VILNIUS UNIVERSITY

Renata Danielienė

RESEARCH OF INTELLIGENT COMPUTER LITERACY TEST DESIGN METHOD

Summary of Doctoral Dissertation
Physical Sciences, Informatics (09P)

Vilnius, 2010

Doctoral dissertation was prepared in 2005–2010 at Vilnius University

Scientific supervisor:

doc. dr. Eugenijus Telešius (Vilnius University, Physical Sciences, Informatics – 09 P)

The Dissertation will be defended in council of Vilnius University Informatics Science trend:

Chairman

prof. habil. dr. Gintautas Dzemyda (Institute of Mathematics and Informatics, Physical Sciences, Informatics – 09P).

Members:

prof. habil. dr. Valentina Dagienė (Vilnius University, Physical Sciences, Informatics – 09P);

prof. habil. dr. Vytautas Kaminskas (Vytautas Magnus University, Physical Sciences, Informatics – 09P);

prof. habil. dr. Genadijus Kulvietis (Vilnius Gediminas Technical University, Physical Sciences, Informatics – 09P);

prof. habil. dr. Rimantas Šeinauskas (Kaunas University of Technology, Technological Sciences, Informatics Engineering – 07T).

Official opponents:

prof. habil. dr. Artūras Kaklauskas (Vilnius Gediminas Technical University, Physical Sciences, Informatics – 09P);

prof. habil. dr. Rimvydas Simutis (Vilnius University, Physical Sciences, Informatics – 09P).

The official defence of the dissertation will be held at 9 a.m. on 22th June, 2010, in a public session of the Council of Informatics Science trend in the 10 auditorium at Kaunas Faculty of Humanities of Vilnius University.

Address: Muitinės str. 12, LT-44280, Kaunas, Lithuania

The summary of doctoral dissertation was sent-out on __ May, 2010.
The doctoral dissertation is available at the library of Vilnius University.

VILNIAUS UNIVERSITETAS

Renata Danielienė

INTELEKTUALAUS KOMPIUTERINIO RAŠTINGUMO TESTŲ KONSTRAVIMO METODO TYRIMAS

Daktaro disertacijos santrauka
Fiziniai mokslai, informatika (09P)

Vilnius, 2010

Disertacija rengta 2005–2010 metais Vilniaus universitete

Mokslinis vadovas:

doc. dr. Eugenijus Telešius (Vilniaus universitetas, fiziniai mokslai, informatika – 09P)

Disertacija ginama Vilniaus universiteto Informatikos mokslo krypties taryboje

Pirmininkas

prof. habil. dr. Gintautas Dzemyda (Matematikos ir informatikos institutas, fiziniai mokslai, informatika – 09P).

Nariai:

prof. habil. dr. Valentina Dagienė (Vilniaus universitetas, fiziniai mokslai, informatika – 09P);

prof. habil. dr. Vytautas Kaminskas (Vytauto didžiojo universitetas, fiziniai mokslai, informatika – 09P);

prof. habil. dr. Genadijus Kulvietis (Vilniaus Gedimino technikos universitetas, fiziniai mokslai, informatika – 09P);

prof. habil. dr. Rimantas Šeinauskas (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija – 07T).

Oponentai:

prof. habil. dr. Artūras Kaklauskas (Vilniaus Gedimino technikos universitetas, fiziniai mokslai, informatika – 09P);

prof. habil. dr. Rimvydas Simutis (Vilniaus universitetas, fiziniai mokslai, informatika – 09P).

Disertacija bus ginama viešame Informatikos mokslo krypties tarybos posėdyje 2010 m. birželio mėn. 22 d. 9.00 val. VU Kauno humanitariniame fakultete 10 auditorijoje.
Adresas: Muitinės g. 12, LT-44280, Kaunas, Lietuva.

Disertacijos santrauka išsiuntinėta 2010 m. gegužės mėn. __ d.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje.

Vilnius University

Renata Danielienė

SUMMARY OF DOCTORAL DISSERTATION

The relevance of the topic. Pen and paper knowledge testing when test takers get questions on paper and give their answers on it has long become morally outdated. Nowadays computer-based testing is used for testing knowledge in various areas. Initially simple computer-based tests of fixed length were used. Later on, as there arose the need for security of questions and more effective assessment of test takers‘ knowledge and individual generation of questions, computer-based testing was improved looking for solutions to arising problems. As a consequence, several questions design methods were developed.

Today testing of knowledge using computer-based tests is widely used. Even in educational institutions some exams are being replaced by tests. Therefore, such testing should be done qualitatively. The greatest problem occurs when the testing system needs to choose the most appropriate tests design method. Qualitative testing is especially important in specialized testing systems, such as testing people‘s computer literacy knowledge. That is why this work studies the main test design methods described in literature, their main properties and situations when and which method to use. Having done the analysis of these methods, a solution to computer literacy testing is proposed, which enables to use it according to people‘s occupations and area of activities, thus reducing frequency of question exposure and compact test composition. Compact test composition is considered as a complex term: (1) test shortening according to the test completion condition; (2) test items delivery according to occupational parameters. The new computer literacy test design method is proposed. So far, there has not been developed a specific way for selection of questions for testing people of different occupations with different skills in this area. Usually test questions are selected in a fixed order, random order or adaptively assessing reliability, the length of the test, the size of the question bank, the number of test takers, etc. The level of knowledge in the proposed test design method can be assessed by the help of variables that have not yet been used in testing, i.e. occupation, education, age, experience of work with the computer, etc. For the research of these variables (quantitative and qualitative) intelligent methods will be applied.

Most commonly knowledge tests are not intended for specialized knowledge testing therefore the subject of work “Research of Intelligent Computer Literacy Test Design Method” was chosen. Usually, testing systems of lower- and medium-importance use random function for selection of questions. Sometimes in systems of such level overlapping of questions or various other restrictions are considered. In high- level testing systems adaptive tests are usually used. In such tests questions are constructed according to how the test taker was answering previous questions, regarding the difficulty of the question. Discrimination and/or guessing factors can also be taken into consideration.

Adaptive tests originally were used to define person‘s intelligence level. The first such test was made by A. Binett in 1905 and it used to be on paper. Later adaptive tests started to be used for testing knowledge in other areas as well (Reckase, 1989),

supplementing this model with factors of difficulty, discrimination and guessing that are defined using IRT – Item Response Theory (Baker, 2001; Embretson & Reise, 2000).

While testing computer literacy knowledge it is not enough to construct tests using only the function of random question selection or certain restrictions. Prior to testing computer literacy knowledge it is expedient to know test taker‘s current level of knowledge which can be gained evaluating certain information about the test taker. Scientists studying adaptive tests try to solve the problem by training test taker before testing, giving them pre-tests, etc. Such approaches are described by Mansoor Al-A'ali (2007a; 2007b).

Hence adaptive as well as non-adaptive tests are used to test general knowledge of certain education; however, in both cases the initial level of test taker‘s knowledge is not known. Therefore, it is necessary to study existing test design methods, reveal their main properties and propose a test design method intended for testing computer literacy skills.

The plane of problem research. The topic of test design became of wide interest after appearance of computers. The first adaptive tests were designed and used in a paper version – they were A. Binett‘s tests (Reckase, 1989). Informatics specialists who have been investigating computer-based testing have suggested new methods that enable to carry out tests in both low-level and high-level testing systems (Bartram, Hambleton 2006; Danielienė, Telešius, 2009; Luecht, 2006; Wilson, Engelhard, 2000). Design principles of adaptive testing have been studied by Cyntia G. Parshall (2000, 2002), Wim J. Van der Linden (2000, 2004, 2006), Howard Wainer (1987, 2000a, 2000b, 2007), David Thissen, Robert J. Mislevy (1990), Al-A'ali, Mansoor (2007a, 2007b), Thompson, N. A. (2006, 2007, 2008) and others.

Investigating tests with the help of statistical methods, adaptive tests using Item Response Theory (IRT) were created. IRT is a comprehensive statistical theory about questions to test takers and test properties, and how these properties are related to abilities that are being assessed by the sequence of questions in the test (Hambleton et al. 1991; Harvey, Hammer, 1999; Partchec, 2004; Fraley et al., 2000; Wiberg, 2004).

Principles of non-adaptive tests design have been studied by Tim Davey (1995), Nathan A. Thompson (2008), Lillian C. Folk, J. Zachary March, Robin D. Hurst (37) and others. In such testing systems attention is usually focused on a lesser overlapping of questions in questionnaires, also on the order of question selection or restrictions of question selection.

In Lithuania and in other countries a number of various empirical studies of computer literacy have been carried out by such experts as Eugenijus Telešius, Alfredas Otas, Raimundas Jasinevičius and others (2003, 2004, 2005, 2006, 2007), Lina Markauskaitė and Valentina Dagienė (2004), Diana Šaparnienė (2003), Renata Danielienė (2006). Such kind of researches were carried out by various groups of scientists also, for example testing centres or institutes in the field, e.g. Cedefop (2004), Ministry of Education and Science of Lithuania (2004), Centre of Information Technologies of Education (2003) and KTU Centre of Computer Literacy (2004, 2005, 2006, 2007), European Commission ICT Skills Monitoring Group (2002), etc.

The research object of the work – Test design methods.

The aim of the work – To design an intelligent test design method that could be applied for testing people‘s computer literacy skills regarding parameters of their current skills level.

