oamk.fi/yhteistyo_ja_palvelut/tki/, date of retrieval 25.12.2012.

Oracle BI Administration Tool Help. (2011). Glossary. OBIEE 11g Release 1 (11.1.1).

Information Management Act,

http://www.vm.fi/vm/en/04_publications_and_documents/03_documents/20110902Actonl/Tietohallintolaki_englanniksi.pdf, date of retrieval 30.5.2012.

Väestörekisterikeskus, http://www.vaestorekisterikeskus.fi, date of retrieval 27.12.2012.