

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION
Federal State Autonomous Educational Institution
higher education
NORTH CAUCASIAN FEDERAL UNIVERSITY

I APPROVE

Dean of the Faculty of Mathematics
and computer science named after
Professor N.I. Chervyakov,
Candidate of Physical and
Mathematical Sciences, Associate
Professor
Grobova T.A.

ASSESSMENT TOOLS FUND FOR DISCIPLINE
"Artificial Intelligence Ethics and Safety / Этика и безопасность
искусственного интеллекта"

Direction of training	02.04.01 Mathematics and Computer Science
Profile	Training Program for Electronic Information
Year of commencement of studies	<u>2026</u>
Form of study	full-time
Implemented in the 2nd semester	

Introduction

1. Purpose: The assessment toolkit is designed to provide a methodological basis for organizing and conducting ongoing assessments in the course "Artificial Intelligence Ethics and Safety / Этика и безопасность искусственного интеллекта" Ongoing assessments in this course are a systematic method for assessing students' knowledge, skills, and abilities. The objectives of ongoing assessments are to obtain initial information on the progress and quality of competencies acquired, as well as to encourage regular, targeted student work. This work is intended to develop a certain level of competencies.

2. FOS is an appendix to the program of the discipline " Artificial Intelligence Ethics and Safety / Этика и безопасность искусственного интеллекта"

3. Developer: Department of Computational Mathematics and Cybernetics.

4. An examination of the organophosphorus compounds was carried out.

Members of the expert group:

Chairman

Babenko M. G., Head of the Department of Computational Mathematics and Cybernetics

Members of the commission:

Peleshenko V.S., Associate Professor of the Department of Computational Mathematics and Cybernetics

Peleshenko T. A., Associate Professor of the Department of Computational Mathematics and Cybernetics

Representative of the employer organization Kornienko S. A., head of the branch department FSUE "Main Radio Frequency Center" in the Southern and North Caucasian Federal Districts districts

Expert opinion: the fund of assessment tools corresponds to the educational program in the field of training 02.04.01 Mathematics and computer science profile "Training Program for Electronic Information» and is recommended for conducting ongoing monitoring of students' academic performance and midterm assessment.

5. The validity period of the FOS is determined by the period of implementation of the educational program.

1. Description of the criteria for assessing competence at various stages of its development, description of assessment scales

Competence (ii), indicator (s)	Levels of development competence (s),			
	The minimum level has not been reached (Unsatisfactory) 2 points	Minimum level (satisfactorily) 3 points	Intermediate level (Fine) 4 points	High level (excellent) 5 points
<i>Competence: UK-5 Able to analyze and take into account cultural diversity in the process of intercultural interaction</i>				
ID-1UK-5 Analyzes and draws conclusions on social, ethical, scientific and technical issues arising in professional activities. ID-2UK-5 Objectively evaluates the diversity of cultures and identifies their individual characteristics	has difficulty understanding the basic concepts related to ensuring information and psychological security of the individual, society and the state, the concepts of information warfare, information warfare and the forms of their manifestation in the modern world	is mainly oriented in the basic concepts related to ensuring information and psychological security of the individual, society and the state, the concepts of confrontation, information warfare and the forms of their manifestations in the modern world	is familiar with the basic concepts associated with ensuring information and psychological security of the individual, society and the state, the concepts of information warfare, information warfare and the forms of their manifestation in the modern world	is fully versed in the basic concepts related to ensuring information and psychological security of the individual, society and the state, the concepts of information warfare, information warfare and the forms of their manifestation in the modern world
<i>Competency: OPC-2. Able to create and explore new mathematical models in the natural sciences, improve and develop concepts, theories, and methods</i>				
ID-1OPK-2 Conducts research into new mathematical models in the natural sciences ID-1OPK-2 Develops new and improves existing concepts, theories and methods	Unable to use fundamental knowledge in the field of numerical methods in future professional activities. Does not possess the skills to use numerical methods to solve professional problems.	Partially able to use fundamental knowledge in the field of numerical methods in future professional activities. Has minimal skills in using numerical methods to solve professional problems.	In general, he is able to use fundamental knowledge in the field of numerical methods in future professional activities. Has a good command of the skills of using numerical methods to solve professional problems.	Able to effectively use fundamental knowledge in the field of numerical methods in future professional activities. Has a high level of skills in using numerical methods to solve professional problems.
<i>Competence: PC-8 Capable of participating in the management of projects for the creation of</i>				