In order to achieve the aim, the following theoretical and practical tasks had to be done:

1. To perform an analysis of existing test design methods and to summarize the results of the analysis.
2. To investigate ways of test design used in high-level computer literacy testing systems.
3. To suggest ways for determining the initial level of computer literacy skills.
4. To create a computer literacy test design method which would compactly generate test questions according to test taker's occupation and area of activities.
5. To carry out research of specific skills of people of various occupations working with a personal computer, considering to find subsets of skills depending on test taker occupation.
6. To create a prototype of test assembly and composition subsystem and to carry out experimental research.

The research was carried out in 2005–2010.

Research methods: theoretical analysis based on results and conclusions of works by different scientists; system analysis; simulation; assessment; generalization. For clustering statistical data the applied programme processing self-organizing maps *Viscovery® SOMine®* was chosen. The results of the research were processed using *Microsoft Excel*.

The structure of the work

The first part ("Computer-based testing and research of test design methods") overviews computer-based testing and assessment, analyzes typical computer literacy testing systems and describes design methods of computer-based tests. While analyzing these methods, the main properties are distinguished, which theory each method is based on, and what impact that has on other significant factors, such as assessment characteristics, reliability, security, scope of the test and number of test taker, characteristics of item bank, etc. It presents the comparison of described test design methods. In this part typical computer literacy testing systems are also analyzed: their structure, properties and functions are examined, and their advantages and drawbacks are revealed.

The second part ("Analysis of principles and impact parameters of computer literacy testing systems") describes requirements for computer literacy testing systems, the need to develop a new computer literacy test design method is justified, and on the grounds of the conclusions of the first part possibilities of application of new intelligent methods in computer literacy tests are presented. It is looked into what test taker dependent parameters could influence results of the test and a solution to use statistical data to estimate the test taker's initial level of knowledge is presented. For processing of statistical data artificial neural networks are proposed to be used.

In the third part ("Computer literacy test design method") the suggested computer literacy test design method is described. In accordance with the conclusions of the first and second parts, Computerized adaptive tests (CAT) and Automated tests assembly (ATA) methods were proposed as starting ground for the new method design. The test design method is improved by adding supplementary test taker' parameters, such as

occupation, age and others. By applying such test design, questions are even more adapted to test taker and thus their skills are assessed more accurately. By using this test design method, first the test taker‘s initial level of knowledge is estimated according to the statistical data of the Lithuanian ECDLTE testing system. By using the people‘s skills questionnaire data tests forms are made.

In the fourth part (“Experimental-practical application of the computer literacy test design method”) it is shown how solutions found out in the theoretical part can be applied in a typical computer literacy testing system. Here analysis of questionnaire and statistical data is given (real commercial data from the ECDLTE testing system are used, more than 40000 records), as well as the developed prototype is described. Experiment results and generalized research results are presented, the work results are summed up and further perspectives of computer literacy testing system exploitation are discussed.

Propositions maintained in the work

- It is advantageous to use test takers‘ occupational and activity parameters for a compact and rational test composition.
- It is advisable to classify examinees by using artificial neural networks.
- The proposed method can be implemented in a testing system, which allows to make computer literacy tests more efficiently without damage to the fullness and the quality criterions of the test.

Scientific novelty of the work

The new computer literacy test construction method was designed on the basis of the test takers’ experience and features of the existing computer-based test construction methods. This method uses artificial neural networks to assess the likely skill level before the test. The obtained data are used for an intellectual test construction and delivery to the test taker. The difference between this method and non-adaptive method is that that the test is being designed using test taker‘s parameters and compact test composition. Compact test composition is considered as a complex term: (1) test shortening according to the test completion condition; (2) test items delivery according to occupational parameters. Questions are delivered according to the test taker occupation from specific area. The test shortening provides economic benefit and occupancy oriented questions create a psychological comfort to test takers.

Practical relevance of the work

The research results obtained can be used to test people‘s computer literacy skills at educational institutions (schools, colleges, universities, etc.), various educational centres or training companion. The new test design method has been implemented at VšĮ (Public Enterprise) “Information Technology Institute” in the ECDLTE testing system (authorized at the ECDL[1] Foundation level).

[1] ECDL (*European Computer Driving Licence*) is a European computer user‘s certificate certifying that its owner has enough knowledge about IT and is competent to use a personal computer and most common software.

Publication of the results of the scientific work. The results of the scientific work have been published in six scientific publications.

Articles in publications included in the databases of Institute of Scientific Information (ISI):

1. Danielienė R. (2006). Research of typical infrastructure of ECDL. *E-learning and teaching of information technologies and informatics, computer literacy, ECDL and EUCIP frameworks. Databases and information systems: 7th International Baltic Conference on Databases and Information Systems, Vilnius, July 3–6, 2006,* p. 9–16. ISBN 9986-19-920-4.
2. Danielienė R., Telešius E. (2008). Analysis of Computer-Based Testing Systems. *Conference Information: Conference on Human System Interactions, Conference proceedings, May 25–27, 2008 Crakow, Poland,* p. 960–964. ISBN 1-4244-1543-8.

Articles in publications included in other international databases approved by the Research Council of Lithuania:

3. Danielienė R., Telešius E. (2005). ECDL testavimo sistemos formalizavimo problemos. *Informacijos mokslai*. 34 tomas, p.13–17. ISSN 1392-0561.
4. Danielienė R., Telešius E. (2009). Internetinės ECDL testavimo sistemos inovatyvūs sprendimai. *Informacijos mokslai*. 50 tomas, p. 257–261. ISSN 1392-0561.

Articles in other peer-reviewed periodicals and successive international and foreign publications:

5. Danielienė R., Telešius E. (2008). Analysis of Computer-Based ECDL Testing. *In: Nunes, M. B., McPherson, M. (eds.) e-Learning 2008*. p. 243–246. IADIS Press, Amsterdam. ISBN 978-972-8924-58-4.
6. Danielienė R., Telešius E. (2006). Research of a typical e-services based ECDL testing infrastructure. *In: Milosz, M., Muryjas, P. (eds.) Varia Informatica. Technologie i bezpieczenstwo*. p. 111–122. Lublin, Poland. ISBN 978-83-922646-5-1.

The structure and scope of the research.

The dissertation consists of an introduction, 4 parts, and conclusions. The main material of the dissertation is 195 pages, including 20 tables, and 50 figures. Additionally 10 appendices are presented. The list of references contains 17 pages.

SHORT OVERVIEW OF THE DISSERTATION CONTENT

The introduction presents the basic information, regarding the research object of the dissertation, the relevance of the topic, the tasks to be tackled, research methods, scientific novelty and practical relevance of the results, and publications.

Part 1. Computer-based testing and research of test design methods

This part presents the concept of Computer-based testing, and the overview of test design methods and typical testing systems. The section "The Conception of Computer-based Assessment and Testing of Knowledge" reviews the literature related to computer-based testing: history of computer-based testing is presented, low-, medium- and high-level testing systems are described, properties and drawbacks of computerized and online testing as well as typical high-level computer-based testing conceptual architecture of a computer-based testing system are presented. There are also presented knowledge assessment definitions. In the newest literature about high-level testing systems there are frequent recommendations and remarks about the need of innovations in traditional test design methods since traditional knowledge testing systems fail to assess all abilities and skills that are necessary for the 21st century society and working environment (Pellegrimo et al., 2004). Researchers point out that in the education system there is a great assessment gap between what is required for knowledge assessment and what is happening behind the walls of the educational institution. Therefore, it is necessary to develop new theories and methods for high-level testing systems, which would allow to better assess test taker‘s knowledge (Komza, 2009). For this purpose organizations, such as International Test Commision (ITC), or the European Commission's Joint Research Centre (JRC) have been established, and projects related to testing quality are underway. For example, there are more than 20 projects underway at Cambridge University whose aims are to investigate the impact on knowledge assessment by new technologies (Harding, 2006). Some other organizations (e.g. Becta, 2006) or research groups (e.g. Ridgway, McCusker, & Pead, 2006) have presented research projects concerning increasing online testing.

The section "Analysis of computer-based tests design methods" presents the overview of test design methods. During the analysis test design methods were grouped according to the similarity of their functions and each of them was described in detail. While analyzing these methods, the main properties are distinguished, which theory each method is based on, and what impact that has on other significant factors, such as assessment characteristics, reliability, security, scope of the test and number of test takers, characteristics of the test item database, etc. It presents the comparison of described test design methods.

In the section "Comparative analysis of computer literacy testing systems" a comparative analysis of high- and medium-level typical testing systems is performed: their structure, properties and functions are examined, and their advantages and drawbacks are revealed. After performing the analysis of these testing systems, it was found out that (1) those testing systems use nearly the same type of questions, (2) the test taker's initial level of knowledge is not assessed, (3) they use traditional test design methods where questions are presented in a random order, and (4) both commercial and free testing systems test general knowledge in a certain area.

The section “The use of question types in computer literacy and other knowledge testing systems” describes a comparative analysis of question types used in computer literacy and other knowledge testing systems. The aim of the analysis was to find out what types of questions are usually used in testing systems of various areas. The results obtained suggest what types of questions can be used in the intelligent computer literacy test design method in question. For the objective assessment of knowledge it is not recommended to stick to only one type of questions, but instead to use more various assessment methods of training results (Andziulienė, 2004).

The results of different authors’ works showed that both adaptive and non-adaptive test design methods are used to assess general knowledge in a certain education area; however, neither of the test design methods provides tools that would help to estimate test takers‘ initial level of knowledge. For high-level testing systems there is the need to develop new theories and methods that would enable to better assess test takers‘ knowledge.

