



## Data bases

### Educational subject description sheet

#### Basic information

<b>Field of study</b> Bioinformatics	<b>Education cycle</b> 2022/23
<b>Speciality</b> -	<b>Subject code</b> BD000000BBIS.L8.0440.22
<b>Organizational unit</b> The Faculty of Biology and Animal Science	<b>Lecture languages</b> english
<b>Study level</b> First-cycle programme	<b>Mandatory</b> optional
<b>Study form</b> Full-time	<b>Block</b> major subjects (conducted) in foreign languages
<b>Education profile</b> General academic	<b>Disciplines</b> <b>Subject related to scientific research</b> No
	<b>Subject shaping practical skills</b> Nie
<b>Teacher responsible for the subject</b>	Wojciech Jakubowski
<b>Other teachers conducting classes</b>	Wojciech Jakubowski
<b>Period</b> Semester 4	<b>Examination</b> graded credit
	<b>Activities and hours</b> laboratory classes: 30
	<b>Number of ECTS points</b> 3.0

#### Goals

C1	To familiarize students with the organization of databases ranging from the flat-files to relational ones. Especially for the flat files of the nucleic or protein sequences are discussed. Some methods of matching of two or more sequences are written in the form of python scripts. The Needleman-Wunsch algorithm is discussed too. The structure of the relational databases with the methods of their normalization is presented. On the end the SQL scripts and their application to tested database is discussed and practised.
----	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Subject's learning outcomes

Code	Outcomes in terms of	Effects	Examination methods
<b>Knowledge - Student knows and understands:</b>			
W1	Student has mastered the method of processing data bases of nucleic and protein sequences, is able to determine the fitting of two sequences - understands the importance of the penalty function, can write a script that searches for sequences containing the answer patterns. Student knows the structure of a relational database, can provide database normalization and formulate a simple queries in SQL language.	BI_P6S_WG07, BI_P6S_WG08	written credit, project, presentation
<b>Skills - Student can:</b>			
U1	Student knows the databases structure of nucleic and protein sequences, can in Python read alphanumerical sequences, understands and knows the algorithms for matching two sequences - can use the various functions of punishment, knows sequential algorithms for searching databases. Student has mastered the structure of relational databases, understands the importance of normalization of the database, he can write a SQL statement for editing and searching the existing relational database.	BI_P6S_UW01, BI_P6S_UW04, BI_P6S_UW05, BI_P6S_UW06, BI_P6S_UW10	written credit, project
<b>Social competences - Student is ready to:</b>			
K1	Student is prepared to work with different database systems used in bioinformatics, can interact and work in a group, sees the need for constant updating of skills.	BI_P6S_KK01, BI_P6S_KK03	written credit, project

## Balance of ECTS points

Activity form	Activity hours*	
laboratory classes	30	
class preparation	38	
project preparation	10	
presentation/report preparation	10	
consultations	2	
<b>Student workload</b>	<b>Hours</b> 90	<b>ECTS</b> 3.0
<b>Workload involving teacher</b>	<b>Hours</b> 30	<b>ECTS</b> 1.0
<b>Practical workload</b>	<b>Hours</b> 30	<b>ECTS</b> 1.0

\* hour means 45 minutes

## Study content

No.	Course content	Activities
1.	1. Flat-file and relational databases. Theoretical basis of information storage. Data types. 2. The Database of sequences of nucleic acids and proteins; FASTA, GenBank and Swiss-Prot databases. 3. The database of nucleic acid and proteins sequences; The sequence comparison algorithms. 4. Needleman-Wunsch algorithm. 5. Algorithms for sequence alignment - scoring matrices PAM and BLOSUM. 6. Algorithms for searching of the database of nucleic acid and proteins sequences. 7. Relational databases - basic concepts. 8. Relational databases - normalization of the relational schema. 9. Relational databases - design sample database schema. 10. SQL language - introduction; the select command; the counting functions. 11. SQL language - select command, joining of the tables. 12. SQL language - select command, grouping. 13. Test of knowledge of SQL language. 14. SQL language - DDL, DML and DCL commands. 15. The relational database of sequences of nucleic acid and protein.	laboratory classes

## Course advanced

### Teaching methods:

problem-solving method, project-based learning (PBL), teamwork, computer lab/laboratory, discussion, classes

Activities	Examination methods	Percentage in subject assessment
laboratory classes	written credit, project, presentation	100%

## Literature

### Obligatory

1. Date C.J. An Introduction to Database Systems, Addison Wesley 2003
2. Hernandez M. J.: Database Design for Mere Mortals: A Hands-On Guide to Relational Database Design, Addison-Wesley Professional 2013
3. Higgs, P. G., Attwood T. K. Bioinformatics and Molecular Evolution; Blackwell Publishing Ltd 2005

### Optional

1. NCBI Sample GenBank Record: <https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html>

## Kierunkowe efekty uczenia się

Kod	Treść
BI_P6S_KK01	Absolwent jest gotów do krytycznej oceny posiadanej wiedzy i jej aktualizacji
BI_P6S_KK03	Absolwent jest gotów do stosowania metod badawczych właściwych dla bioinformatyki, ma znajomość rozwoju dziedzin nauki i dyscyplin naukowych
BI_P6S_UW01	Absolwent potrafi stosować zaawansowane techniki informatyki: pracować w środowiskach różnych systemów operacyjnych, stosować różne programy użytkowe, tworzyć proste programy komputerowe oraz projektować bazy danych biologicznych i zootechnicznych
BI_P6S_UW04	Absolwent potrafi poszukiwać i wykorzystywać informacje z zakresu biologii, statystyki matematycznej i informatyki, także w języku obcym
BI_P6S_UW05	Absolwent potrafi samodzielnie wykonywać zadania badawcze lub ekspertyzy z zakresu biologii, zootechniki i informatyki pod kierunkiem opiekuna naukowego oraz w ramach pracy grupowej i wykorzystywać przy tym dostępne źródła informacji, w tym elektroniczne
BI_P6S_UW06	Absolwent potrafi stosować metody informatyczne do opisu i interpretacji wyników uzyskanych w analizie danych biologicznych i hodowlanych
BI_P6S_UW10	Absolwent potrafi samodzielnie wykonywać proste projekty badawcze z zakresu bioinformatyki
BI_P6S_WG07	Absolwent zna i rozumie w stopniu zaawansowanym zagadnienia z zakresu problemów właściwych dla bioinformatyki oraz zna ich powiązania z innymi dyscyplinami przyrodniczymi i możliwościami ich wykorzystania w praktyce
BI_P6S_WG08	Absolwent zna i rozumie w stopniu zaawansowanym zagadnienia z zakresu kategorii pojęciowych i terminologii informatycznej, biologicznej, matematycznej, fizycznej, chemicznej, informatycznej i rolniczej oraz podstawowe procesy zachodzące w cyklu życia urządzeń