

**UNIVERSITAS STUDIORUM FLUMINENSIS
UNIVERSITY OF RIJEKA**

uniri



**ARTIFICIAL INTELLIGENCE TOOLS USAGE POLICY
AT THE UNIVERSITY OF RIJEKA, CROATIA**

Rijeka, January 2024

Pursuant to Article 34, Paragraph 1 of the Statute of the University of Rijeka (CLASS: 030-01/23-01/05, Reg. No.: 2170-137-01-23-2, the consolidated text from October 2, 2023), the University of Rijeka Senate at its 91st session held on January 23, 2024, adopted the following

ARTIFICIAL INTELLIGENCE TOOLS USAGE POLICY AT THE UNIVERSITY OF RIJEKA

INTRODUCTION

Artificial intelligence (AI) is one of the key factors in the transition towards a digital society and economy. It is rapidly developing and altering all the aspects of life and business, creating the prerequisites for the improvement of the quality of life, as well as for global prosperity and economic development, while also posing considerable challenges – which is all well documented by the Council of Europe. AI is a very adaptable technology that takes on different forms and new applications. Universities are thus obliged to monitor the developments of AI, enable its responsible and appropriate usage, but also to raise awareness of possible challenges caused by this disruptive and transformative technology.

Science and higher education are invited to encourage the usage of available and the development of new AI tools and systems which will be compatible with the **United Nations Sustainable Development Goals (UN SDGs)** that, together with digitalization and a resilient single market, constitute some of the fundamental principles of the **European Union (EU) Treaties** and of the priority goals of the EU policies. The necessity of systematizing existing knowledge about the way AI systems can contribute to the achievement of the sustainable development goals is particularly emphasized, as is the systematic approach to ensuring the sustainability of the AI systems themselves.

Supporting scientific progress and knowledge dissemination by adopting innovative technologies and models for their responsible application, **the University of Rijeka advocates a responsible, ethical, open, transparent, and innovative approach to the usage of AI tools and other advanced digital technologies.** In this framework, **the University of Rijeka Council of Honor**, at its session held on April 18, 2023, issued a *Statement on the responsible usage of artificial intelligence tools*, emphasizing the need for a constructive discussion about a responsible, ethical, and transparent usage of AI tools with the staff and students of the University, and welcoming all institutional and individual efforts towards the responsible usage of AI tools in learning and teaching, including their permitted, critical, and informed usage while respecting the highest ethical principles. Following this statement, **in this policy document, adopted by the Senate, the University of Rijeka further elaborates the basic terms, defines the aims of AI tools usage in teaching and research, states the fundamental principles of their usage at the University, defines the activities of the University in this area as well as the stakeholders involved in these activities.** Hence, the University expresses its determination to maximally use the advantages and potentials of new AI-based technologies, while at the same time taking into account the highest ethical standards by promoting continuous education and developing critical thinking of all included stakeholders.

Given the accelerated growth and development and the technological evolution of AI and AI-based digital and other technologies, the University of Rijeka shall revise and update its AI Tools Usage Policy regularly, at least once a year.

GOALS OF AI TOOLS USAGE AT THE UNIVERSITY OF RIJEKA

AI-based tools (cf. list of basic terms given at the bottom of this document) are computer programs or systems that use techniques and algorithms from the AI field to perform tasks that usually require human intelligence. Examples of AI tools include machine learning models, chatbots, big data analytical tools, image and speech recognition software, data management systems, etc. AI tools play a major role in process automation, productivity improvement, and providing solutions to complex tasks that require analysing and reasoning based on large volumes of data. These tools can perform various tasks, including pattern recognition, natural language processing, planning, decision making, recommending, data analysis, and many more, which is why they can be used in sectors such as healthcare, finance, manufacturing, transportation, marketing, education, etc. Depending on their complexity and purpose, AI tools can be implemented as a standalone program, or they can be integrated into larger applications or systems.