Part 2. Analysis of principles and impact parameters of computer literacy testing systems

In this part requirements for computer literacy testing systems are described, the need to develop a new computer literacy test design method is justified, and on the grounds of the conclusions of the first part possibilities of application of new intelligent methods in computer literacy tests are presented. It is looked into what test taker dependent parameters could influence results of the test and a solution to use statistical data to determine the test taker‘s initial level of knowledge is proposed. For statistical data processing artificial neural networks are proposed to be used.

In the section „The need of computer literacy testing method” work statistics of Lithuanian ECDL (European Computer Driving Licence) testing system is presented, which demonstrates that testing has been increasing in the last few years, and so it is necessary to provide more qualitative testing services. The testing system stores various information related to tests, questions, examiners, testing centres, test takers and so on. At the end of the section a hypothesis is formulated that statistical data about a test taker could have influence on the test results.

In the section „Requirements for computer literacy testing system” on the basis of the analysis of the ECDL Foundation approved documents and quality assurance system documents for the Lithuanian ECDL testing system, a requirements table for computer literacy testing system is made, according to which tests for test takers should be designed adaptively regarding the data they present. Similarly, the difficulty of questions and their overlapping in the test have to be evaluated. Question forms have to be equivalent. The length of the test can be variable if the set passing/failure level is reached or when the maximum number of questions is given.

In the section „Possibilities of application of new intelligent methods in computer literacy tests” on the basis of advantages of test design methods and requirements defined for computer literacy tests, a table of methods suitability for computer literacy tests is made. According to that table ATA (Automated Test Assembly) method is the most suitable method which should be improved by introducing certain adaptivity features.

In addition, in this section testing processes are compared, when together with traditional test design methods the description of the method in question is presented and

two new parts are included in testing. They are estimation of test taker's initial level of knowledge, and delivery of questions using statistical, version and test taker's parameters (e.g. occupation, gender, age, experience). By using such test design method questions are individualized and in this way cheating possibilities are minimized. In addition, various parameters are attached to questions, which help to design test forms. Such test design technique satisfies the requirements for computer literacy testing system most. It would also enable to assess test takers' real computer skills more accurately. In order to choose questions more accurately, before the assessment process it is advisable to conduct a survey about preferable ways to perform activities on the computer for test takers of different occupations.

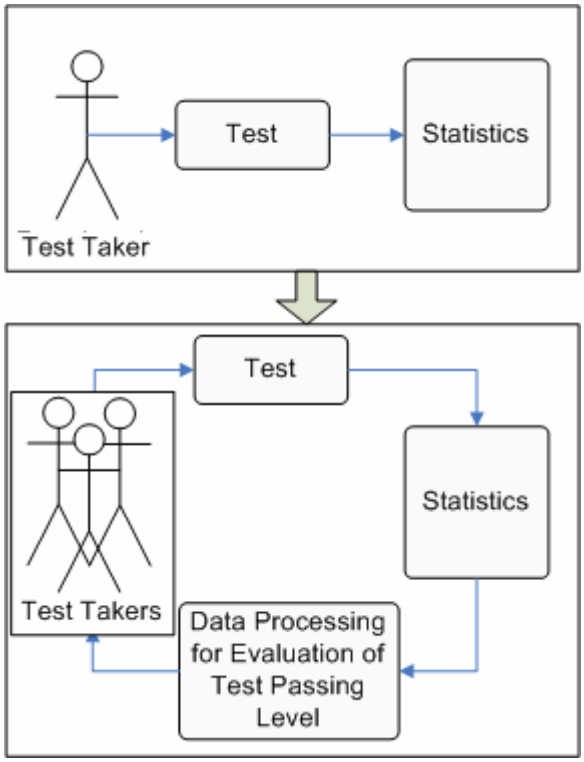

Figure 1. Use of statistical data

In the section "Analysis of impact parameters" it is shown that statistical data of an existing testing system, such as data on tests, test takers and their test results, is possible to be used not only for analysis of statistical data but also for testing other people. It is demonstrated that certain test taker's parameters have impact on passing/failure of the test. Therefore, instead of typical analysis of statistical data it is suggested to use test taker parameters for setting a new possible test taker's passing level as shown in Figure 1.

The section "Data clustering" describes transformation of multidimensional data (statistical testing data) into two-dimensional space and their display on the plane, i.e. data processing and visualization method using self-organizing neural networks (SOM). SOM maps can be used to visually represent data clusters and express projections of multidimensional data into space of smaller dimension, usually the plane (Haykin, 1999; Dzemyda, Kurasova, Žilinskas, 2008; Verikas, Gelžinis, 2003). This section also carries out realization analysis of self-organizing neural networks. For data processing Viscovery® SOMine® programme was chosen.

Thus after evaluation of the conclusions of the first part and requirements for computer literacy testing systems, parts of possible computer literacy test design process are presented supplementing the standard test process with two parts: estimation of the initial level of skills and test taker describing parameters. The testing process should be adaptive so that test forms were as equivalent as possible. It is proved that statistical testing data can be used to estimate the initial skills level of test taker. It is also proved that for processing and visualization of statistical data self-organizing neural networks can be used.

Part 3. Computer literacy testing process model

The third part of the thesis is devoted to the computer-based computer literacy testing process model. In accordance with the conclusions of the research carried out in the previous parts a simplified computer literacy testing process model is being designed which is shown in Figure 2.

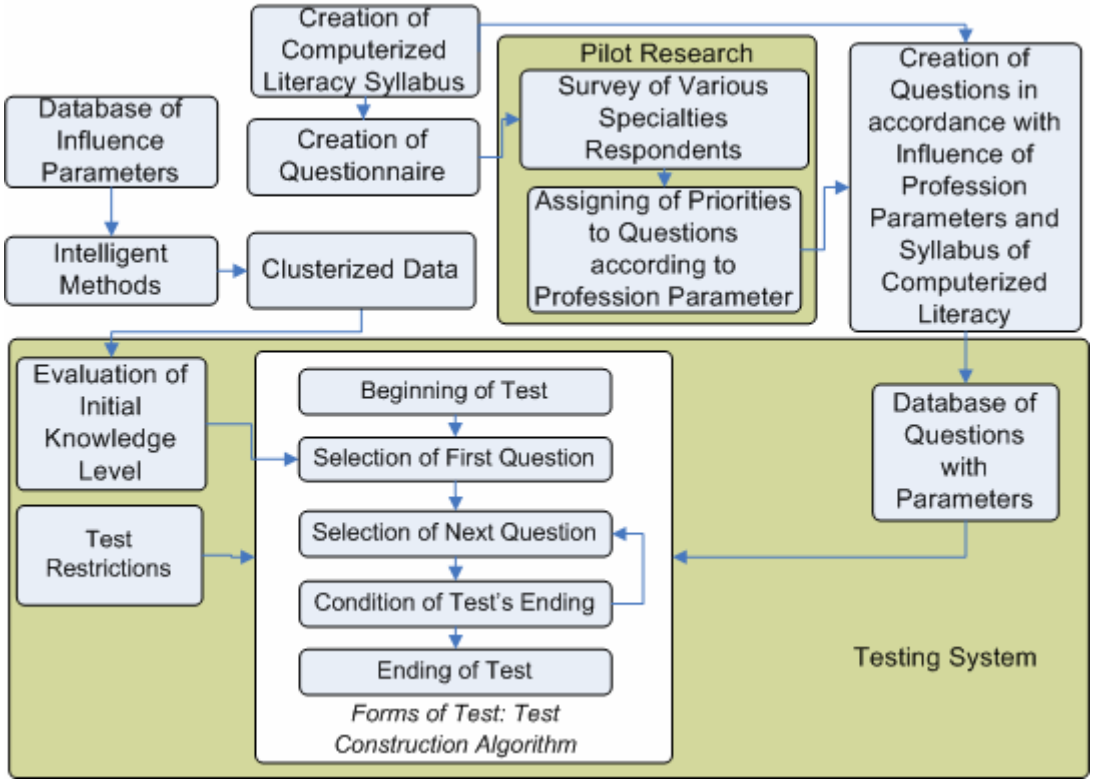

Figure 2. Model of computer literacy testing process

In this model, first of all, data are collected in two stages. During the first stage statistical data are clustered (on their basis a test taker's initial level of knowledge will be estimated), and during the second stage a survey about people's of different occupations preferable ways to perform activities on the computer is conducted (on that data test forms will be created). With processed statistical data and survey data ready the test will be designed using test construction restrictions.

Before designing test questions, first requirements for test questions are defined. Since the test is designed considering parameters of programme versions and difficulty, each item of computer literacy questionnaire syllabus should have at least three questions with different levels of difficulty. And each question with a different level of

difficulty can also be presented for different programme versions. Figure 3 shows the structure of the programme and coherence of questions.

