

70230401 - 计算语言学硕士专业

乌兹别克斯坦共和国高等教育、科学和创新部

70230401 - 计算语言学硕士专业

资格要求

塔什干 - 2024年

乌兹别克斯坦共和国高等教育、科学和创新部 (O'zR OTFIV)

第 272 号命令

2024年 08月 01日

“70230401 - 计算语言学”硕士专业

开发与提交单位：

以阿里舍尔·纳沃伊命名的塔什干国立乌兹别克语与文学大学

批准与实施单位：

根据乌兹别克斯坦共和国高等教育、科学和创新部 2024年 08月 01日 第 277 号命令批准实施。

引入单位：

乌兹别克斯坦共和国高等教育、科学和创新部。

本资格要求根据《高等教育国家教育标准。基本规定》、《高等教育专业与方向分类》、《乌兹别克斯坦共和国国家资格框架与行业资格框架》、职业标准以及人才用人单位的建议制定，属于正式的规范性方法文件。

在乌兹别克斯坦共和国境内正式印刷和出版该资格要求的权利属于乌兹别克斯坦共和国高等教育、科学和创新部。

2

“70230401 - 计算语言学”硕士专业

目录

序号	内容	页码
1.	基本特征	4
1.1.	适用范围	4
1.1.1.	资格要求的适用性	4
1.1.2.	资格要求的主要使用者	4
1.2.	职业活动描述	4
1.2.1.	职业活动领域	4
1.2.2.	职业活动对象	5
1.2.3.	职业活动类型	5
1.2.4.	职业任务	5
2.	对核心能力的要求	7
3.	对实习的要求	9
4.	课程目录结构	10

序号	内容	页码
文献资料	11	
签署页	13	

3

1. 基本特征

“70230401 - 计算语言学”硕士专业

70230401 计算语言学硕士专业的人才培养采用全日制教育形式。专业教学基于学分制（Credit-Module System）组织。硕士学制为 2 年。

完整掌握其理论和实践课程、并通过国家最终认证考试的高等学历毕业生，将被授予“计算语言学专家”硕士学位，并颁发国家样式的正式高等教育毕业证书。

1.1. 适用范围

1.1.1. 资格要求的适用性

本资格要求代表了所有培养 70230401 计算语言学专业硕士的研究生高等教育机构必须遵守的要求总和。

1.1.2. 资格要求的主要使用者：

- 负责本硕士专业资格要求、教学计划和教学大纲的制定与更新，并在自身权责范围内对毕业生培养质量负责的高等教育机构管理人员（校长、副校长、教务处长、院长及系主任）和教授教师团队；
- 学习该硕士专业教学计划和课程大纲的高等教育机构学生；
- 对硕士毕业生准备水平进行评估的国家认证委员会；
- 负责向高等教育机构提供资金保障的机关；
- 负责高等教育体系认证和质量控制的授权国家机关；
- 人才订购方、用人单位及各企业机构；
- 报考高等教育机构的本科毕业生及其他相关利益方。

1.2. 职业活动描述

1.2.1. 职业活动领域：

70230401 计算语言学硕士专业是一门涉及“计算语言学”、“语言学”和“编程基础”教育方向的交叉学科专业。该专业旨在培养能够高效解决语言问题的专业语言学家。培养的核心能力包括：开发旨在解决语言问题的语料库、主要包括翻译软件、语音识别系统、乌兹别克语语音合成器、文本词法和语义分析软件在内的语言学计算机程序和系统；构建语言本体；开发电子词典；研发语言教学平台；创建语料库并完善乌兹别克语国家语料库。同时，还涵盖了基于人工智能的软件开发、自然语言处理（NLP）、构建语言模型所需的知识、技能和熟练度，以及研究多义词、多功能词和同音异义词的分析方法与技术的能力。

4

“70230401 - 计算语言学”硕士专业

1.2.2. 职业活动对象：

- 相关部委系统内的各类组织机构；
- 国际组织及其在乌兹别克斯坦共和国的代表处；
- 软件产品开发公司及互联网公司；
- 各类国有和非国有企业；
- 高等教育系统内的机构；
- 国家和地方管理机关的各部門；
- 与科学、技术和工艺问题相关的学术界、科研机构和组织；
- 旅游信息中心；
- 共和国国有和非国有出版印刷厂；
- “电子政务”中的信息检索系统；
- 开发信息系统语言组件（信息检索领域、智能系统、机器翻译、电子学习、语音自动分析与合成算法、社交网络组建）的国有和私营公司；
- 各机构和部门的语言鉴定与评估部门；
- 职业学院、学术中学（lyceum）和高等教育机构。

1.2.3. 职业活动类型

- 软件开发活动；
- 科学研究活动；
- 组织管理活动；
- 在高等教育、进修和再培训教育机构开展科研教学活动，以及在中等专业、职业教育机构开展教学活动（按规定程序）；
- 硕士毕业生具体从事的职业活动类型，由高等教育机构与教育过程的相关利益方合作确定。

