

Programme Name: Big Data Solutions

Programme Description

Degree awarded	Master of Science in Software Engineering
Specialization	Big Data Solutions
Mode of Study	Full-Time
Language of Instruction	English
Programme Duration	2 years (120 ECTS), official start date September, 1
Entry Requirements	<p>Academic Entry Requirements: Bachelor Degree or equivalent degree and qualification.</p> <p>English Language Requirements: English as a native language / certificate TOEFL (paper 500 and better; web 55 and better) or IELTS (5.5 or better) or Equivalent Certificate / TPU Entrance Test.</p> <p>Selection process: All individuals are selected on their results of TPU Entrance Exams. Additional selection criteria: GPA in Bachelor Programme; relative merits and abilities of the applicant, approved by certificates.</p>
Fees and Funding	General TPU policies apply. Please see regulations that apply to this programme or make an enquiry to the international department.

Introducing Your Degree

Big Data is recognized as one of the most important areas of future technology, and is fast gaining the attention of many industries, since it can provide high value to companies.

Data sets grow rapidly – in part because they are increasingly gathered by cheap and numerous information-sensing mobile devices, aerial (remote sensing), software logs, cameras, microphones, radio-frequency identification (RFID) readers and wireless sensor networks. The world's technological per-capita capacity to store information has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 exabytes (2.5×10¹⁸) of data are generated. One question for large enterprises is determining who should own Big Data initiatives that affect the entire organization.

The educational programme "Big Data Solutions" will help students understand the Big Data analytics capabilities and potential benefits and support them seeking to formulate more effective data-driven analytics strategies.

Programme Overview

This programme provides an in-depth coverage of topics in Big Data from data generation, storage, management, transfer, to analytics, with focuses on the state-of-the-art technologies, tools, architectures, and systems that constitute Big Data computing solutions in high-performance networks. Real-life Big Data applications in various domains (particularly in sciences and finance) are introduced as use cases to illustrate the development, deployment, and testing of a wide spectrum of emerging Big Data Solutions.

Learning Outcomes

Manipulation, storage, and analysis of large-scale data; large-scale distributed file systems like HDFS (Hadoop Distributed File System); large-scale databases including SQL and NoSQL; MapReduce algorithm design. Python and parallel programming. Big Data visualization. Working with Distributed Systems and Cloud Computing.

Competences and skills

Mathematics and technical knowledge and skills in the exploration, modeling, analysis and use of the latest Big Data tools and techniques.

Management skills in Big Data systems implementation and Big Data services.

Research skills in analytics and optimization, focusing on predictive modeling, data mining, business analysis, marketing analytics and others.

Core (mandatory) Courses:

1. Introduction to Big Data
2. Programming languages (Python)
3. Data Bases
4. Foreign Language (Russian)
5. Philosophical and methodological problems of science and technology
6. Data Analysis Methods
7. Distributed Systems and Cloud Computing
8. Large Scale Data Bases
9. Big Data Programming Tools

Elective (optional) Courses:

10. Machine Learning// Imaging: Data Analytics and Pattern Recognition
11. Big Data Analytics// Parallel Programming
12. Data Visualization// Knowledge Management Systems
13. Web Data Mining// Special Topics in Big Data

Degree Requirements

To be awarded to Master Degree a student should successfully complete all programme courses and modules and defend his/her Master thesis.

Career Opportunities

Career fields and types of organizations: science, business.

Positions: Big Data analyst, Big Data programmer.

Internships: CERN (the European Organization for Nuclear Research), SAS Institute.

Facilities and Equipment:

1. 3 servers with Big Data processing software (HP DL385p Gen8, 2 processors 6320 (2.8GHz-16MB) 8-Core Processor Option Kit, 6 Memory modules 8GB 2Rx4 PC3L-10600R-9 , RAID controller P420i (512MB) FBWC RAID 0,1,1+0,5,5+0, 11 HDD 500GB SC 6G 7.2K LFF SATA HotPlug Midline Drive 1y war, Flash drive 120GB 6G SATA VE 3.5in SCC EV G1 SSD)
2. Computed Cluster Complex T-Platforms (HPC-0011102-001) (Computing nodes: 24, processors: 48 (Intel XEON 5150), Computing cores: 96 (2.66Ghz), RAM: 192ГБ, HDD: 2880ГБ, Storage system: 5ТБ, System network: Infiniband 4x, 24 ports, Auxiliary network: Gigabit Etherhet, 48 ports, Service network: ServNet, 25 ports, Peak performance: 1.02 TFLOPS)
3. Computational cluster "SKIF 2" (computing nodes: 39, processors: 78 (Nehalem 4C X5570 2.93G 8M 6.4G), cores: 320 (2.9Ghz), total amount of RAM: 479GB, system network: Infiniband 4x , 39 ports, auxiliary network: Gigabit Etherhet, 39 ports, peak performance: 3.01 TFLOPS).

Academic Exchanges

A part of the Programme can be studied at TPU partner university. Please see all possibilities and regulations at www.ciap.tpu.ru

Programme Director:

Gubin Evgeny I. (e-mail: gubine@tpu.ru, mob. +7 906–958–72–50)

Academic Staff

Implementation of the basic educational programme of Master studies is provided by qualified teachers. Scientists doctorate or PhD with 75% of the teachers providing the training process in the direction of the Magistracy.

The general management of the scientific content and the educational part of the master's programme is carried out by Professor of Science. Direct supervision of master students performed scientific leaders who have a degree and (or) the academic status or leadership experience in the field.

Programme Structure

N	Modules	Credits
<i>Semester 1</i>		
1	Introduction to Big Data	3
2	Programming languages (Python)	6
3	Data Bases	3
4	Foreign Language (Russian)	3
5	Philosophical and methodological problems of science and technology	3
6	Research work	9
<i>Semester 2</i>		
7	Data Analysis Methods	3
8	Distributed Systems and Cloud Computing	6
9	Large Scale Data Bases	3
10	Big Data Programming Tools	3
11	Foreign Language (Russian)	3
12	Research work	9
<i>Semester 3</i>		
13	Machine Learning	4
14	Big Data Analytics	6
15	Data Visualization	5
16	Web Data Mining	6
17	Research work	6