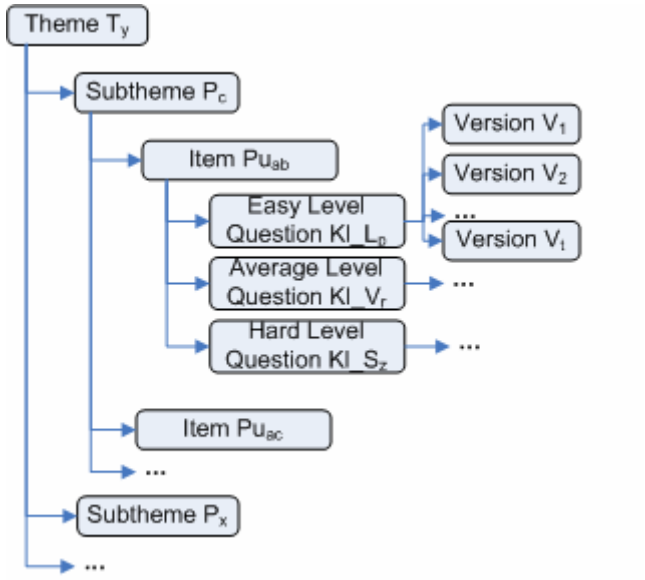

Figure 3. Structure of the syllabus and questions

Having formulated questions, test forms are made where questions should be selected according to the data of the survey using statistical (level of difficulty, item exposure rate), version (some questions can be presented using different programme versions and the test taker can choose which to use for the test), and test taker's (occupation, gender, age, experience parameters) data. The formation of question record scheme is shown in Figure 4.

Figure 4. Structure of question

Thus each question should have not only its qualifying parameters, but also test taker identifying parameters:

Q_record = [Kl, Ta, S, V, Pd, Tp, Φ(Pr, M, L, Pa)],
where Q_record is the question record in the database,
Kl is the question,
Ta is the correct answer,
S is the level of difficulty,
V is the programme version used for the test,
Pd is question exposure rate,
Tp is the content restricting parameter,

Pr is the occupation identifier,
M is the test taker's age interval,
L is the test taker's gender,
Pa is the test taker's work experience with the computer.

In order to formulate the test compactly, it is suggested to present questions similarly as in adaptive tests, however, starting not with a question of medium difficulty, but with a question set by the programme according to the test taker's data given prior to the test. During the test if the question is answered correctly, a more difficult question is presented; and if the question is answered incorrectly, an easier question is given (see Figure 5).

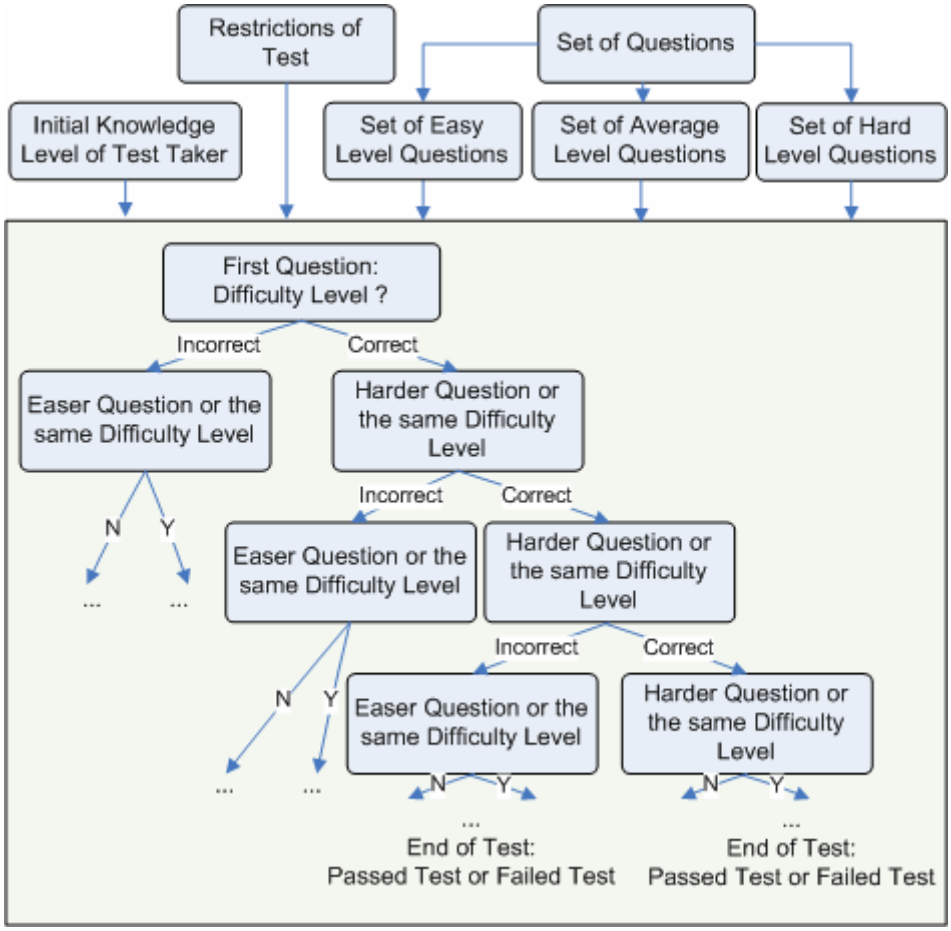

Figure 5. Process of question formation

In principle, during the whole test, test takers should be presented with equal number of questions of all levels of difficulty. However, in this case the test can be of variable length if the test taker answers a certain number of questions correctly and collects the minimum number of points. In such a case the test can be terminated producing the test result of passing. The test can also be finished (terminated) in case the test taker does not answer a certain number of questions (i.e. exceeds the number of points of unanswered questions); in that case there is no point in continuing the test and it is terminated producing the test result of failure.

Knowing the test taker's initial level of skills, with ready questions and test restrictions the test formation process is begun. Depending on the level of skills the first

test question of corresponding difficulty is chosen. If the test taker answers correctly, he is given a more difficult question or a question of the same difficulty (if there are no more difficult questions). If the test taker answers incorrectly, he is given an easier question or a question of the same difficulty (if there are no easier questions). The process is continued until the condition of the end of the test is reached, which can be: the test is passed if the limit of the test end is achieved; the test is passed (or failed) if the maximum number of possible questions has been given; and the test is failed if the test failure limit is reached. After each given question it is checked if the minimum number of points has been collected (in case of correct answers), and if there is still a chance to collect the minimum number of points (in case of incorrect answers). The algorithm of the newly designed computer literacy test design method looks like this:

```
Parameters provided by Test taker: Version_of_program, Profession, Age, Gender,
County
[Amount_of_Questions, Amount_of_difficulty_Levels] <- LoadData(MYSQL_DB)
Limits_of_Question_Difficulties =
CalculateLimitsOfQuestions(Amount_of_Questions, Amount_of_difficulty_Levels)
Min_Amount_of_Points =
CalculateMinAmountOfPoints(Limits_of_Question_Difficulties)
Allowed_Amount_of_lost_Points =
CalculateAmountOfAllowedLostPoints(Limits_of_Question_Difficulties)
Scored_Points = 0
Lost_Points = 0
Knowledge_Level = CalculateKnowledgeLevelUsingStatisticalData()
Difficulty_Level_of_Next_Question = GetInnitialDifficulty(Knowledge_Level)
Questions_Form = GetQuestionsFormByProfession(Profession, Age, Gender,
County, Knowledge_Level)

  While (Scored_Points <Min_Amount_of_Points) AND (Lost_Points
    <= Allowed_Amount_of_lost_Points) do
    Question_Difficulty = Difficulty_Level_of_Next_Question
    Subtheme = GetUnusedSubtheme(Question_Difficulty)
    Question_id = GetQuestion(Version_of_program, Subtheme, Questions_Form)
    Number_of_Answer =  GetAnswerOfTestTaker(Question_id)
      If (IsAnswerCorrect(Number_of_Answer)) Then
      Scored_Points  += GetWeightOfQuestion(Question_id)
      Else
      Lost_Points += GetWeightOfQuestion (Question_id)
      EndIf
      Difficulty_Level_of_Next_Question =
      CalculateDifficultyOfQuestion([Limits_of_Difficulty])
  EndWhile
```

The features of computer literacy test design method

1. The computer literacy test design method is intended for such testing systems that are used for testing people's computer skills.

2. The computer literacy test design method differs from other methods in the following aspects:

- prior to the test the test taker's initial level of knowledge is estimated,

- the test form is generated according to the test taker's occupation,
- test questions are selected according to the level of difficulty regarding the answer to the previous question.

3. Under certain circumstances the computer literacy test design method allows to decrease the number of questions in the test.

4. The use of the computer literacy test design method increases test security and it can be claimed that questions are generated for each individual test taker.

5. Test takers get test forms of the same difficulty.

6. During the test, parameters, such as statistical, version, test taker's occupation and others are estimated.

Part 4. Experimental-practical application of the computer literacy test design method

In this part analysis of questionnaire and statistical data is given, as well as the developed prototype is described. Experimentally researched and generalized research results are presented, the work results are summed up and further possibilities of application of computer literacy testing system are discussed.

In the section „Clustering of ECDL test takers' data" research of application of statistical data for the estimation of test takers' initial level of knowledge is carried out. For this purpose the following parameters are used: test taker's age, gender, county, place of work, post, module name, module version, the number of questions presented, the number of correct answers, the length of the test, date and the number of tests one person has taken. The experiment is carried out using statistical information of one ECDL module (the module "Use of computer and handling files"). First statistical data are prepared for clustering, which is done with the help of the SOMine programme (see Section 2.4). The data are entered into the table where each test taker is defined by certain parameters that may influence the passing level of the test (see Section 2.3). The parameters (in this case gender, county, place of work, job, answering level, test level and the code of module version) with non-numerical values will be coded. This is necessary to do because neurons (test takers) are classified according to their features on the two-dimensional map, which forms regions interlinked according to their values (each group of neurons finds itself next to the group with similar features).

