# STUDY MODULE DESCRIPTION FORM

| Name of the module/subject | Code |
|---|---|
| **Applications of information technologies** | **1010331571010334978** |

| Field of study | Profile of study (general academic, practical) | Year /Semester |
|---|---|---|
| **Information Engineering** | **(brak)** | **4 / 7** |
| **Elective path/specialty** | **Subject offered in:** | **Course (compulsory, elective)** |
| **Information Technologies** | **Polish** | **obligatory** |

| Cycle of study: | Form of study (full-time,part-time) |
|---|---|
| **First-cycle studies** | **full-time** |

| No. of hours | No. of credits |
|---|---|
| Lecture: **30** Classes: **-** Laboratory: **-** Project/seminars: **15** | **4** |

| Status of the course in the study program (Basic, major, other) | (university-wide, from another field) |
|---|---|
| **(brak)** | **(brak)** |

| Education areas and fields of science and art | ECTS distribution (number and %) |
|---|---|
| **technical sciences** | **4 100%** |

## Responsible for subject / lecturer:

prof. dr hab. inż. Czesław Jędrzejek
email: czeslaw.jedrzejek@put.poznan.pl
tel. 61 665 3532
Wydział Elektryczny
ul. Piotrowo 3A, 60-965 Poznań

## Prerequisites in terms of knowledge, skills and social competencies:

| | | |
|---|---|---|
| 1 | **Knowledge** | K_W04: mStudent has organized knowledge with theoretical foundations of basic program constructions, algorithm implementations, paradigms and programming styles, software verification methods, formal languages, compilers, platforms.<br>K_W08: possesses structured and theoretically founded knowledge of databases and data warehouses;<br>K_W12: K_W12: has ordered and methodological knowledge of software engineering |
| 2 | **Skills** | K_U02: is able to work independently and in a team, is able to estimate the time needed for the commissioned tasks, able to develop and implement a schedule of work to ensure deadlines,<br>K_U03: is able to develop documentation of engineering tasks and prepare a text containing a discussion of the results of this task<br>realizacji tego zadania |
| 3 | **Social competencies** | K_K04: is aware of responsibility for his/her own work and a willingness to comply with the principles of teamwork and shared responsibility |

## Assumptions and objectives of the course:

To acquaint students with the algorithms and methods of extracting information from text. Practical application of lead extraction systems using correlation words: Indri, Terrier. Practical analysis of the results obtained with the construction of systems based on semantic vocabularies / ontologies: Yago2, Reverb, Nell. Storage, access and processing so. NoSQ databases.

## Study outcomes and reference to the educational results for a field of study

### Knowledge:

1. Student has organized knowledge with theoretical foundations of basic program constructions, algorithm implementations, paradigms and programming styles, software verification methods, formal languages, compilers, platforms. - [K_W05]
2. Student is familiarized with state of the art and current trends in computer science. - [K_W19]
3. Student knows common IT engineering technology. - [K_W18]

### Skills:

1. Student is able to use software platforms and environments for simple programs encoding, running and testing in imperative, object-oriented and declarative programming languages - [K_U10]
2. Student is able to prepare requirements, to create object model and to evaluate uncomplicated IT system, including system functions and relations between system elements. - [K_U16]
3. Student is able to choose and to implement proper technologies - [K_U22]

### Social competencies:

| |
|---|
| 1. understands the need and knows the opportunity of continuous training (second-and third-degree, postgraduate courses) ? improvement of language, professional, personal and social skills - [K_K01] |

## Assessment methods of study outcomes

Lecture: written examination checking the knowledge of basic algorithms for information extraction and semantic search.

Project: demonstration of the application made ??by the leading semantic search systems, Terrier.Wykład: egzamin pisemny sprawdzający znajomość podstawowych algorytmów ekstrakcji informacji i wyszukiwania semantycznego.

Projekt: pokaz działania aplikacji zrealizowanych przy pomocy wiodących semantycznych systemów wyszukiwawczych, Terrier.

## Course description

Lecture. Semantic processing of information. Algorithms and methods for extracting information from text. Types of information: structured and unstructured semistrukturalna. Method of LSA (Latent Semantic Analysis) and SVM. Natural language processing methods. Measures of the quality of the extraction.

Tools that use correlations of words: Indri, Terrier. Systems based on the construction of semantic vocabularies/ontologies: Yago2, Reverb, Nell. Search by concepts (focused crawling). Tools: GATE, OmniFind. Search-engine Lucene. Semantic extraction pf legal information (e-discovery). The IBM Watson.

Project. Application of LSA, the extended semantics. Projects using Indri, Terrier: query language and the use of quality function. Examples of different tokenizers. The analysis of the results for extraction quality measurement. Search of terrorist content on the Internet.

## Basic bibliography:

1. 1.Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze An Introduction to Information Retrieval, Cambridge UP, 2009

2. W. Bruce Croft, Donald Metzler, and Trevor Strohman, Search Engines: Information Retrieval in Practice Addison Wesley; 1 edition (2009)

3. Articles referring to Yago2, Reverb, Nell, Terrier

## Additional bibliography:

1. Dokumentation: Gate, Terrier i Omnifind

## Result of average student's workload

| Activity | Time (working hours) |
|---|---|
| 1. Lecture | 30 |
| 2. Independent work on the subject of the lecture. | 25 |
| 3. Preparation to project | 15 |
| 4. Doing project | 15 |
| 5. Exam preparation | 15 |

## Student's workload

| Source of workload | hours | ECTS |
|---|---|---|
| Total workload | 100 | 4 |
| Contact hours | 45 | 2 |
| Practical activities | 30 | 1 |