Thus, under the supervision and expert guidance of trained researchers and teachers and through their responsible and ethical application, AI tools usage is increasingly expanding, whereby they assist and accelerate processes and procedures in education, as well as in research and innovation (R&I).

AI tools are already widely used in teaching, especially for:

- Designing and developing the structure and content of curricula for individual courses or entire study programs;
- Preparing exam questions;
- Using chatbots and other generative AI tools as virtual assistants in class (e.g. for personalized help, answering questions or identifying learning resources and materials), with the development of appropriate digital twins;
- Writing student papers, whereby it must be insisted that students state and cite the AI tools they used and for which parts of the paper they were used, and that the parts created this way are used specifically for the development of critical thinking and critical questioning of the reliability of information as well as of the results generated by AI systems and tools.

AI tools are a key enabling technology that is extensively used in R&I to help with:

- Mining, analysing, summarizing, and synthesizing information and large volumes of data from pieces of literature or in their translation;
- Brainstorming of ideas, pattern recognition and correlation identification, and recommendation of appropriate work methodologies;
- Writing and editing (style, grammar, proofreading) of scientific papers and project applications, especially for non-native English-speaking researchers;
- Reviewing papers and projects;
- Writing program codes, generating datasets, images, visualizations, simulations, 3D models, etc.;
- Optimizing and automatizing experimental settings;

- Optimizing the elements of complex systems or automated classification (e.g. of artistic compositions);
- Providing medical diagnoses, designing proteins, encoding genetic sequences, searching for solutions for antibiotic resistance, and many other applications in biomedicine and healthcare;
- Improving weather forecasts and developing innovative materials;
- Facilitating and accelerating administrative tasks;
- Communicating and disseminating scientific results.

The advantages of using AI tools in teaching and research are as follows:

- Increased efficiency, productivity, and application of innovative methods in scientific research and teaching: AI tools can facilitate routine tasks such as searching for relevant sources of information in the given field, as well as finding works via advanced recommendation systems and their preliminary comparison and analysis, which increases the productivity of educational and research processes and allows students and teachers to focus on more creative and more complex tasks. What is more, generative AI tools can help suggest adequate wordings for scientific papers or project applications, as well as edit the respective texts. In fact, many scientific journals often already have a published policy on the AI tools usage, whereby they approve their use for improving the text, for creating and/or modifying images, graphs, and visual elements, and for managing bibliographic references - all with a mandatory indication of the activities and parts of the content of the papers for which AI tools were used, while the authors themselves are responsible for the final content, especially for the scientific statements and conclusions of the papers. On the other hand, publishers as well as research funding organisations would have to clearly and openly declare when they use AI tools in the selection of projects or papers or when reviewing them.
- Rapid analysis of large volumes of data: AI can analyse and draw conclusions from large datasets in different scientific domains and fields quicker and more accurately than humans. In fact, algorithms of machine learning and deep learning enable pattern recognition in large volumes of data, data classification and data clustering, predictions, and other automated processes which can accelerate scientific analyses and lead to discoveries and insights. In this framework it is important to use appropriate unbiased datasets that will not lead to the development of wrong models and false conclusions. Obviously, the datasets should also be openly available and structured in accordance with the FAIR principles of open science (Findable, Accessible, Interoperable, and Reusable). Since all of the above is useful in various fields, it should enable teachers and students to dedicate more time to creative interpretation rather than routine data processing.
- Personalization and recommendations: AI can provide targeted and personalized recommendations and services based on users' preferences and browsing history, which improves the quality of responses and aids the specificity of the results.
- Solving complex problems: AI tools can solve complex problems in fields such as medical diagnostics, research, and engineering, encouraging an interdisciplinary and team approach to problem solving.
- Continuous availability: AI chatbots and virtual assistants can provide services and support to users throughout the day and every day of the week, opening up the potential for the improvement of numerous services.

