

Serbia Accelerating Innovation and Growth Entrepreneurship Project

Terms of Reference and Scope of Services

for

Artificial Intelligence (AI) Risk Specialist

1. Background information

Serbia has received a loan from the World Bank (WB) for the **Serbia Accelerating Innovation and Growth Entrepreneurship (SAIGE) Project**¹ (Project) to support the strengthening of the Serbian Innovation and Entrepreneurship ecosystem. The main objective of the Project is to improve (i) the relevance and excellence of scientific research; and (ii) innovative entrepreneurship and access to finance for enterprise growth, as a way of contributing to Serbia's growth and competitiveness. The European Union Instrument for Pre-accession Assistance (IPA-II) is also providing complimentary financing support for the SAIGE Project. The Project includes the following three components:

Component 1: Research Sector Reforms

Component 2: Enterprise Acceleration

Component 3: Project Implementation, Monitoring, and Capacity Building.

The Ministry of Education, Science and Technological Development (MoESTD) until October 2022 was responsible for the overall Project coordination and implementation and execution of Component 1.2 RDI Reforms. After the April 2022 elections, a new line ministry – Ministry of Science, Technological Development, and Innovation (NITRA, hereinafter: the Client) has been tasked (in October 2022) as the successor, with the overall Project coordination and implementation. The Project is coordinated (and implemented) by the Project Implementation Unit (PIU) at NITRA comprising a PIU manager, project officers for the three Project components, experts for environmental and social standards, including ethics, M&E and RDI transformation officer, and other technical and administrative staff.

In addition to the above, in mid-2023, Serbia requested allocation of additional funds for the SAIGE Project to support artificial intelligence (AI) development in Serbia in support of Serbia's commitment to AI development, recognizing the importance of AI, digitalization, and robotics in various sectors, including the economy, healthcare, public services, the environment, etc. specifically focusing on the legal framework, enhancing workforce skills in AI, and improving the regulatory environment, coverage of AI topics in education and research, and the availability of computing power.

Serbia's high-level plan is to support AI programs/activities by allocating additional funds that will focus on three main areas: green development based on IT (*green transition*), transport systems and the auto industry (*e-mobility*), and the application of AI in clinical medicine (*clinical healthcare*).

This plan also includes support for Innovation Fund and Science Fund AI programs, capacity building for the AI institute and BIO4, the development of an AI investment promotion platform, activities related to regulations, support to innovation development, hackathons, grant schemes, innovation vouchers, training programs and capacity building among others.

Under the SAIGE AF, the Science Fund (SF) and Innovation Fund (IF) are planning to support enterprises and scientific researchers through grants in deploying artificial intelligence (AI) and cutting-edge technologies to drive innovation and improve services. However, as the adoption of AI increases, so does

¹ <https://projects.worldbank.org/en/projects-operations/project-detail/P170185>

the need to ensure AI systems, processes and applications are deployed according to Trustworthy AI principles.²

2. Objective of the Assignment

This Terms of Reference (ToR) document outlines the scope of services and key responsibilities, qualifications, and application process for the AI Risk Specialist position (hereinafter: the Consultant) to be engaged by the Client and supported by the World Bank.

The Consultant will play an important role in identifying, assessing, classifying and mitigating AI risks and supporting the mitigation of environmental and social risks related to AI systems creation and applications within submitted projects for funding. The primary objective of this role is to ensure the understanding of how AI technologies proposed by the applicant work, so the Client team can ensure these are created, trained and utilized in a manner consistent with Trustworthy AI principles, minimizing potential negative AI-related impacts.

3. Scope of work

- A) The first task will be development of the Risk Assessment Framework (RAF) for AI covering specifically AI-related risks and possible mitigation measures for the Project. The RAF will also be complemented by an “exclusion list” of activities to be excluded from Project financing. This will be done in a manner that is consistent with the basic framework for risk management as stipulated in the Financing Agreement and as expanded in the POM.

The following should be considered for the development of RAF:

- (i) develop the AI framework/guidelines against which the grant project proposals will be assessed and monitored;
- (ii) validate that all financed projects and activities under the Project include a data protection impact assessment and that applicable data protection law and good international practice;
- (iii) validate that any AI activity to be implemented under the Project that is designated to be high risk includes an AI risk impact assessment, to ensure compliance with Trustworthy AI principles;

² “Trustworthy AI” has the meaning assigned to it in the Project Loan Agreement. That definition is reproduced in the Appendix, which is incorporated by reference into these ToRs as defined under the Project means “collectively, the principles and requirements for robust and accountable AI systems, processes and applications throughout the lifecycle of the AI based on Serbia’s Ethical Guidelines for Development, Implementation and Use of Robust and Accountable Artificial Intelligence, including the following: (a) Explainability and verifiability; (b) Dignity; (c) Prohibition to cause damage; (d) Fairness and non-discrimination; (e) Human agency, oversight, determination and control; (f) Technical reliability, safety and security; (g) Privacy, personal data protection and data management; (h) Transparency; (i) Diversity, non-discrimination and equality; (j) Social and environmental wellbeing; (k) Accountability and responsibility. Additionally, “Trustworthy AI” principles incorporate the following principles, as included in other good practice frameworks recognized by Serbia including the “UN General Assembly Resolution on Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development” A/78/L.49 (UNGA Resolution) and the “UNESCO Recommendation on the Ethics of Artificial Intelligence” SHS/BIO/PI/2021/1 (UNESCO Recommendation), including, inter alia: (l) Proportionality and do no harm; (m) Due process; (n) Freedom of expression; (o) Sustainability; (p) Awareness and literacy; and (q) Multi-stakeholder and adaptive governance and collaboration.

