REPUBLIC OF UZBEKISTAN

MINISTRY OF HIGHER EDUCATION, SCIENCE AND INNOVATION

QUALIFICATION REQUIREMENT

FOR THE UNDERGRADUATE EDUCATION FIELD OF

60230400 - COMPUTATIONAL LINGUISTICS

Tashkent – 2024

Order No. 218

June 25, 2024

Educational Field: 60230400 - "Computational Linguistics"

DEVELOPED AND INTRODUCED BY:

Tashkent State University of Uzbek Language and Literature named after Alisher Navoiy.

APPROVED AND IMPLEMENTED BY:

Approved by Order No. 218 of the Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan dated June 25, 2024.

IMPLEMENTED BY:

Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan.

This Qualification Requirement has been developed in accordance with the "State Educational Standard of Higher Education. Basic Provisions", "Classifier of Fields and Specializations of Higher Education", the National Qualifications Framework and Sectoral Qualifications Frameworks of the Republic of Uzbekistan, professional standards, and proposals from employers, and serves as an official regulatory and methodological document.

The official right to publish this Qualification Requirement within the territory of the Republic of Uzbekistan belongs to the Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan.

TABLE OF CONTENTS

1. General Description

Training of bachelors in the educational field **60230400 - Computational Linguistics** is conducted in the full-time form of education. Education is organized on the basis of the credit-module system. The standard duration of the undergraduate program is 4 years.

1.1. Scope of Application

1.1.1. Application of the Qualification Requirement

The Qualification Requirement represents a set of requirements mandatory for all higher education organizations training bachelors in the educational field 60230400 - Computational Linguistics.

1.1.2. Main Users of the Qualification Requirement:

- Administrative staff of higher education organizations (rector, vice-rectors, head of the academic department, deans, and heads of departments) and the professor-teaching staff who are responsible for developing and updating qualification requirements, curricula, and training programs for this educational field, and who ensure the effective implementation of the educational process and are accountable for the preparation level of graduates within their competence;
- Students of higher education organizations pursuing the curriculum and training programs of this educational field;
- State Attestation Commissions evaluating the preparation level of undergraduate graduates;
- Authorized state bodies governing education;
- Bodies providing financing for higher education organizations;
- Authorized state bodies supervising quality control and accreditation of the higher education system;
- Employers and ordering organizations/enterprises;
- Applicants enrolling in higher education organizations, their parents, and other interested parties.

1.2. Description of Professional Activities

1.2.1. Spheres of Professional Activities:

This field pertains to the educational domains of "Computational Linguistics", "Linguistics", and "Fundamentals of Programming". It encompasses a complex set of tasks associated with training professional linguists capable of working effectively in developing philological, predominantly linguistic computer software and systems aimed at solving language-related challenges. This includes translation software, speech recognition systems, Uzbek speech synthesizers, morphological and semantic text analysis tools, the creation of linguistic ontologies, the

formation of electronic dictionaries, the development of linguodidactic platforms, the creation of language corpora, and the optimization of the National Corpus of the Uzbek language. It also includes creating artificial intelligence-based software, natural language processing (NLP), and the formulation of linguistic models.

A graduate who has fully mastered the theoretical and practical courses and successfully passed the final state attestation will be awarded the bachelor qualification (degree) of **"Computational Linguist"** along with official state-pattern document(s) of higher education.

1.2.2. Objects of Professional Activities:

- Scientific research and scientific-technical institutes;
- Organizations within the system of relevant ministries;
- Software development companies and internet companies;
- Various state and non-state enterprises;
- Departments of state and local administration bodies.

Graduates of the bachelor's degree in 60230400 - Computational Linguistics possess the skills to address complex tasks in organizations involved in: developing the linguistic and software support for digital products in science fields; creating electronic dictionaries and e-book databases; data digitization; compiling lexical datasets for dictionary electronic formatting; developing automatic text editing and analysis software and its linguistic processing tools; creating automatic parallel text translation software; developing the Uzbek language corpus; utilizing electronic information technologies in education; implementing data processing, storage, transmission, and dissemination tools/technologies; and building modern programming formats.

1.2.3. Types of Professional Activities:

- Software development activity;
- Scientific research activity;
- Organizational and managerial activity, and similar fields.

1.2.4. Professional Tasks

In accordance with the 6th qualification level of the National Qualifications Framework for the educational field 60230400 - Computational Linguistics, as well as the spheres, objects, and types of professional activities of a bachelor, the graduate must be capable of performing the following professional tasks:

In software development activity:

- Mastering methods for creating linguistic software and utilizing them effectively;
- Having detailed information about linguistic software and employing them efficiently in professional activities;
- Creating databases and linguistic support systems;
- Developing and using field-appropriate systems in scientific and practical activities;
- Knowing and adhering to the code of professional ethics.