FUNDAMENTAL PRINCIPLES OF AI TOOLS USAGE AT THE UNIVERSITY OF RIJEKA

In addition to these great benefits of AI tools usage, there are also some unquestionable challenges and risks linked to their application, such as:

- Relying on repetitive patterns without understanding them, and lack of transparency of the obtained information and results;
- Increased possibility of plagiarism and fabrication and manipulation of the results of scientific papers;
- Inability to repeat the results (the reproducibility issue);
- Absence of links to the sources of the used information;
- Biases and mistakes (hallucination) in data and algorithms;
- Possibility of replicating false information and counterfeits, consequently eroding the public's trust in science and higher education;
- Unresolved issues related to intellectual property and copyrights;
- Issues related to possible violations of privacy and data safety;
- Possibility of prolonging regional and national inequalities in the availability of scientific infrastructure due to the costs and energy requirements related to the development and usage of AI systems and tools, but also due to the availability of data for the training of AI models.

Intending to maximize the advantages of AI tools usage, and at the same time minimizing the potential risks of their application, the fundamental principle is that the human, i.e., the teacher and the student of the University who uses the AI tools, is responsible for AI usage and the quality and reliability of the thus obtained results. This means that the AI tools users themselves are responsible for the scrutiny and interpretation of data and results obtained by using these tools, for the authorships of the papers, of seminars, and of other educational and scientific publications, as well as for their reviewing and for the wording of scientific hypotheses and research goals.

The usage of AI tools at the University of Rijeka must therefore be based on:

- the principles of fairness, equity, honesty, and accountability, which also implies democratic participation in the development and usage of AI tools, hereby comprising the examination, debate, control, and users' feedback;
- compliance with the laws, bylaws, regulations, and standards regulating the subject matter, in particular, the Harmonised Rules on Artificial Intelligence, i.e., the Artificial Intelligence Act (AI Act) Proposal of the European Commission adopted in December 2023, and the preservation of fundamental rights;
- personal responsibility, integrity, conscientiousness, and commitment to the truth of the users – teachers and students;
- the principles of ethical and reliable use focused on people and their values;
- accuracy and technical (including cyber) security and robustness, while avoiding bias using verifications based on reliable samples, testing appropriate to the purpose of use, standardized certification of AI tool reliability, and traceability of AI models and systems;
- manageability and external independent and impartial control with the possibility of contestation of AI tool results;
- research, professional, and teaching work that strives for excellence based on:
 - current and future safety and sustainability of life, health, and well-being of people, the environment, and property (including care for cultural, historical, and archaeological

- heritage) as well as of the social needs and interests;
- the highest standards of expertise;
- the principles of academic integrity and the integrity and dignity of individuals – researchers and/or research subjects and research profession;
- human and scientific autonomy, objectivity, accuracy, impartiality, modesty, and dignity;
- transparency (including transparency and traceability of funding sources) and openness concerning the data sources used to train the AI models and how the AI systems make decisions;
- protection of personal and confidential data with informed consent to research and participation in it;
- protection of privacy, intimacy, and social, cultural, and linguistic diversity, including (but not limited to) the prohibition of the use of AI systems and tools for social ranking, automatic behavioural profiling, and mass surveillance;
- inclusiveness, non-discrimination, and solidarity among people and generations, including the minimization of differences in the availability of resources and infrastructure for the use of AI tools;
- enabling reproducibility and explainability of the procedures used, which in some cases implies the consistent need to determine the honesty and morality of the used AI algorithms;
- fostering openness of innovation and knowledge sharing;
- establishing instruments for raising AI literacy in society;
- zero tolerance for all forms of fraud and plagiarism;
- striving for sustainable economic, social, and environmental development (i.e., without present or future biological, physical, or moral danger to humans, animals, plants, the environment, and resources);
- publishing, editing, and reviewing/assessing based on:
 - integrity in determining and citing all authors of the paper, which implies the need for each use of AI tools to be indicated clearly, to make clear in every paper which parts were obtained by using AI and which sources they are based on, and to critically evaluate the data and results obtained in this way;
 - objectivity, honesty, argumentation, and confidentiality;
- other principles of ethics that have to include:
 - the issue of professional responsibility concerning the public and the profession itself, which includes care for further development, improvement, and dissemination;
 - intellectual property issues;
 - questions of the legitimacy of research results;
 - avoidance of conflict of interest;
- objective and truthful communication and raising awareness and building public trust in the results of the use of AI tools while encouraging synergy and cooperation of all interested stakeholders, including the promotion of AI expertise in public and government services;
- a clear, explicit, and irrefutable definition of responsibility, which in no case can be transferred from a human to an AI machine or tool;
- international cooperation and coordination, especially in the field of AI standardization and regulation;