After that statistical data are processed using SOMine programme and self-organizing diagrams are created (real commercial data from the ECDLTE testing system are used, more than 40000 records). In this case the programme singles out four main clusters (see Figure 6). Cluster A means the set of those who failed, cluster B means the set of those who passed with no mistakes, clusters C and D mean the set of those who passed, cluster C represents data of those people who passed the test faster, whereas cluster D means people who passed the test much more slowly.

Figure 7 shows detailed diagrams, where colour markings indicate which cluster a test taker gets into according to a certain parameter.

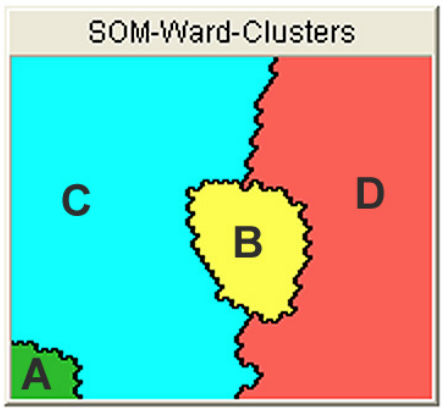

Figure 6. Main clusters

At the bottom of each diagram there is a scale. For example, in the diagram "gender" values can be 1 (female), represented by the red colour, or 0 (male), represented by the blue colour.

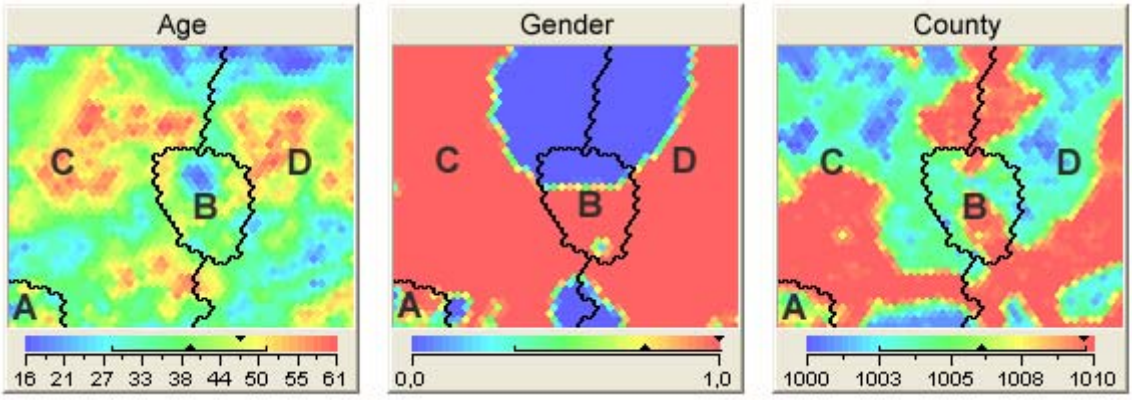

Figure 7. Clustered data[2]

Having the test taker's initial data and using cluster diagrams, it is possible to deduce how a person tends to do tests and try to predict his initial level of knowledge and on the base of that to choose the first question.

Table 1. Test takers' data for clustering

Age	Gender	County	Workplace	Job	Date	Module	Module Code
38	1	1010	1275	19	2009	2	144
30	1	1005	400	8	2009	2	143
36	1	1009	1666	17	2009	2	144
50	1	1004	250	8	2009	2	139
34	0	1000	33	10	2009	2	139
37	1	1010	54	12	2009	2	140
40	1	1005	75	14	2009	2	139
43	1	1009	96	15	2009	2	140

For instance, Table 1 presents data about new test takers. Each person defining parameter is coded, e.g., gender: number 0 - male, 1 – female, 1004 means Panevėžys county, module code 144 corresponds the code of the Lithuanian test version of module 2 Windows XP.

[2] Full diagram is presented in the dissertation

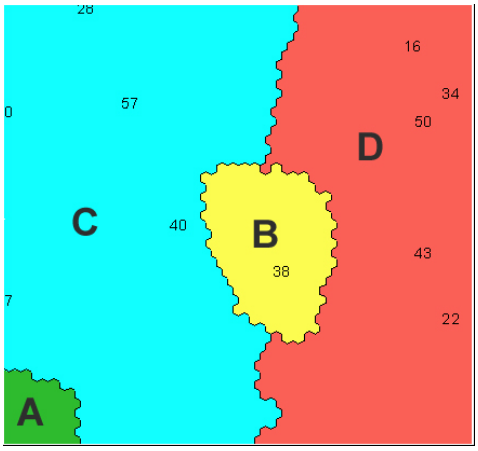

Figure 8. Clustering of new test takers

After importing the data from Table 1 into the SOMine programme, it is possible to know at once which cluster data of each new test taker got into and this allows to choose the first question. For example, the one who got represented into cluster B (test takers passing the test without mistakes) can be given a more difficult question; the one who got represented into cluster C (people who tend to do the test faster but are allowed to make a certain number of mistakes) can be given a medium difficulty question; and the one who got represented into cluster D can be given an easy question since people who get represented into this cluster tend to make several mistakes and do the test longer than the average test taking time. Those test takers that get represented into cluster A can be given recommendations according to which people represented in this cluster statistically make mistakes most frequently and advised to do a self-check test prior to taking the test (see Figure 8).

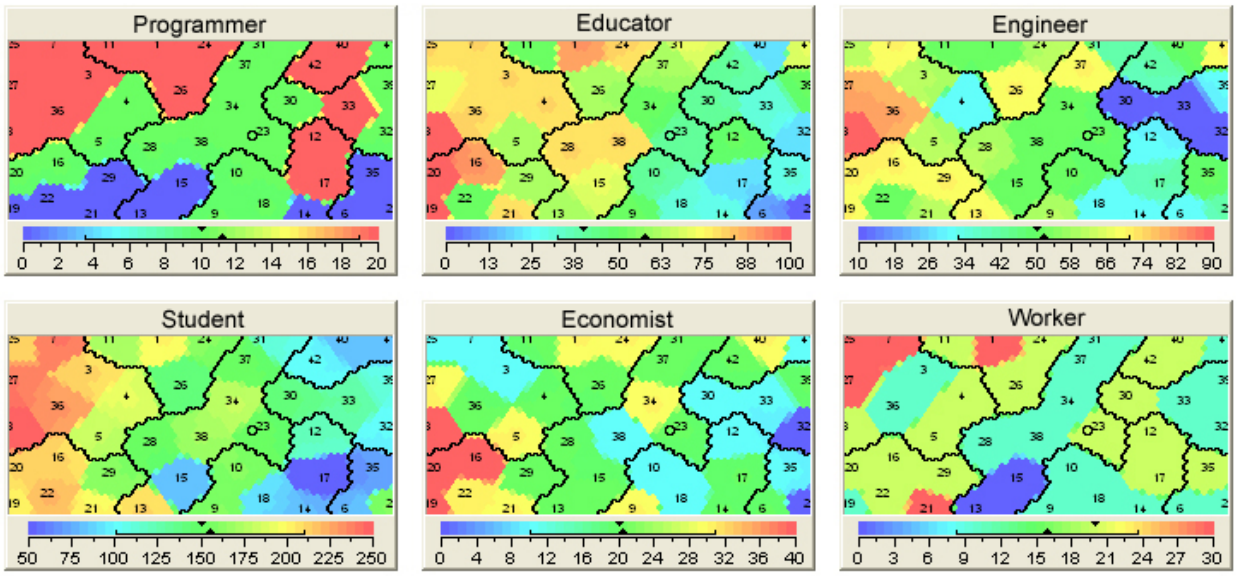

Figure 9. Distribution of computer activities performed by certain occupations according to frequency

The section „Results of the pilot research" describes the research of activities with the computer performed by people of different occupations. During the survey before

answering questions the respondents gave short personal information: occupation, gender, age, job and person assessment of computer skills.

After the first part of the survey it turned out that the majority of teachers, office workers and economists seldom or never use more sophisticated, unordinary operations, such as tools of the control panel, setting monitor options, review of folder trees, programme deleting, disk formatting, etc.

With the help of the SOMine programme the results of the most frequent answer by the survey respondents from each occupation were clustered (see Figure 9). By studying each diagram of the figure separately, it is possible to learn how differently a representative of each occupation performs particular actions. It can be seen from different diagrams how many participants were questioned and the number of people who answered a certain question.

On the base of the research results for each answer of the survey question (in the answers there were given all possible ways of task performing) a test question with four options was created.

