| Subject code | Credits |
|---|---|
| INF3033 | 4 |

**Title**

INFORMACIJOS MODELIAVIMAS IR PAIEĢKA

**Title in English**

**INFORMATION MODELING AND RETRIEVAL**

**Subject goal and annotation**

The course aims to study the theories and techniques used in modelling of information and text based retrieval. It discusses classical and modern techniques of knowledge modelling and their application for information retrieval. Students start from the basic notions and concepts, and finish with the newest challenges in the field.

**Prerequisites**

Programming basics

**Relationship between the learning outcomes of the Programme and learning outcomes of the subject**

| Learning outcomes of the Programme | Learning outcomes of the subject | Criteria for measuring the achievement of learning outcomes |
|---|---|---|
| 3. Knowledge of basic and advanced computer science and its application. | Ability to use information retrieval models, indexing, text preprocessing. | Student demonstrates the ability to model and develop simple IR system |
| 7. Formalization and specification of real-world problems, and ability to describe them at an abstract level | Ability to define problems formally | Students is able to define formal IR model, explain and apply it. |
| 8. Perform interdisciplinary research and development in Internet systems area, apply results in practical applications. | Application of the newest models and methods in information retrieval. | Student demonstrates skills in implementing new IR models. |
| 10. Analysis, design and development of advanced Internet systems. | Ability to build, use and improve production level IR systems. | Student demonstrates ability to install, configure and modify production level search system. |
| 12. Analysis, design and development of diverse software systems. | Ability to build, use and improve production level IR systems | Student demonstrates ability to install, configure and modify production level crawling and search system. |

**Subject content**

| | Lecture topics and contents | Hours |
|---|---|---|
| 1. | Introduction. | 2 |
| 2. | Documents preprocessing. | 5 |
| 3. | Documents indexing. | 5 |
| 4. | Information modelling: logical, vector, probabilistic and other models. | 8 |
| 5. | Web search: crawling, PageRank. | 4 |
| 6. | HMI in IR, information visualization. | 2 |
| 7. | Multimedia search | 2 |
| 8 | Newest tendencies in IR. | 2 |
| | **Total** | **30** |

**Practical work contents**

Development of different components of IR systems.

## Evaluation of study results

| |
|---|
| Final written exam (50%), mid-term written exam (15%), and assessments of laboratory (practical) work (35%). |

## Distribution of subject study hours

| | |
|---|---|
| Lectures | 30 |
| Laboratory work | 30 |
| Individual studies (including studies in groups, preparation for the mid-term and final exams) | 48 |
| **Total** | **108** |

## Recommended literature

| No | Authors of publication and title | Number of copies available | | |
|---|---|---|---|---|
| | | *in the Library of VMU* | *in specialized publication collections at VMU* | *in other libraries* |
| ***Basic materials*** | | | | |
| 1. | C. D. Manning, P. Raghavan, H. Schütze. *Introduction to Information Retrieval.*2008 | On-line book: http://www-csli.stanford.edu/~hinrich/information-retrieval-book.html | | |
| 2. | C. J. Van Rijsbergen. *Information Retrieval. 1979* | On-line book: http://www.dcs.gla.ac.uk/Keith/Preface.html | | |
| 3. | Baeza-Yates R. and Ribeiro-Neto B. *Modern Information Retrieval.*1979 | | 1 | |
| 4. | R. Belew. *Finding Out About.*2001 | http://cseweb.ucsd.edu/~rik/foa/l2h/ | | |
| ***Supplementary materials*** | | | | |
| 1. | The Official Google Blog, Yahho!Search blog, Wolfram Alpha blog | http://googleblog.blogspot.com/, http://www.ysearchblog.com/, http://blog.wolframalpha.com/ | | |

## Subject prepared and coordinated by

| |
|---|
| Prof. Tomas Krilavičius |