- the development of risk assessment methods and the truthful public presentation of these risks;
- if necessary, the precautionary principle (no-safety – no usage);
- precautionary measures including:
 - o protection measures and sufficiently long supervision of the impact on health, public safety, and the environment;
 - o anticipation of the possible negative impacts of AI tools usage and the use of proper avoidance and mitigation measures;
 - o the requirement that each research funding application also contains a risk assessment with a ban on research funding that could lead to violations of fundamental rights and ethical principles.

ACTIVITIES IN THE FIELD OF AI AT THE UNIVERSITY OF RIJEKA

The University of Rijeka will:

- Encourage the development of human resources in the field of AI, especially through workshops and other structured and supported educational activities that will develop digital and information literacy of teachers and students with the aim of an effective, responsible, and ethical use of AI tools, with the related development of critical thinking and highlighting the need to adhere to the above-defined principles of AI tools usage. The framework for literacy in the field of AI will include the basic software and hardware tools, the techniques for the evaluation of the results obtained via AI tools, communication with AI systems and chatbots, the usage of AI tools in cooperation with others, the creation of content by using AI tools, security, legal and ethical aspects of using AI tools, as well as specific knowledge for certain domains of application of AI tools.
- Encourage attracting researchers with expertise in the field of the use of AI tools, among others by using the model of returnees of expatriate scientists.
- Encourage the development of teaching and assessment methods that use AI as well as the corresponding development of innovative and interdisciplinary curricula and courses, while taking into due account the necessary critical thinking and understanding.
- Encourage, also via an active support to the activities of the Center for Artificial Intelligence and Cyber Security, the activities of the EDIH Adria project, and support the maintenance and further development of the research, in particular digital, infrastructure of the University, the R&I activities in the field of AI and the related interdisciplinary fields, especially the projects and the implementation of AI tools in cooperation with stakeholders from the public and business sectors, all aimed also at the consequent development of the regional innovation ecosystem.
- Encourage the establishment of startup and spin-off companies of teachers and students of the University based on the application of AI.
- Monitor the impact of AI on the development of research careers, especially those of young scientists (doctoral and postdoctoral students), as well as on academic integrity.
- Insist on a flexible approach and find an appropriate balance between the use of new digital technologies and the responsible and ethical use of AI tools. Indeed, these tools have to be seen as an aid to research and teaching, not as a substitute for teachers and researchers. In this

framework, teachers are allowed to choose whether they want to use AI tools in teaching and to what extent the employment of AI tools contributes to achieving or improving the learning outcomes they are aiming at.