In scientific research activity:

- Participating in research activities using modern field methods and technologies at research institutes and centers;

- Studying and systematizing literature and other scientific-technical information related to computational and corpus linguistics;
- Collecting, analyzing, and providing evaluations of research by Uzbek and foreign scientists regarding linguistic and software support, natural language processing, and literary language modeling;
- Purposely searching and analyzing information about the latest scientific achievements on the internet;
- Preparing and editing scientific articles, lectures, monographs, and textbooks; compiling abstracts and bibliographies;
- Preparing scientific projects and developments, and participating in project execution.

In organizational and managerial activity:

- Organizing and managing production activities;
- Organizing, conducting, and actively participating in scientific seminars, conferences, and symposia;
- Demonstrating the ability to manage projects, plan production processes and resources, analyze unforeseen potential risks, and lead project teams;
- Showing capabilities in organizing corporate training based on e-learning and m-learning technologies, and developing corporate databases;
- Utilizing modern information technologies, information systems, and resources in management; creating software products;
- Developing innovative and start-up projects.

2. Requirements for Professional Competences

- Having systematic philosophical knowledge related to worldviews, being able to analyze them independently and apply them in professional activities, and analyzing concepts of reality and existence;
- Knowing topical issues of state policy and being able to analyze socio-economic problems and processes;
- Having a general understanding of the Strategy of New Uzbekistan, reforms, current political changes, and processes; thoroughly knowing the decrees and resolutions of the President of the Republic of Uzbekistan on improving the educational system, resolutions and orders of the Cabinet of Ministers, and state educational standards;
- Understanding the nature of professional documents and operations in one foreign language; possessing essential knowledge of exact sciences within the scope of professional activities and using them on a modern scientific basis;
- Being able to apply information technologies in professional work; understanding the nature and significance of information technologies in an information society; recognizing the dangers of information attacks and threats; and adhering to primary information security requirements;
- Possessing the ability to independently analyze social problems and processes; being able to master new tasks, work on self-improvement, and organize labor on a scientific basis;
- Relying independently on individual knowledge to comprehend and analyze problems of social and personal importance;
- Mastering the basic methods and tools for retrieving, storing, and processing information from the internet; possessing computer skills as an information management tool;
- Possessing skills in preparing and editing scientific articles, and organizing and conducting scientific seminars and conferences;
- Having skills in executing project work to participate in grants and projects announced by state, non-state, and non-profit organizations;

- Being able to use regulatory-legal documents in one's work, make sound independent decisions in professional duties, and accurately perceive and analyze socio-political situations;
- Comprehending the distinct norms of the literary language and distinguishing them from regional dialects;
- Possessing natural language processing and morphological analysis skills, with the ability to tag parts of speech;
- Possessing skills in syntactic text analysis, creating phrase models, and formulating formal syntax;
- Comprehending the social nature of literary language; identifying phonetics as a factor in forming literary pronunciation; distinguishing between language and speech, sound and phoneme; and recognizing relationships and contradictions in pronunciation and orthography;
- Participating in project execution, performing IT laboratory assignments, conducting similar practical training sessions, and analyzing data;
- Mastering the elementary stage of a foreign language and possessing skills in utilizing information and communication technologies;
- Having the qualification to conduct theoretical and practical training sessions for students in the field of computational linguistics based on state educational standards and curricula; having a clear concept of new information related to computational linguistics education;
- Mastering the nature of statistical analysis, statistical software, descriptive statistics, and transformational analysis;
- Knowing the syntax of the Python programming language, being able to develop software in it, and possessing skills in applying programming technologies in practice;
- Being able to use Database Management Systems, create databases for linguistic software in DBMS like MySQL, SQL, and Oracle, and knowing how to build system architecture;
- Knowing how to optimize the database of the developing Uzbek National Corpus, knowing how to create its lexicographic database, possessing tokenization skills for corpus text entries, and being able to build specialized/private language corpora;
- Being able to develop and deploy oral, written, and automatic translation systems in practice using computer systems; applying automatic translation tools; analyzing and researching machine translation theory; and mastering skills to create and optimize linguistic processing tools, databases, and software for translation systems; knowing how to use parallel corpora and construct their knowledge and data bases;
- Possessing qualifications to establish international relations, conduct negotiations, meetings, international correspondence, and electronic communications;
- Knowing how to create and optimize the linguistic and cognitive components of information and intelligent systems for various purposes (thesauri, linguistic ontologies, databases, knowledge bases, syntactic analysis systems/parsers, semantic analysis tools, morphological analyzers);
- Being able to use mathematical principles in formalizing linguistic knowledge and performing the analysis and synthesis of linguistic structures and linguistics expertise;
- Developing the ability to promote the Uzbek language globally using digital technologies;
- Identifying topical problems and providing solutions regarding modern software that processes natural language;
- Utilizing search and communication systems effectively, and possessing the qualification to create them;
- Possessing the ability to organize accelerated instruction and service delivery for taught computer programming languages.