1.2.4. 职业任务

根据国家资格框架中 70230401 计算语言学专业的第 7 级资格要求，并结合硕士职业活动的领域、对象和类型，硕士毕业生应具备履行以下职业任务的能力：

5

“70230401 - 计算语言学”硕士专业

软件开发活动中：

- 掌握语言软件的开发方法并能有效利用；
- 掌握有关语言软件的详细知识，并在职业活动中高效应用；
- 在科学和学术实践活动中发展并应用符合该领域的系统；
- 必须了解并遵守职业道德规范。

科学研究活动中：

- 组织科研活动；
- 开展科学和应用研究，分析实验结果并据此得出具备科学依据的结论，发掘科学创新；
- 撰写和编辑科学论文、报告、专著和教材，针对所研究的课题撰写科学评论，编制摘要和文献目录；
- 在科学文献和互联网中，有针对性地搜寻和获取关于最新科学、设计和技术成就的信息；
- 针对相关专业课题，开发科学项目，以及所解决的科学问题和任务的概念与理论模型。

组织管理活动中：

- 组织和管理生产流程；
- 组织、举办并积极参与学术研讨会、会议和专题讨论会；
- 具备项目管理、生产流程和资源规划、潜在突发风险分析以及管理项目团队的能力；
- 具备基于电子学习（e-learning）和移动学习（m-learning）技术组织企业培训并发展企业数据库的能力。

在高等教育、进修和再培训教育机构开展科研教学活动，以及在中等专业、职业教育机构开展教学活动（按规定程序）：

- 具备在高等和中等专业、职业教育体系机构中，讲授专业培养方向内所规定的课程的理论、实践和实验室课程（包括在教学工作坊内进行）的能力；
- 掌握相关课程的教学法；
- 具备利用现代信息和教学技术、技术化教学手段开展非传统型教学活动的能力；
- 具备编制、准备和规范讲授课程所需教学方法文件的能力；
- 具备利用技术化教学手段来开展相关课程教学活动的能力；
- 具备通过自主学习和创造性探索，在所讲授学科、教学活动方法及其他领域不断进行自我完善的能力；
- 在高等教育、再培训和进修、职业教育机构中开展本专业的教学及教学方法活动；
- 熟练掌握电子（e-learning）、移动（m-learning）和远程信息技术及教学方法综合体；
- 必须定期提高自身的教学与科学技能及资质；
- 必须具备制定连续教育体系国家教育标准、资格要求及其他教学规范文件的能力。

“70230401 - 计算语言学”硕士专业

2. 对职业核心能力的要求

- 必须具备关于科学世界观的知识体系，了解通用方法论学科的基础以及国家政策的当前紧迫问题；
- 具备独立分析社会问题和过程的能力；
- 能够使用一种外语理解与职业活动相关的法律文件和工作实质，了解科学研究和教学方法论，并能将其应用在基于现代科学基础的职业活动中；
- 能够独立掌握新知识，开展自我提升，并以科学为基础组织自身劳动；
- 能够对所学知识进行创造性的批判审视与分析，并能在科研活动中加以利用；
- 能够依托个人知识理解并分析具有社会和个人意义的问题；
- 能在自身活动中利用规范性法律文件，并在职业活动中做出合理的独立决策；
- 掌握从互联网获取、存储和处理信息的主要方法与工具，具备将计算机作为信息管理工具的操作技能；
- 能够使用信息技术，理解信息社会条件下信息技术的本质与意义，认识到信息攻击的风险与威胁，具备遵守信息安全基本要求的能力；
- 能够准备用于申请专利或著作权登记证书的规范性文件；

“70230401 - 计算语言学”硕士专业

- 必须具备编写项目申请书的技能，以参与由国家、非国有和非营利组织宣布的项目；
- 了解科研教学研究的本质、对研究的特征描述、研究策略及其选择路径和方法、基于定性和定量的研究及其特征、问题的选择路径和方法、统计分析的本质、统计分析程序、描述性统计、转换分析、相关性以及与计算语言学教育相关的最新信息；
- 硕士生应能够形成研究课题，选择适合课题的科研方法，为科研课题准备可行性论证，并能高效利用科学文献；
- 具备开展科学和应用研究、处理实验结果并据此得出具备科学依据的结论、准备和编辑科学论文、组织和举办学术研讨会、会议和专题讨论会以及开发科学项目的技能；
- 认识文学语言的特有规范，并具备将其与方言区分开来的技能；
- 独立掌握方法论和语言教学法的现代知识，能够预见教学方法问题并找到解决方案；
- 必须了解如何完善正在创建的乌兹别克语国家语料库的基础库，了解如何创建其词典编纂基础，能够为语言语料库基础库对文本进行分词（tokenization）、进行词性标注（POS tagging），并具备创建专用语言语料库的技能；
- 具备为了生产和实践活动目的而进行自然语言文本处理（语言标注、编写形式化词汇注释、自动标引和摘要，即在保持大容量文本核心内容的同时将其精简为小容量文本）的技能；