- Following the above principles of using AI tools, support the development and implementation of the ethical and legal frameworks as well as the quality standards for the application of AI, while encouraging an open and transparent public debate on the benefits and challenges of using AI.
- Based on the principles of open science and open innovations, advocate the availability of large (especially locally relevant) volumes of data under the authority of state and public administration to the academic research groups working on R&D in the field of AI, and advocate the development of public digital infrastructure.
- Actively foster international cooperation and a global inclusive access to AI and associated infrastructure in line with the fundamental human rights and the United Nations Sustainable Development Goals, especially within the university networks (YUFE, YERUN, and others) in which the University is an active stakeholder.
- Establish an Advisory Body that will monitor the development of AI tools and their implications for the teaching and research processes at the University. This Body will also regularly propose a revision and update of the present AI Tools Usage Policy at the University of Rijeka to the University Management and the Senate. The Advisory Body will be appointed by the Rector, and it will consist of:
 - the Vice-rector in charge of digitalization (President),
 - the Vice-rector in charge of innovation (Vice President),
 - the Head of the Center for Artificial Intelligence and Cyber Security,
 - a representative of the University Library who is an expert in the field of open science,
 - an expert on the issues of the ethical use of AI (e.g. the Head of the EDIH Adria Ethics Advisory Board),
 - a representative of the University Council of Honor,
 - two experts in the field of AI,
 - an employee of the University Rectorate in charge of the administrative support to the activities of the Advisory Body,
 - If necessary, representatives of stakeholders and experts from outside the university community can also be included in the membership.

STAKEHOLDERS OF THE AI TOOLS USAGE AT THE UNIVERSITY OF RIJEKA

Stakeholders of the AI tools usage in the above-mentioned frameworks are the University of Rijeka management, the management of the University constituents, teaching staff and researchers, administrative and professional staff, students of the University, and partners of the University in the regional innovation ecosystem as well as international partners of the University. All these stakeholders should take care in their scopes of a responsible, ethical, open, and transparent usage of AI tools, with a strict identification of the content for which the AI tools were used and with a clear responsibility of the users themselves for the final results of the performed activities.

The **University of Rijeka management** directs and supervises the activities in the field of AI at the University as defined by this policy document and following the principles defined above, takes

care of the development of these activities and principles, and, through the described Advisory Body, scrutinises and proposes updates and revisions to this Policy to the University Senate.

The **management of the University constituents** proposes to their faculty or academic councils the adoption of strategic guidelines and policies for the usage of AI tools at their institutions in accordance with the scientific and artistic fields in which they operate and directs and supervises the operations related to those documents.

The **teaching staff and researchers of the University** actively participate in workshops and other educational activities to build and improve their skills in using the AI tools and use these tools in teaching and research, while observing the principles of their use as defined by this Policy and the relevant documents of their constituents.

The **administrative and professional staff of the University** actively improve their skills in using the AI tools and use them in accordance with this Policy and the strategic guidelines of their constituents.

The **University of Rijeka students** actively educate themselves on the AI tools usage and use them in their study and other activities with a consequent application of the principles defined by this Policy and taking into account the recommendations of the constituents where they study.

The **partners of the University in the regional innovation ecosystem from the economy, the local or regional government, the civil sector, the NGOs, and the wider community** cooperate with the employees and students of the University on the implementation of activities that include AI tools usage, thereby contributing to the regional, national and international economic and social development, also based on the principles of the quintuple innovation helix framework, of smart specialization, and of a two-way and circular knowledge transfer. The employees and students of the University also actively cooperate with the **international partners of the University** in such activities, especially within the framework of university networks (YUFE, YERUN, and others) in which the University is an active stakeholder.

BASIC TERMS

Artificial intelligence (AI)

- The relevant references do not contain one generally accepted and comprehensive definition of the term artificial intelligence. There are very simple definitions that mention "intelligent machines" or "machines that function in a way that is considered intelligent", and more elaborated definitions that it is a scientific branch that encompasses computer technologies inspired by the way people use their nervous system and bodies to feel, learn, reason, and function. Most of the definitions agree that it is about machines and software that, for a human-given set of objectives and with various levels of autonomy, adopt and improve the execution of their tasks based on the capability to generate outputs such as content, predictions, recommendations, or decisions which influence the environment they are physically or digitally interacting with. This includes machine learning approaches (including supervised, unsupervised, and reinforced learning using various methods, including deep learning), logic and knowledge-based approaches (including knowledge representation, inductive (logical) programming, knowledge bases, inductive and deductive (symbolic) reasoning modules and expert systems) and statistical approaches, Bayesian evaluation, as well as search and optimization methods.