The section „Results of experimental testing" describes the results of the experimental testing. The testing was performed using traditional method and the newly designed testing method. Using the traditional testing when questions are presented in the random order, the calibration of the question base was performed. The value is calculated using the formula:

$$P = p1 + p2 + ... + p36 \qquad (1)$$

where

$$p_n = \begin{cases} 0, \textit{ if answered correctly} \\ 1, \textit{ if answered incorrectly} \end{cases}$$

By using the test design method suggested in this work, the value is calculated using the formula:

$$P = S + V + L \qquad (2)$$

where $S = \sum k_s \cdot p_n$, is the sum of points of difficult questions which can vary from 0 to 30;

k_s =5, is the coefficient of difficult questions; i = 1 … 6;

$V = \sum k_v \cdot p_m$, is the sum of points of medium-difficulty questions which can vary from 1 to 18;

k_v = 3, is the coefficient of medium-difficulty questions; j = 1 … 6;

$L = \sum k_l \cdot p_z$, is the sum of points of easy questions which can vary from 1 to 6;

k_l = 1, is the coefficient of easy questions; r = 1 …6;

The use of IRT (Item Response Theory) enables to find out the probability of the correct answer to the question which is defined as the difference between b and Θ. The function of the simplest IRT model uses only these two elements. As each question is defined only by one parameter, that of difficulty (b), the model is called one-parameter model or 1-PL. The model was developed and popularized by a Danish mathematician Georg Rasch (1901-1980) and very often it is called Rasch model.

$$P(\Theta) = \frac{1}{1 + e^{-(\Theta - b)}} \qquad (3)$$

where P(Θ) is the probability that the test taker with Θ abilities will answer correctly the question whose difficulty is b (Wainer H. et al., 2007, Partchec I. 2004, Hambleton R. K, 1991; McDonald, 1999).

Using the formula 3 the worst, the best and the average test performance scenarios were simulated (see Figure 10). At best, the test taker answered an average-level question incorrectly and lost 3 points out of possible 54, his initial level of knowledge was the highest. At worst, the test taker answered 3 questions (average, easy and difficult) incorrectly and lost 9 points, but this was enough to pass the test. The test taker's initial level of knowledge was average, but because of the mistake, he had to answer the maximum number of questions (in this case 18). In the average case the test taker did not answer 5 questions (1 average and 4 easy questions) and lost 9 points. The test taker's initial level of knowledge was the lowest.

Research of tests results showed, that if more than 16 questions were presented, correct answer probability after 16th question remains almost the same.

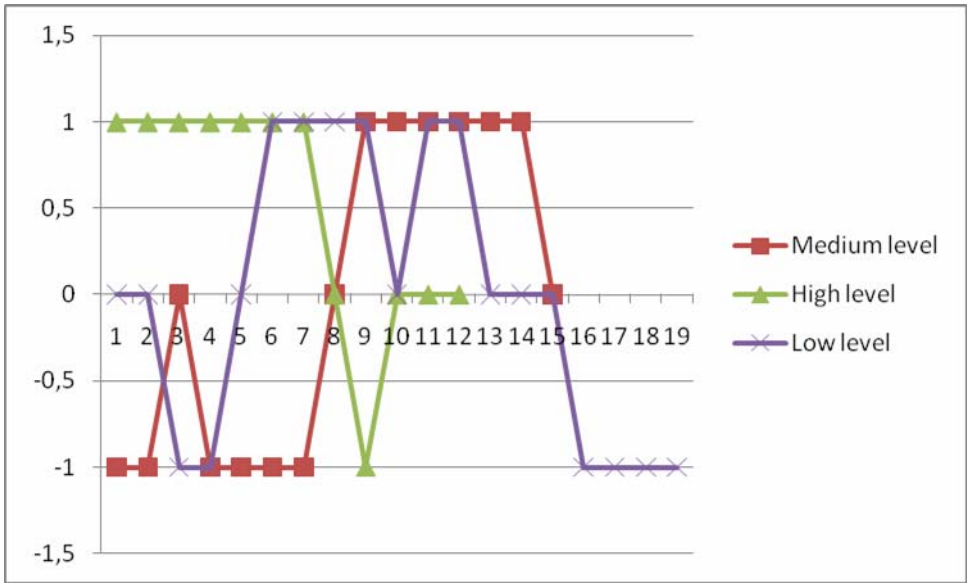

Figure 10. The best, the average and the worst test performance scenarios

Summing up, it can be stated that the shortening of the test up to a certain number does not influence assessment. This is also proved in literature by conclusions made by other scientists, namely, that further questions do not influence the value of assessment after answering to 13–15 test questions (Al-A'alli M., 2007, Gouli E. et al., 2001, Eggen & Straetmans, 2000).

Therefore, during the suggested question formation the test is being formulated individually for each test taker according to how he answered previous questions. Since the test is adaptive, the length of the test can be variable. In this way testing time is economized and the number of question exposures is minimized, which decreases possibility of cheating and question leak.

CONCLUSIONS

The review of scientific literature, analysis of research methods and application of computer literacy test design method and the research carried out supported the **propositions maintained in the dissertation**.

1. For high-level testing systems there is the need to develop new theories and methods that would enable to better assess test takers‘ knowledge.
2. Both adaptive and non-adaptive test design methods are used to assess general knowledge in a certain area; however, neither of the methods provides tools that would help to assess and later apply test takers‘ initial skill level of computer literacy.
3. After performing the analysis of typical computer literacy high-level testing systems, it was observed:
   a. those testing systems use nearly same types of questions;
   b. test takers‘ initial level of knowledge is not assessed;
   c. they use traditional test design methods.
4. The test takers‘ initial level of knowledge should be estimated prior to the test and for this purpose statistical testing data have to be used that are processed and visualized using self-organizing neural networks.
5. The developed test design method has the features of automated test assembly as well as those of adaptive test design:
   a. frequency of question exposure is controlled and test takers are given test forms of the same difficulty;
   b. the test of variable length is compactly formed for each test taker individually regarding the occupational parameter.
6. Considering the results of the experiment it was found out that:
   a. the way of the work with the computer depends on the occupation; therefore, each question can be given occupational weights. These weights help to select questions for a test taker based on characteristic computer skills of people of different occupations;
   b. clustering of statistical data using self-organizing neural networks helped to estimate test taker‘s initial computer literacy skills level;
   c. shortening of an adaptively constructed test up to certain limit does not have a significant influence on the final test results.
7. The proposed test design method ensures sufficient efficiency performing computer literacy testing considering the main necessary and sufficient requirements.
8. The designed testing system has a practical value of performing testing of computer literacy skills. The system prototype was implemented at Information Technologies Institute and used for pilot testing.
9. It is recommended to use the new developed test design method for possible implementation in the new version of ECDL test system.

Renata Danielienė

Vilniaus universitetas

DAKTARO DISERTACIJOS SANTRAUKA

Temos aktualumas. Žinių testavimas rankiniu būdu, kai testuojamieji klausimus gauna popieriuje ir jame pateikia savo atsakymus, jau seniai yra morališkai pasenęs. Šiuo metu įvairių sričių žinioms tikrinti yra naudojamas kompiuterizuotas testavimas. Iš pradžių buvo naudojami paprasti kompiuterizuoti fiksuoto ilgio testai. Vėliau, atsiradus poreikiui apsaugoti klausimus, efektyviau įvertinti testuojamųjų žinias, generuoti klausimus kiekvienam individualiai, kompiuterizuotas testavimas buvo tobulinamas siekiant išspręsti nuolat kylančias problemas. Taip atsirado ne vienas klausimų konstravimo metodas.

Pastaruoju metu žinių tikrinimas kompiuterizuotais testais yra plačiai naudojamas. Netgi mokymo įstaigose kai kurie egzaminai yra keičiami testais. Taigi tokie testavimai turi būti atliekami kokybiškai. Didžiausia problema kyla tuomet, kai reikia testavimo sistemai parinkti tinkamiausią testų konstravimo metodą. Ypač svarbu atlikti kokybišką testavimą specializuotose testavimo sistemose, pavyzdžiui, testuojant žmonių kompiuterinio raštingumo žinias. Dėl to šiame darbe bus nagrinėjami literatūroje aprašomi pagrindiniai testų konstravimo metodai, išskiriamos jų pagrindinės savybės ir situacijos, kada kurį metodą naudoti. Atlikus šių metodų analizę, kompiuteriniam raštingumui testuoti bus pasiūlytas sprendimas, kurį naudojant pagal žmonių profesijas ir veiklos sritį teste būtų kompaktiškai formuojami klausimai, tuo pačiu sumažinant klausimų parodymų dažnį. Naujas kompiuterinio raštingumo testų konstravimo metodas siūlomas dėl to, kad specifinio klausimų atrinkimo būdo, būtent testuojant skirtingų profesijų žmones, turinčius skirtingus įgūdžius šioje srityje, nėra sukurta. Paprastai testų klausimai yra parenkami fiksuota, atsitiktine tvarka arba adaptyviai, įvertinus patikimumą, testo ilgį, klausimų banko dydį, testuojamųjų kiekį ir pan. Siūlomame metode žinių lygiui nustatyti gali padėti kintamieji, kurie dar nebuvo vertinami testuojant, t. y. profesija, išsilavinimas, amžius, darbo kompiuteriu patirtis ir pan. Šiems kintamiesiems (kiekybiniams ir kokybiniams) tirti taikomi intelektiniai metodai.

Tradiciniai testų konstravimo būdai nepateikia priemonių, kaip testuoti būtent kompiuterinio raštingumo įgūdžius, todėl ir buvo nuspręsta atlikti intelektualaus kompiuterinio raštingumo testų konstravimo metodo tyrimą. Dažniausiai žemesnės ar vidutinės svarbos testavimo sistemose klausimams atrinkti yra naudojama atsitiktinės atrankos funkcija. Kartais tokio lygio sistemose yra įvertinamas klausimų persidengimas ar įvairūs apribojimai. Aukšto lygio testavimo sistemose dažniausiai naudojami adaptyvūs testai. Atliekant tokius testus klausimai yra konstruojami pagal tai, kaip testuojamasis atsakinėjo į ankstesnius klausimus, įvertinus klausimo sunkumą. Taip pat gali būti atsižvelgiama ir į diskriminacijos ir (arba) į spėjimo faktorius.