<i>information systems and software systems at various stages of the life cycle</i>				
ID-1PK-8 Manages the development, development and maintenance of information systems and software systems	has difficulty analyzing sources and classifying and assessing information security threats	mainly analyzes sources and provides classification and assessment of information security threats	analyzes sources and provides classification and assessment of information security threats	fully analyzes sources and provides a classification and assessment of information security threats
ID-2PK-8 Knows and applies software development trends in his/her professional activities				

The assessment of the level of development of competence in the discipline is carried out on the basis of the "Regulations on the current monitoring of academic performance and midterm assessment of students in higher education programs - bachelor's programs, specialist programs, master's programs - at the federal state autonomous educational institution of higher education "North Caucasus Federal University" in the current version.

ASSESSMENT TOOLS FOR CHECKING THE LEVEL OF DEVELOPMENT OF COMPETENCIES

Task number	Correct answer	Contents of the question	Competence
full - time education			
1.	AI ethics is the field that studies the moral and ethical aspects of the development and use of AI systems.	What is AI ethics?	At K- 5 OPK-2 PC-8
2.	Key concerns include bias, transparency, accountability and data protection.	What are the main issues in AI ethics?	At K- 5 OPK-2 PC-8
3.	The principles include respect for autonomy, nonmaleficence, beneficence and justice	What ethical principles should be considered when developing AI?	At K- 5 OPK-2 PC-8
4.	Using diverse datasets, testing for bias and auditing models.	How to prevent bias in AI systems?	At K- 5 OPK-2 PC-8
5.	Transparency means understanding how an AI system works and how it makes decisions.	What is transparency in AI?	At K- 5 OPK-2 PC-8
6.	Establishing feedback and complaints mechanisms and transparency in decision-making	How to ensure accountability in AI systems?	At K- 5 OPK-2 PC-8
7.	Risks include diagnostic errors, treatment bias, and data privacy.	What are the risks associated with using AI in healthcare?	At K- 5 OPK-2 PC-8

8.	Issues include bias in assessment, access to education, and data privacy.	What ethical issues arise when using AI in education?	At K- 5 OPK-2 PC-8
9.	Use of encryption, anonymization and data minimization	What measures can be taken to protect data in AI?	At K- 5 OPK-2 PC-8
10.	The standards include UNESCO recommendations and EU guidelines on AI.	What international standards govern AI ethics?	At K- 5 OPK-2 PC-8
11.	AI security is the protection of AI systems from malicious influences and errors.	What is AI safety?	At K- 5 OPK-2 PC-8
12.	Risks include cyber attacks, programming errors and unforeseen consequences	What are the main security risks in AI?	At K- 5 OPK-2 PC-8
13.	Use of secure protocols, regular updates and monitoring.	What measures can be taken to protect AI from cyberattacks?	At K- 5 OPK-2 PC-8
14.	Methods include penetration testing and vulnerability analysis.	What methods are used to test AI safety?	At K- 5 OPK-2 PC-8
15.	Consequences include financial losses, damage to reputation and physical harm.	What consequences can errors in AI systems have?	At K- 5 OPK-2 PC-8
16.	Approaches include testing, verification and validation of systems.	What approaches are used to ensure AI trustworthiness?	At K- 5 OPK-2 PC-8
17.	Risks include	What are the risks associated with autonomous AI systems?	At K- 5

	unpredictable behavior and lack of control.		OPK-2 PC-8
18.	Using quality checks, testing and regular updates	What measures can be taken to prevent errors in AI?	At K- 5 OPK-2 PC-8
19.	The standards include ISO guidelines on AI safety.	What international standards govern AI safety?	At K- 5 OPK-2 PC-8
20.	The tools include vulnerability analysis and penetration testing programs.	What tools are used to analyze AI safety?	At K- 5 OPK-2 PC-8
21.	Issues include liability, intellectual property and data protection.	What legal issues arise when using AI?	At K- 5 OPK-2 PC-8
22.	Laws vary, but include GDPR in the EU and CCPA in the US.	What laws regulate the use of AI in different countries?	At K- 5 OPK-2 PC-8
23.	Risks include lawsuits and fines for discrimination	What are the legal risks associated with bias in AI?	At K- 5 OPK-2 PC-8
24.	Use of legal advice and compliance with existing laws	What steps can be taken to ensure compliance with legal requirements when developing AI?	At K- 5 OPK-2 PC-8
25.	The agreements include OECD recommendations on AI.	What international agreements regulate the legal aspects of AI?	At K- 5 OPK-2 PC-8
26.	Tools include legal databases and consulting services.	What tools are used to analyze AI legal risks?	At K- 5 OPK-2 PC-8
27.	Consequences include fines, lawsuits and reputational damage.	What are the consequences of violations of legal requirements in AI?	At K- 5 OPK-2 PC-8