AI alignment

- It is focused on ensuring that AI systems are structured, developed, and trained in accordance with the aims, preferences, or ethical principles of humans. The main aim of AI alignment is to ensure that AI systems understand and function in line with what humans want or think is ethical. This primarily refers to the alignment of AI systems with human values, including ethical principles, and it includes the development of AI systems that make decisions and take action in accordance with the moral and ethical guidelines that society has established. AI alignment is crucial for the building of AI systems, which are not only technologically advanced but also safe, reliable, and useful for society.

Artificial intelligence and education (AI&ED)

- It encompasses different links between AI and education that include concepts such as "learning with AI", "learning about AI" and "preparation for AI". The concept of learning with AI is also called artificial intelligence for education.

Artificial intelligence in education (AIED)

- The academic research field, conceived in the 1980s, which primarily studies AI tools that facilitate support to learning activities (learning with AI).

AI-based systems

- Software or hardware systems based on technologies and algorithms from the AI field. They use, for example, machine learning, deep learning, natural language processing methods, and similar analysis approaches for large volumes of data. They enable learning from data, drawing conclusions, and generating new knowledge as well as making decisions. AI-based systems are applied in various domains of human action, such as medicine, engineering, finances, logistics, education, security, space research, and others. These systems are also called AI systems or intelligent systems.

AI tools

- The name used for all tools using AI technology. The focus is currently on tools based on generative artificial intelligence that enable the creation of new content, such as, for example, texts in a given style, programming codes, visualizations and realistic images, artwork, music, videos, 3D models, simulations, etc.

Artificial general intelligence (AGI)

- A level of AI capable of understanding and applying knowledge in a wide spectrum of different tasks. AGI can function in any domain and cope with different tasks without the need for special adjustments or programming for each new task. Such systems can understand, learn, apply knowledge, and make decisions in a way that is similar to human intelligence.

Artificial narrow intelligence (ANI) (narrow AI, specialized AI)

- The term refers to an AI type specialized in performing specific domain tasks, such as face recognition, language translation, playing chess, or data analysis. Narrow AI is specialized and specifically aimed at solving one or a few related problems and it does not have the ability of general understanding or knowledge application outside these specific tasks. ANI systems can solve tasks based on previously programmed rules or based on large volumes of data. However, such systems lack the understanding of the deeper context and aims of those tasks.

Bias

- Bias is the preference towards a preconception in favour of or against a person, a thing, or an

attitude. Although it is expected that systems based on AI are logical and correct, they can be biased as well. In fact, machine learning algorithms reflect the bias of the datasets on which they are trained, so that they can adopt and express biased solutions and attitudes that often violate ethics. For example, language models can generate texts that contain prejudice, stereotypes, and discriminatory expressions, which is a consequence of learning from internet content that can contain such elements. Machine translation systems can, in turn, transfer prejudice and bias from one language to another, which can result in inappropriate and offensive translations.

Big data

- Large heterogeneous datasets of different origin that rapidly change and grow. Big data analysis is arranged so that people without computer programs can conduct it. Machine learning algorithms and deep neural networks are used in this framework, enabling to find data patterns and relations as well as deriving new knowledge.

Chatbot

- A computer program, i.e., an application that can communicate with the users via a textual conversation. It is used for answering questions and providing information in various domains. Besides the systems developed for more specific, narrow domains, such as, for example, providing information about products or services, chatbots like ChatGPT and Google Bard, which can converse about any topic, are also being developed.

Deep learning

- A machine learning type based on applying multi-layer neural networks. Deep learning algorithms enable the learning and finding of features of large datasets, i.e., the recognition of the occurring patterns. Deep learning is applied in different tasks, such as, for example, content recommendation, diagnostics based on medical images, genetic data analysis, risk identification of certain diseases, new material development, etc.