Adaptyvūs testai iš pradžių buvo naudojami nustatant žmogaus intelekto lygį. Pirmąjį tokį testą 1905 m. sukūrė A. Binett, ir šis testas buvo pateikiamas popieriniu variantu. Vėliau adaptyvūs testai buvo pradėti naudoti testuojant žinias ir kitose srityse (Reckase, 1989), papildant šį modelį sunkumo, diskriminacijos ir spėjimo faktoriais, kurie yra nustatomi naudojant užduočių sprendimo teoriją IRT – *Item Response Theory* (Baker, 2001; Embretson ir Reise, 2000).

Testuojant kompiuterinio raštingumo žinias nepakanka testus konstruoti naudojant atsitiktinės klausimų atrankos funkciją ar pagal tam tikrus apribojimus. Prieš kompiuterinio raštingumo žinių tikrinimą tikslinga išsiaiškinti esamą testuojamojo žinių lygį, kurį būtų galima sužinoti įvertinus tam tikrą informaciją apie testuojamąjį. Mokslininkai, tyrinėjantys adaptyvius testus, šią problemą bando spręsti prieš testavimą apmokydami testuojamuosius, pateikdami jiems žinių tikrinimo testus ir pan. Tokius bandymus savo straipsniuose aprašo Mansoor Al-A'ali (2007a; 2007b).

Taigi, tiek adaptyvūs, tiek neadaptyvūs testai yra naudojami norint testuoti tam tikros srities bendrojo išsilavinimo žinias. Tačiau ir vienu, ir kitu atveju iš pradžių nėra žinomas pradinis testuojamojo žinių lygis. Dėl to būtina ištirti egzistuojančius testų konstravimo metodus, apibūdinti jų savybes ir pasiūlyti testų konstravimo metodą, skirtą testuoti kompiuterinio raštingumo įgūdžius.

Problemos ištyrimo lygmuo. Testų konstravimo tema imta plačiai domėtis atsiradus kompiuteriams. Iki tol egzistavo paprasti popieriniai testai. Netgi pirmieji adaptyvūs testai buvo sukurti ir naudojami popieriniu variantu – tai buvo A. Bineto testai (Reckase, 1989). Kompiuterinį testavimą tyrinėję informatikos srities specialistai siūlė naujus metodus, leidžiančius atlikti testus tiek žemo, tiek aukšto lygio testavimo sistemose (Bartram, Hambleton, 2006; Danielienė, Telešius, 2009; Luecht, 2006; Wilson, Engelhard, 2000). Adaptyvių testų sudarymo principus nagrinėjo Cyntia G. Parshall (2000, 2002), Wim J. Van der Linden (2000, 2004, 2006), Howard Wainer (1987, 2000a, 2000b, 2007), David Thissen, Robert J. Mislevy (1990), Al-A'ali, Mansoor (2007a, 2007b), Thompson, N. A. (2006, 2007, 2008) ir kiti.

Remiantis statistiniais metodais buvo sukurti adaptyvūs testai, kuriuose yra naudojama užduočių sprendimo teorija (*Item Response Theory* – IRT). IRT yra bendra statistikos teorija apie klausimus testuojamiesiems ir testo savybes, bei kaip tos savybės susijusios su gebėjimais, kurie yra įvertinami teste pateikiant klausimus (Hambleton et al., 1991; Harvey, Hammer, 1999; Partchec, 2004; Fraley et al., 2000; Wiberg, 2004).

Neadaptyvių testų sudarymo principus tyrinėjo mokslininkai Tim Davey (1995), Nathan A. Thompson (2008), Lillian C. Folk, J. Zachary March, Robin D. Hurst (37) ir kiti. Tokiose testavimo sistemose dažniausiai dėmesys skiriamas tam, kad klausimai testų formose kuo mažiau persidengtų, taip pat į klausimų atrinkimo tvarką teste arba klausimų atrinkimo apribojimus.

Kompiuterinio raštingumo srityje tiek Lietuvoje, tiek kitose šalyse buvo vykdomi įvairūs kompiuterinio raštingumo empiriniai tyrimai, kuriuos atliko ekspertai, tokie kaip Eugenijus Telešius, Alfredas Otas, Raimundas Jasinevičius ir kt. (2003, 2004, 2005, 2006, 2007), Lina Markauskaitė ir Valentina Dagienė (2004), Diana Šaparnienė (2003), Renata Danielienė (2006). Tokius tyrimus vykdė ir įvairios mokslininkų grupės, pavyzdžiui, Cedefop (2004), LR švietimo ir mokslo ministerijos (2004), Švietimo informacinių technologijų centro (2003) ir KTU Kompiuterinio raštingumo centro (2004, 2005, 2006, 2007), Europos Komisijos IKT įgūdžių monitoringo grupės (2002) ir pan.

Darbo objektas – testų konstravimo metodai.

Darbo tikslas – sukurti intelektualų testų konstravimo metodą, kurį būtų galima pritaikyti testuojant žmonių kompiuterinio raštingumo įgūdžius, įvertinant jų turimų įgūdžių parametrus.

Siekiant šio tikslo, atliekami tokie teoriniai ir praktiniai uždaviniai:

1. Išanalizuoti egzistuojančius testų konstravimo metodus ir apibendrinti šios analizės rezultatus.
2. Ištirti aukšto lygio kompiuterinio raštingumo testavimo sistemose naudojamus testų konstravimo būdus.
3. Pasiūlyti priemones nustatyti pradinį testuojamųjų kompiuterinių įgūdžių lygį.
4. Sukurti kompiuterinio raštingumo testų konstravimo metodą, kuris pagal testuojamojo profesiją ir veiklos parametrus kompaktiškai formuotų testų klausimus.
5. Atlikti įvairių profesijų žmonių darbo su asmeniniu kompiuteriu įpročių tyrimą, siekiant nustatyti kompiuterinių įgūdžių priklausomybę nuo testuojamojo profesijos.
6. Sukurti testavimo sistemos klausimų konstravimo prototipinę realizaciją ir atlikti eksperimentinį tyrimą.

Tyrimas buvo atliekamas 2005–2010 metais.

Tyrimo metodai: teorinė analizė, grįsta įvairių mokslininkų darbų rezultatais ir išvadomis; sisteminė analizė; modeliavimas; įvertinimas; apibendrinimas. Statistiniams duomenims klasterizuoti, siekiant nustatyti pradinį testuojamųjų žinių lygį, pasirinkta saviorganizuojančių neuroninių tinklų apdorojimo taikomoji programa *Viscovery® SOMine®*. Tyrimų rezultatai apdoroti naudojant skaičiuoklę *Microsoft Excel*.

Darbo struktūra

Pirmoje dalyje („Kompiuterinis testavimas ir testų konstravimo metodų tyrimas“) apžvelgiamas kompiuterinis testavimas ir vertinimas, analizuojamos tipinės kompiuterinio raštingumo testavimo sistemos, aprašomi kompiuterinių testų konstravimo metodai. Analizuojant šiuos metodus, išskiriamos pagrindinės savybės, kokia teorija kiekvienas iš jų remiasi ir kokios įtakos tai turi kitiems svarbiems faktoriams, tokiems kaip vertinimo charakteristikos, patikimumas, saugumas, testo ilgis ir testuojamųjų skaičius, klausimų banko charakteristikos ir kt. Pateikiamas aprašytų testų konstravimo metodų lyginimas. Šioje dalyje taip pat išanalizuotos tipinės kompiuterinio raštingumo testavimo sistemos: išnagrinėta jų struktūra, savybės ir funkcijos, nurodyti privalumai bei trūkumai.

Antroje dalyje („Kompiuterinio raštingumo testavimo sistemų principų ir įtakos parametrų analizė“) aprašomi reikalavimai kompiuterinio raštingumo testavimo sistemoms, pagrindžiamas poreikis kurti naują kompiuterinio raštingumo testų konstravimo metodą ir, remiantis pirmo skyriaus išvadomis, pateikiamos naujų intelektualių metodų taikymo galimybės kompiuterinio raštingumo testams. Nagrinėjama, kokie nuo testuotojo priklausantys parametrai gali turėti įtakos testo rezultatams, ir pateikiamas siūlymas panaudoti statistinius duomenis nustatant pradinį testuojamojo įgūdžių lygį. Statistiniams duomenims apdoroti siūloma naudoti dirbtinius neuroninius tinklus.

Trečioje dalyje („Kompiuterinio raštingumo testavimo proceso modelis“) aprašomas siūlomas kompiuterinio raštingumo testų konstravimo metodas. Naudojant pirmo ir antro skyrių išvadas siūloma remtis adaptyvaus testų sudarymo (CAT) ir kompiuterinio automatizuoto testų surinkimo (ATA) metodais. Testų konstravimo metodas tobulinamas pridedant papildomus testuojamųjų profesijos, amžiaus ir kitus parametrus. Taikant tokį testų konstravimą, klausimai dar labiau adaptuojami testuojamiesiems, ir dėl to galima tiksliau įvertinti jų įgūdžius. Naudojant naują testų

konstravimo metodą, iš pradžių nustatomas pradinis testuojamojo žinių lygis pagal statistinius lietuviškos ECDLTE testavimo sistemos duomenis. Naudojant žmonių įgūdžių apklausos duomenis, konstruojamos testų formos.