28.	Approaches include regular audits and consultations with lawyers	What approaches are used to ensure the legal security of AI?	At K- 5 OPK-2 PC-8
29.	Risks include patent and copyright infringement.	What are the risks associated with intellectual property in AI?	At K- 5 OPK-2 PC-8
30.	Use of patents, copyrights and non-disclosure agreements	What measures can be taken to protect intellectual property in AI?	At K- 5 OPK-2 PC-8
31.	Issues include job losses, discrimination and impact on society	What social issues arise from the use of AI?	At K- 5 OPK-2 PC-8
32.	Measures include worker retraining and social support programs.	What measures can be taken to mitigate the social impacts of AI?	At K- 5 OPK-2 PC-8
33.	a) Risks include stress, addiction and social isolation.	What are the risks associated with the impact of AI on mental health?	At K- 5 OPK-2 PC-8
34.	Approaches include the development of socially responsible systems and societal impact assessments	What approaches are being used to ensure the social responsibility of AI?	At K- 5 OPK-2 PC-8
35.	Initiatives include UN digital development programmes.	What international initiatives are aimed at addressing the societal challenges of AI?	At K- 5 OPK-2 PC-8
36.	Tools include sociological research and impact assessment.	What tools are used to analyze the social impacts of AI?	At K- 5 OPK-2 PC-8
37.	The consequences include social tensions	What are the consequences of ignoring AI's social problems?	At K- 5 OPK-2

	and reduced trust in technology.		PC-8
38.	Approaches include developing systems that take into account social factors	What approaches are used to ensure social safety of AI?	At K- 5 OPK-2 PC-8
39.	Risks include cultural homogenization and loss of traditional values.	What are the risks associated with the impact of AI on culture?	At K- 5 OPK-2 PC-8
40.	Measures include developing culturally sensitive systems and supporting cultural programming	What measures can be taken to preserve cultural diversity when using AI?	At K- 5 OPK-2 PC-8
41.	The issues include questions of consciousness, free will, and moral responsibility.	What philosophical issues arise when using AI?	At K- 5 OPK-2 PC-8
42.	Theories include utilitarianism, deontology, and virtualism .	What philosophical theories underlie AI ethics?	At K- 5 OPK-2 PC-8
43.	Risks include loss of control and unpredictable behavior	What are the risks associated with creating strong AI?	At K- 5 OPK-2 PC-8
44.	Approaches include developing systems taking into account philosophical principles	What approaches are used to ensure philosophical safety of AI?	At K- 5 OPK-2 PC-8
45.	Discussions include discussions in academic circles and at conferences.	What international debates are taking place on the philosophical aspects of AI?	At K- 5 OPK-2 PC-8

46.	Tools include philosophical texts and discussions in academic journals.	What tools are used to analyze the philosophical implications of AI?	At K- 5 OPK-2 PC-8
-----	---	--	--------------------------

2. Description of the rating scale

Under the rating system, student performance in each course is assessed through ongoing assessments and midterm assessments. The student assessment system is based on a set of assessments designed to test the material covered, optimally spaced throughout the course.

3. Competency assessment criteria *

A student receives an "excellent" grade if they have a comprehensive understanding of the fundamental concepts related to ensuring information and psychological security for individuals, society, and the state, as well as the concepts of information warfare, information confrontation, and their manifestations in the modern world. They also fully analyze sources and classify and assess information security threats.

A student is given a grade of "good" if he, Understands the fundamental concepts related to ensuring information and psychological security for individuals, society, and the state, as well as the concepts of information warfare and information confrontation, and their manifestations in the modern world. Analyzes sources and classifies and assesses information security threats.

A "satisfactory" grade is awarded to a student who has a basic understanding of the fundamental concepts related to ensuring information and psychological security for individuals, society, and the state, as well as the concepts of information warfare, information confrontation, and their manifestations in the modern world. They primarily analyze sources and classify and assess information security threats.

A student is given a grade of "unsatisfactory" if they struggle with the basic concepts of ensuring information and psychological security for individuals, society, and the state, as well as the concepts of information warfare and information confrontation, and their manifestations in the modern world. They struggle to analyze sources and classify and assess information security threats.