Explainability

- In AI, explainability is defined as the ability to understand and interpret decisions, behaviours, and results generated by AI systems, such as machine learning or deep learning models. Explainability is crucial in the context of AI because it enables users to understand how and why a certain model makes a decision. Explainability is a term similar to interpretability but, while interpretability refers to the understanding of the inner functioning of an AI model, explainability includes providing explanations for the decisions made by the model that are understandable to humans. An interpretable model enables the understanding of how a model makes a certain decision, whereas an explainable model provides explanations of why a model made a certain decision. Explainability is an important aspect of the ethical development and application of AI, especially in situations where AI systems make decisions important to people.

Generative AI

- A branch of AI that automatically creates content such as text, images, video, program codes, music, and others, in response to a textual prompt via a suitable interface. Generative AI uses deep learning algorithms and neural networks to create new content based on existing datasets. Generative models can create diverse and creative content such as, for example, texts in different styles, high-quality images and photographs on a given topic, realistic simulations, or music that sounds like it was composed by a human.

Large language models (LLM)

- Models based on AI that can present and generate new text. For the development of large language models, deep neural networks trained on large amounts of text are applied, during which tens of millions or even hundreds of millions of parameters are used. Based on the large amounts of textual content, models can learn how to generate complex and diverse texts. Such models have also proven to be good in solving tasks in the field of natural language processing, such as machine translation, text summarization, sentiment analysis, text classification, named entity recognition, etc. Examples of such models are GPT-4 or BERT.

Machine learning (ML)

- A branch of AI that deals with the development of learning algorithms based on given datasets. Machine learning algorithms analyse data based on certain properties of that data (a set of features) and solve different tasks without being explicitly programmed for that. The basic idea of machine learning is that computer systems can automatically recognize patterns in data and use them to make decisions and solve problems.

Model training

- A term used in the fields of AI and machine learning for the processes that enable the algorithm or model to learn from data so as to perform certain tasks or make decisions. During the training process, the model learns to recognize patterns in the data and adjust its parameters to achieve the best performances.

Reproducibility

- In AI, the term refers to the possibility of other researchers to reproduce and confirm the research results and experiments conducted by using AI. In the context of machine learning, it is particularly important that the experiment can be repeated so that other researchers can use the same data, models, and methods to verify the results. This is crucial to ensure data validity and reliability and to promote transparency and trust in research based on the use of AI.

Virtual Assistant

- A computer program or an application that is mostly based on AI and provides support for users in performing different tasks, including search prompts, setting reminders and terms, following information, etc. They can access and search through large volumes of data to provide the users with information and solutions. They often provide personalized information and recommendations based on the users' interaction history. They are similar to chatbots, but they often have an expanded functionality and can multicast with the user. Virtual assistants can communicate with users via different interfaces and communication channels, including textual, voice, and visual interactions.

ESSENTIAL BIBLIOGRAPHIC REFERENCES

Bockting, Claudi L.; van Dis, Eva A. M.; van Rooij, Robert; Zuidema, Willem and Bollen, Johan (2023). Living guidelines for generative AI — why scientists must oversee its use. *Nature*, 622(7984), 693-696. doi: 10.1038/d41586-023-03266-1

Brown, Tom; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared D. et al. (2020). Language models are few-shot learners. in „Advances in Neural Information Processing Systems“, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan and H. Lin (ed.), 33, 1877-1901.

Conroy, Gemma (2023). How ChatGPT and other AI tools could disrupt scientific publishing. *Nature*, 622(7982), 234-236. doi: 10.1038/d41586-023-03144-w

Council of Europe on Artificial Intelligence, <https://www.coe.int/en/web/artificial-intelligence>, in particular <https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6> with the related Council documents and decisions, assessed on January 5, 2024.

Devlin, Jacob; Ming-Wei, Chang; Kenton, Lee and Kristina Toutanova (2018). Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805.

Elsevier: „The use of generative AI and AI-assisted technologies in writing for Elsevier“ - <https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier>, assessed on December 26, 2023.