- (iv) validate that any SF AI Research Grant or IF AI Matching Grant implements Trustworthy AI principles and in accordance with the AI exclusion list.

The Consultant will also provide capacity-building activities on AI-related issues, including info day sessions for interested applicants to open SF and IF programs and calls, and provide tailored trainings for the technical and administrative staff on AI-related topics, including BIO4 and AI Institute staff.

B) Risk Assessment Framework implementation

The Consultant will work with the PIU team in the following tasks:

- Support integration of the RAF in the grant verification and selection process of the SF and IF, including via:
 - conducting an initial assessment of the SF and IF grant project proposals/applications selected for the second phase of evaluation related to AI-grant programs and provide initial AI categorization;
 - assess any supported AI activity according to the RAF, and based on the assessment findings, estimate the risk level associated with the activity;
 - for high-risk categorization of AI projects, support the preparation of an AI risk impact assessment;
 - support and cooperate with the NITRA/PIU team and the Bank team as required on AI risk activities.

C) Support to monitoring and enforcement of RAF

Based on the above risk assessment, the Consultant will collaborate with the PIU, SF and IF teams in developing and implementing adequate risk mitigation measures in a manner consistent with Trustworthy AI principles and issues that AI can create, using evidence, logic, and reasoning. By taking recognized risks into account, the proposed measures will aim at improving AI systems to minimize inherent risks.

When developing his/her recommendations on mitigation measures, the AI Specialist will ensure that such recommendations are consistent with Trustworthy AI principles as applicable in the given sectors and the World Bank (WB) Project Operational Manual (POM) and Grant Manuals (GM) developed for the Project. The Consultant will perform online or onsite monitoring reviews of the Grantee's activities to ensure compliance with Trustworthy AI principles and applicable requirements, regulations, and standards. In the case of non-compliance with the RAF during grant implementation, the Consultant will determine and suggest remedial actions, prepare non-compliance reports, and keep track of respective responses/actions undertaken.

The Consultant shall provide support to NITRA in contributing to an improved regulatory environment for AI, and perform other duties as required by the PIU Project Manager, including but not limited to support in preparing documents, participating in meetings, and advising colleagues where relevant and requested.

4. Qualifications and Experience:

- Minimum Master's degree in Computer Science, Engineering, Data Science, Law, or a related field;

- Proven experience in data governance, AI risk assessment and management (including principles of AI systems work, their strengths and limitations, and optimization of their performance and accuracy);
- Working experience related to features of Trustworthy AI principles, as well as relevant laws, policies and procedures;
- Working experience with international guidelines and best practices with respect to AI governance;
- Knowledge of English is mandatory; knowledge Serbian language will be considered an advantage;
- Experience in the technical aspects of AI, such as data science, machine learning, natural language processing, computer vision, and/or robotics;
- Working experience in data quality and bias mitigation would be considered an advantage;
- Working experience with national data protection legal frameworks;
- Previous experience with integration of ethical principles and guidelines into AI design, training and development stages;
- Experience in applying the Human-in-the-Loop (HITL) Control principle and how to implement human oversight and intervention capabilities in AI systems to help maintain control and accountability;
- Experience in applying principles of transparency and explainability in AI systems;
- Implementation of data biases risks, and the ability to advise on improving fairness in AI algorithms, to avoid discrimination, would be considered an advantage;
- Experience in assessing risks to personal data posed by AI systems as well as on information disclosure and public engagement in the creation of AI, would be considered an advantage;
- Awareness and knowledge of the environmental impacts of AI, such as energy consumption, carbon footprint, waste generation, biodiversity loss, etc;

5. Selection criteria

The candidates will be evaluated by applying the following evaluation criteria:

- Knowledge and General experience (40 Points)
- Specific Experience relevant to the Assignment (60 Points)

6. Reporting Obligations

The Consultant will collaborate with other PIU team members and will report to the PIU Project Manger.

7. Terms of Engagement

The Contract will be the Standard World Bank Time-Based Contract for Individual Consultants, entered into between the Consultant and the NITRA, as a Client.

The Consultant will be engaged on a part-time basis, until September 30, 2025. The expected time effort for the assignment is approximately 100 working days (including weekends when needed). The Consultant agrees to be engaged in no more than 48 hours per week cumulatively for this assignment plus any additional assignments/contracts.

The payments for services will be based on monthly timesheets approved by the PIU Project Manager. The PIU will administer the Contract.

8. Input by the Client

The Client will facilitate access to relevant documents as well as communication with appropriate counterparts in a time and manner convenient to both parties.

9. Confidentiality and Conflict of Interest

The Consultant will maintain confidentiality on all information that is not in the public domain and shall not be involved in another assignment that represents a conflict of interest to the prevailing assignment.