Ketvirtoje dalyje („Kompiuterinio raštingumo testų konstravimo metodo eksperimentinis praktinis panaudojimas“) aprašoma, kaip teorinėje dalyje pateiktus sprendimus galima pritaikyti tipinėje kompiuterinio raštingumo testavimo sistemoje. Čia pateikiama anketinių ir statistinių duomenų analizė (analizei buvo naudojami realios ECDLTE testavimo sistemos duomenys apie 40000 įrašų), aprašomas sukurtas prototipas. Pristatomi apibendrinti eksperimento tyrimo rezultatai ir aptariamos tolesnės kompiuterinio raštingumo testavimo sistemos plėtros galimybės.

Darbe ginami teiginiai

- Kompaktiškam ir racionaliam kompiuterinio raštingumo testo konstravimui tikslinga naudoti testuojamųjų asmenų profesinius ir veiklos parametrus.
- Tikslinga testuojamuosius klasifikuoti naudojant dirbtinius neuroninius tinklus, siekiant nustatyti kompiuterinių įgūdžių priklausomybę nuo testuojamojo profesijos.
- Pasiūlytas testų konstravimo metodas (įdiegtas ECDLTE testavimo sistemoje) leidžia efektyviau atlikti kompiuterinio raštingumo testus nepakenkiant testo pilnumui ir testo kokybės kriterijui.

Darbo mokslinis naujumas

Apjungus esamų kompiuterinių testų konstravimo metodų savybes ir patirtį sukurtas kompiuterinio raštingumo testų intelektualaus konstravimo metodas. Remiantis šiuo metodu naudojant saviorganizuojančius neuroninius tinklus pirmiausia įvertinamas galimas testuojamojo įgūdžių lygis, ir pagal šiuos duomenis testuojamajam intelektualiai parenkami klausimai. Nuo neadaptyvių testų metodų šis metodas skiriasi tuo, kad testas yra konstruojamas naudojant testuojamojo parametrus, tuo pačiu klausimai testuojamiesiems formuojami kompaktiškiau nepakenkiant testo pilnumui ir testo kokybės kriterijui. Kompaktiškumas čia suprantamas kaip sudėtinė sąvoka: testo trumpinimas atsižvelgiant į testo baigimo sąlygas ir klausimų parinkimas atsižvelgiant į testuojamojo profesijos ir veiklos parametrus. Testo trumpinimas duoda ekonominę naudą, o testuojamiesiems orientuotas klausimų parinkimas suteikia jiems psichologinį komfortą.

Praktinė darbo nauda

Gauti tyrimų rezultatai gali būti naudojami testuojant kompiuterinius žmonių įgūdžius mokymo įstaigose (mokyklose, kolegijose, universitetuose ir pan.), įvairiuose mokymo ir testavimo centruose ar kvalifikacijos kėlimo sistemoje. Disertacijoje suformuotas testų konstravimo metodas buvo įdiegtas viešojoje įstaigoje „Informacinių technologijų institutas“ ECDL testavimo sistemoje.

IŠVADOS

Mokslinės literatūros apžvalga, tyrimo metodų analizė ir kompiuterinio raštingumo testų konstravimo metodo pritaikymas ir atlikti tyrimai patvirtino disertacijoje ginamus teiginius.

1. Realizuojant aukšto lygio testavimo sistemas yra būtina kurti naujas teorijas ir metodus, kurie leistų tiksliau įvertinti testuojamųjų žinias.
2. Tiek adaptyvūs, tiek neadaptyvūs testų konstravimo metodai yra naudojami įvertinant testuojamųjų žinias iš įvairių probleminių sričių, tačiau nei vienas nagrinėjamas testų konstravimo metodas nepateikia priemonių, kurios padėtų įvertinti ir vėliau panaudoti pradinį kompiuterinio raštingumo lygį.
3. Atlikus tipinių aukšto lygio kompiuterinio raštingumo testavimo sistemų analizę pastebėta:
   a. yra naudojami daugiausia tie patys klausimų tipai;
   b. nėra vertinamas pradinis testuojamųjų žinių lygis;
   c. jose yra naudojami tradiciniai testų konstravimo metodai.
4. Prieš testą turi būti įvertinamas pradinis testuojamojo žinių lygis, ir jam nustatyti tinkami statistiniai testavimo duomenys, kuriems apdoroti ir vizualizuoti naudojami saviorganizuojantys neuroniniai tinklai.
5. Sukurtas testų konstravimo metodas turi automatizuoto testavimo bei adaptyvaus testų konstravimo ypatybių:
   a. kontroliuojamas klausimų pasirodymo dažnis bei testuojamiesiems pateikiamos vienodo sunkumo testų formos;
   b. kintamo ilgio testas kompaktiškai formuojamas kiekvienam testuojamajam individualiai atsižvelgus į jo profesijos ir veiklos parametrus.
6. Atlikus eksperimentą nustatyta, kad:
   a. kompiuteriniai įgūdžiai priklauso nuo profesijos, dėl to kiekvienam testo klausimui galima suteikti profesijos svorius. Šie svoriai padeda testuojamajam parinkti klausimus pagal jo profesijai būdingą veiksmų atlikimo būdą;
   b. statistinių duomenų klasterizavimas naudojant saviorganizuojančius neuroninius tinklus padeda nustatyti pradinį testuojamojo kompiuterinių įgūdžių lygį;
   c. adaptyviai formuojamo testo trumpinimas iki tam tikros ribos neturi esminės įtakos galutiniams testo rezultatams.
7. Siūlomas testų konstravimo metodas užtikrina pakankamą efektyvumą atliekant kompiuterinio raštingumo testavimą, įvertinant pagrindinius būtinus ir pakankamus reikalavimus.
8. Sukurta testavimo sistema turi praktinę vertę, atliekant kompiuterinio raštingumo įgūdžių testavimą. Tai įrodo pasiūlyto testo konstravimo metodo įdiegimas viešojoje įstaigoje „Informacinių technologijų institutas“ vykdant ECDL testavimus.
9. Sukurtą testų konstravimo metodą yra tikslinga naudoti tobulinant ECDL ir kitas aukšto lygio kompiuterinių įgūdžių testavimo sistemas.

Mokslinio darbo rezultatų skelbimas. Mokslinio darbo rezultatai paskelbti šešiose mokslinėse publikacijose.

Straipsniai leidiniuose, įtrauktuose į Mokslinės informacijos instituto (ISI) duomenų bazes:

1. Danielienė R. (2006). Research of typical infrastructure of ECDL. *E-learning and teaching of information technologies and informatics, computer literacy, ECDL and EUCIP frameworks. Databases and information systems: 7th International Baltic Conference on Databases and Information Systems, Vilnius, July 3–6, 2006,* p. 9–16. ISBN 9986-19-920-4.
2. Danielienė R., Telešius E. (2008). Analysis of Computer-Based Testing Systems. *Conference Information: Conference on Human System Interactions, Conference proceedings, May 25–27, 2008 Crakow, Poland,* p. 960–964. ISBN 1-4244-1543-8.

Straipsniai leidiniuose, įtrauktuose į kitas tarptautines duomenų bazes, patvirtintas Lietuvos mokslo tarybos:

3. Danielienė R., Telešius E. (2005). ECDL testavimo sistemos formalizavimo problemos. *Informacijos mokslai*. 34 tomas, p.13–17. ISSN 1392-0561.
4. Danielienė R., Telešius E. (2009). Internetinės ECDL testavimo sistemos inovatyvūs sprendimai. *Informacijos mokslai*. 50 tomas, p. 257–261 ISSN 1392-0561.

Straipsniai kituose recenzuojamuose periodiniuose ar tęstiniuose tarptautiniuose ar užsienio leidiniuose:

5. Danielienė R., Telešius E. (2008). Analysis of Computer-Based ECDL Testing. *In: Nunes, M. B., McPherson, M. (eds.) e-Learning 2008*. p. 243–246. IADIS Press, Amsterdam. ISBN 978-972-8924-58-4.
6. Danielienė R., Telešius E. (2006). Research of a typical e-services based ECDL testing infrastructure. *In: Milosz, M., Muryjas, P. (eds.) Varia Informatica. Technologie i bezpieczenstwo*. p. 111–122. Lublin, Poland. ISBN 978-83-922646-5-1.

CURRICULUM VITAE

PERSONAL DATA

Name: Renata Danielienė
Date and place of birth: March 10, 1980, Telšiai, Lithuania
E-mail: renata@ecdl.lt

INSTITUTION

Vilnius University Kaunas Faculty of Humanities
Department of Informatics
Muitines str. 8, LT-44280, Kaunas, Lithuania
Tel. (+370 37) 422 566

Education background

2005–2009	Doctoral studies at the Department of Informatics, Kaunas Faculty of Humanities, Vilnius University, Kaunas, Lithuania.
2002–2004	Master degree of Informatics, Faculty of Informatics, Kaunas University of Technology, Kaunas, Lithuania.
1998–2002	Bachelor degree of Informatics, Faculty of Informatics, Kaunas University of Technology, Kaunas, Lithuania.

Professional experience

2008– now	Information System Manager at Institute of Information Technologies, Kaunas, Lithuania.
2006–2008	Assistant at Department of Informatics, Kaunas Faculty of Humanities, Vilnius University, Kaunas, Lithuania.
2002–2008	Programmer at Information System Service, Kaunas University of Technology.
2001–2002	Technician at Information System Service, Kaunas University of Technology.

International courses

2007 (May–August)	PhD course at Brandenburg University of Technology, Cottbus, Germany. ERASMUS academic exchange program.