European Commission – DG for Research and Innovation (2023). AI in Science: Harnessing the power of AI to accelerate discovery and foster innovation – Policy Brief. European Union. doi: 10.2777/401605

European Commission High-level expert group on artificial intelligence (2019). Ethics guidelines for trustworthy AI.

European Commission (2021). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM(2021) 206 final, 2021/0106 (COD).

European Parliament - Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies (2019). State of the art and future of artificial intelligence – Brief.

European Research Council (2023). Foresight: Use and impact of Artificial Intelligence in the scientific process. ERC Executive Agency. doi: 10.2828/10694

European Research Council (2023). Current position of the ERC Scientific Council on Artificial Intelligence.

European University Association (2023). Artificial intelligence tools and their responsible use in higher education learning and teaching.

Glenster, Ann Kristin and Gilbert, Sam (2023). Policy brief: Generative AI. Bennett Institute for Public Policy.

Hanson, Brooks; Stall, Shelley; Cutcher-Gershenfeld, Joel; Vrouwenvelder, Kristina; Wirz, Christopher; Rao, Yuhan and Peng, Ge (2023). Garbage in, garbage out: mitigating risks and maximizing benefits of AI in research. *Nature*, 623(7985), 28-31. doi: 10.1038/d41586-023-03316-8

Holmes, Wayne; Persson, Jen; Chounta, Irene-Angelica; Wasson, Barbara and Dimitrova, Vania (2022). Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law. Council of Europe.

Meštrović, Ana (2023). Umjetna inteligencija u otvorenoj znanosti – novi izazovi pred istraživačima [AI in open science – new challenges for researchers]. Lecture at an Open Science Café organized by the University Library having as topic „AI in the context of Open Science“, December 8, 2023. -

available on <https://urn.nsk.hr/urn:nbn:hr:195:319846>, assessed on December 24, 2023.

Ministry of Economy, Entrepreneurship and Crafts (2019). Nacrt Nacionalnog plana za razvoj umjetne inteligencije [Draft of National Plan for the Development of Artificial Intelligence].

Montréal Declaration for a Responsible Development of Artificial Intelligence (2018) - available on <https://montrealdeclaration-responsibleai.com/the-declaration/>, assessed on December 24, 2023.

OECD (2023). Recommendation of the Council on Artificial Intelligence.

OpenAI (2023). GPT4 technical report.

Parrilla, Juan Manuel (2023). ChatGPT use shows that the grant-application system is broken. Nature, 623(7986), 443. doi: 10.1038/d41586-023-03238-5

Partnership on AI - <https://partnershiponai.org/>, assessed on December 25, 2023.

Russell, Stuart J. and Norvig, Peter (2020). Artificial Intelligence: A Modern Approach – 4th ed. Pearson.

UK Government – Department for Science, Innovation and Technology (2023). Policy paper: The Bletchley Declaration by Countries Attending the AI Safety Summit.

UNESCO (2023). Guidance for generative AI in education and research.

University of Cyprus Recommendations for Use of Artificial Intelligence in Educational Processes (2023).

University of Rijeka Council of Honor(2023). Priopćenje o odgovornoj uporabi alata umjetne inteligencije [Statement on the responsible use of artificial intelligence tools] - available on <https://uniri.hr/en/about-university/structure/council-of-honor/>, assessed on December 24, 2023.

USA Whitehouse (2023). FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.

Van Noorden, Richard and Perkel, Jeffrey M. (2023). AI and science: What 1,600 researchers think. Nature, 621(7980), 672-675. doi: 10.1038/d41586-023-02980-0

World Economic Forum (2023). Generative AI has disrupted education. Here's how it can be used for good – UNESCO.

World Economic Forum (2023). The Presidio Recommendations on Responsible Generative AI.

Zelenika, Saša (2011). Etička načela na području tehničkih (inženjerskih) znanosti, tehnologija te prirodnih znanosti [Ethical principles in the field of technical (engineering) sciences, technologies, and the natural sciences]. University of Rijeka.



University of Rijeka, Croatia