3. Requirements for Internships

Qualification internship is aimed at strengthening theoretical knowledge gained in general professional and specialized disciplines, integrating them with practical (industrial production) processes, and forming relevant practical skills, competences, and qualifications.

The following internships are conducted within the educational field:

1. Introductory Internship;
2. Production (Industrial) Internship;
3. Pre-Graduation Internship.

4. Structure of the Course Catalog

No.	Course Code	Names of Educational Disciplines, Blocks, and Activity Types	Total Workload (Hours)	Credits	Semester
1.00		**General Compulsory Subjects**	**5100**	**170**	
1.01	DSh1404	Religious Studies	120	4	1
1.02	F1204	Philosophy	120	4	2
1.03	JMS1304	Physical Culture and Sport	120	4	3
1.04	OEYT1204	Modern History of Uzbekistan	120	4	4
1.05	XT13-410	Foreign Language	300	10	3, 4
		Professional Compulsory Subjects			
1.06	TK1105	Introduction to Linguistics	150	5	1
1.07	DA1105	Fundamentals of Programming	150	5	1
1.08	MA1105	Fundamentals of Logic	150	5	1
1.09	AN1106	Theory of Algorithms	180	6	1

No.	Course Code	Names of Educational Disciplines, Blocks, and Activity Types	Total Workload (Hours)	Credits	Semester
1.10	LL11-210	Lexicology and Lexicography	300	10	1, 2
1.11	STAT1205	Statistics	150	5	2
1.12	TT1205	Typology of Languages	150	5	2
1.13	M1206	Morphology	180	6	2
1.14	DT12-415	Programming Technologies	450	15	2, 3, 4
1.15	SIN1306	Syntax	180	6	3
1.16	MBBT13-515	Database Management Systems	450	15	3, 4, 5
1.17	NA1406	NLP Algorithms	180	6	4
1.18	TMU1505	Mathematical Methods in Linguistics	150	5	5
1.19	KL15-610	Computational Linguistics	300	10	5, 6
1.20	KMT15-610	Big Data Technologies (BigData)	300	10	5, 6
1.21	MT16-710	Machine Translation	300	10	6, 7
1.22	TK16-710	Language Corpora	300	10	6, 7
1.23	MSA1705	Morphological and Syntactic Analyzers	150	5	7
1.24	MITU1705	Intelligent Data Analysis Methods	150	5	7
2.00		**Elective Subjects**	**1200**	**40**	
2.00		Elective Subjects (8 subjects)	1200	40	3, 4, 5, 6, 7

No.	Course Code	Names of Educational Disciplines, Blocks, and Activity Types	Total Workload (Hours)	Credits	Semester
		TOTAL	**6300**	**210**	
		Qualification: Digital Linguist			
		Qualification Internship & Final State Attestation	900	30	8
		TOTAL	**900**	**30**	
		GRAND TOTAL	**7200**	**240**	

Bibliographic Data

- UDC: 002:651.1/7
- OKS 01.040.01
- Group T 55

Keywords: Professional activity type, competence, object of professional activity, sphere of professional activity, undergraduate curriculum and training program (bachelor's program), profile, educational period, learning outcomes, qualification requirements, undergraduate educational process, quality assessment and control, independent education, organizational and managerial activity, qualification internship, graduation qualification work, state attestation, block of training subjects, higher education institution, design, software engineering, artificial intelligence, computational linguistics, lexical-semantic group, corpus linguistics, social networks, ontologies, computational semantics, morphological models, electronic dictionaries, automatic text translation, information-communicative systems, modern research methods, information and modern pedagogical technologies, models and modeling, virtual electronic knowledge resources, didactics, theory.

APPROVAL SHEET

Developers, coordinated main aligned higher education institutions, and consumers of personnel:

DEVELOPED BY:

- Alisher Navoiy Tashkent State University of Uzbek Language and Literature

  Rector: **Sh. Sirojiddinov** (Signature, Stamp) — 2024

COORDINATED WITH:

- Center for Research on Higher Education Development under the Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan

  Director: **M. Boltaboyev** (Signature, Stamp) — 2024

- "Mahalla va Oila" Publishing House

  Director General: **B. Mavlonov** (Signature, Stamp) — 2024

- LLC "KEEP PRINT"

  Director: **A. Narzulloyev** (Signature, Stamp) — 2024