- 了解如何为各种目的创建和完善信息与智能系统的语言及认知组件（叙词表、语言本体、数据库、知识库、句法分析系统剖析器、语义分析程序、词法分析器）；
- 了解如何**开发**和完善语言学资源（文本语料库、词典、语音、词汇和术语数据库）；
- 能够借助计算机系统在实践中开发并引入平行、口头、书面和自动翻译系统，能够应用自动翻译工具，分析和研究机器翻译理论，具备创建和完善翻译软件语言保障、数据库及软件的技能，能够使用平行语料库并了解如何创建其知识库和数据库；
- 能够将文本从外语高质量地翻译为国家语言和官方语言，以及从国家语言和官方语言翻译为外语；
- 能够针对处理自然语言的现代程序确定当前紧迫问题并给出解决方案；

8

“70230401 - 计算语言学”硕士专业

- **开发**自然语言自动处理系统的语言组件（口语语音合成与识别、文本生成、自动翻译、自动摘要与标引）及智能系统；
- 能够在语言知识形式化中利用数学基础，并能对语言结构和语言学知识进行分析与合成；
- 能够执行计算机词典编纂任务：能够创建电子词典数据库，了解如何创建新一代电子词典以及语言学词典的移动应用程序；
- 了解如何准备语言模型和资源，为人工智能进行自然语言建模，并能实践中加以应用；
- 能够创建包含语音和自然语言书面文本自动处理功能的电子系统；
- 掌握 Python 编程语言，能够在数据库（SQL, MySQL, Oracle 等）中工作；
- 了解如何创建符合自然语言处理任务（如智能检索）的语言系统的操作形式化模型；
- 熟练并广泛应用机器学习、自动语言标注、文本句法和语义分析等方法；
- 具备创建语音识别与合成、机器翻译、文本理解和语义分析程序的语言与软件保障及数据库的技能；
- 能够独立完成项目、参与项目执行、了解如何对紧迫问题进行设计、能够开展信息技术（IT）方面的实验室课程并能分析数据；
- 必须具备以强化方法教授所学语言并组织服务的能力。

3. 对实习的要求

科研教学实习旨在巩固公共专业和专门学科的理论知识，并将其与实际生产过程相结合，侧重于形成相应的实际技能、核心能力和职业熟练度。

本硕士专业规定开展科研教学实习。

9

“70230401 - 计算语言学”硕士专业

4. 课程目录结构：

课程资格代码	教学科目、模块和活动类型名称	序号	总负荷学时	学分	学期
1.00	必修科目				
1.01 / ITM 104	科研方法论	1	120	4	1
1.02 / MFO'M 304	特殊科目教学法	2	120	4	3
1.03 / TTQI 105	自然语言处理 (NLP)	3	150	5	1
1.04 / KL 1-2	语料库语言学	4	300	10	1, 2
1.05 / PDT 1-2	Python 编程语言	5	300	10	1, 2
1.06 / MBDM 1-3	数据库与数据挖掘 (Data Mining)	6	420	14	1, 2, 3
小计：		1410	47		
2.00	选修科目		390	13	2, 3
选修科目 (3门科目)		390	13	2, 3	

课程资格代码	教学科目、模块和活动类型名称	序号	总负荷学时	学分数	学期
获授资格：	数字语言学家、教学研究员				
3.00	科研活动		1800	60	1, 2, 3, 4
总计			3600	120	

10

UDK: 002:651.1/7

OKS 01.040.01

“70230401 - 计算语言学”硕士专业

文献资料

组别 T 55

关键词：

职业活动类型、核心能力、职业活动对象、职业活动领域、硕士教学计划和教学大纲（硕士项目）、专业方向、学习期限、科学研究过程、人工智能、计算语言学、词汇-语义组、语料库语言学、社交网络、本体、计算机语义学、词法模型、电子词典、文本自动翻译、信息沟通系统、现代研究方法、信息和现代教学技术、模型与建模、学科工作大纲、科学研究组织、虚拟电子知识资源、教学法、理论。

11 & 12（签署与盖章部分）

开发人员、达成一致的相关主要高等教育机构及用人单位：

开发单位：

以阿里舍尔·纳沃伊命名的塔什干国立乌兹别克语与文学大学

（公章：乌兹别克斯坦共和国高等教育、科学和创新部 / 塔什干国立乌兹别克语与文学大学）

校长： Sh. Sirojiddinov（签字）

2024年

达成一致：

乌兹别克斯坦共和国高等教育、科学和创新部下属高等教育发展研究中心

（公章：高等教育发展研究中心）

主任： M. Boltaboyev（签字）

2024年

乌兹别克斯坦国立世界语言大学

（公章：乌兹别克斯坦世界语言大学）

校长： I. Tuxtasinov（签字）

2024年

“KEEP PRINT” 有限责任公司（MCHJ）

（公章：塔什干州赞吉阿塔区“KEEP PRINT” 有限责任公司）

总经理： A. Narzulloyev（签字）

2024年