COURSE UNIT DESCRIPTION

Course unit title	Code
Ethical artificial intelligence for social good	

Lecturer(s)	Department, Faculty
Coordinating: Assoc. Prof. Dr. Valentas Gružauskas **Other:** Assoc. Prof. Dr. Miroslavas Seniutis Neringa Gaubienė	Faculty of Mathematics and Informatics Faculty of Philosophy Faculty of Law

Study cycle	Type of the course unit
First	Interdisciplinary studies (Individualised studies)

Mode of delivery	Semester or period when it is delivered	Language of instruction
Face-to-face	Autumn semester	English

Requisites	
Prerequisites: ability to read English at a minimum as B2 level according to the Common European Framework of Reference for Languages.	**Co-requisites (if relevant): -**

Number of ECTS credits allocated	Student's workload (total)	Contact hours	Individual work
5	135	66	69

Purpose of the course unit: programme competences to be developed		
The purpose of the "Ethical Artificial Intelligence for Social Good (EAI4SG)" course is to equip students with a multidisciplinary understanding of artificial intelligence (AI) applications in societal contexts. The module aims to develop students' capabilities to create and evaluate AI solutions that promote social good, integrating ethical, legal, and technical perspectives. It focuses on applying AI in ways that are economically viable, environmentally sustainable, and socially equitable, enhancing students' ability to tackle real-world challenges aligned with sustainable development goals.		
Learning outcomes of the course unit	**Teaching and learning methods**	**Assessment methods**
Will know the concepts related to the research field of Ethical Artificial Intelligence for Social Good and will be able to apply them by recognizing, defining, and critically evaluating the problems of ethically aligned AI implementation in various social contexts.	Lectures: **problem-based teaching** using interactive digital learning techniques such as real-time polls, Q&A sessions etc. to deepen understanding of AI concepts.	Active participation in lectures and seminars or laboratory work (including collaboration in teams and the stimulation of evidence-based discussions).

Will **be able to find and critically assess, interpret, and systematize relevant information** from national and international primary and secondary sources.	Seminars: **text analysis, group discussion and real-world case studies** to explore practical applications of AI and their ethical and legal implications.	Active participation in lectures and seminars or laboratory work (including collaboration in teams and the stimulation of evidence-based discussions).
Will be able understand and apply an interdisciplinary perspective in solving complex ethical issues related to specific cases of AI implementation, analysing the causes and consequences of these problems. The focus will be on the application of AI, rather than on the fundamental knowledge of algorithms.	Laboratory work: machine learning approaches for classification and regression tasks, including training. Deep learning approaches for image and language processing with a focus on inference.	AI-Driven Mapping (utilize AI-powered software with public access to create a concept map). Presentation and Q&A, feedback and reflection. Classification task based on machine learning approaches and selected datasets, with a focus of sensitivity analysis.
Will **be able to constructively collaborate** in interdisciplinary and intercultural groups, organize the work of groups and teams by empowering communities and building networks of partnerships.	Seminars and individual work: engage in **team-based learning** with a collaborative group project focused on using AI to tackle real-world issues, enhancing both teamwork and interdisciplinary problem-solving skills.	Team project presentation
Will **be able to perceive** and **address complex societal issues** applying principles of human rights, equal opportunities, and ethics.		Team project report
Will **be able to independently organize and plan** their professional activities, reflect on their own and others' achievements, and make decisions adapting to environmental and societal changes.		

Course content: breakdown of the topics	**Contact hours**							**Individual work: time and assignments**	
	Lectures	Tutorials	Seminars	Workshops	Laboratory work	Internship/ work placement	**Contact hours, total**	**Individual work**	**Assignments**
Outlining key paradigms for Ethical AI in Social Good (EAI4SG) research.	2		2				**4**	**6**	Read and prepare to discuss the text: - Josh Cowls (ed.), 2021, Artificial Intelligence for Social Good, Springer Cham, pp. 1-6; 45-65. AI-Driven Mapping - for details, please see section Assestment cryteria

									No. 2;
Exploring AI applications across diverse fields, including industry, governance, healthcare, education, social services, media, and politics.	2		2				**4**	**6**	Read and prepare to discuss the text: - Chui, M., Harryson, M., Valley, S., Manyika, J., & Roberts, R. (2018). Notes from the AI frontier applying AI for social good, pp. 1- 42.
Frameworks for the development process of AIbased technological solutions.	2		2				**4**	**6**	Read and prepare to discuss the text: - European Commission. (2020). AI Watch Assessing Technology Readiness Levels for Artificial Intelligence https://publications.jrc.ec.europa.eu/repository/handle/JRC122014 - Cowls, J., King, T. C., Taddeo, M., & Floridi, L. (2019). Designing AI for social good: Seven essential factors, pp. 1-26.
Introduction into the team project including exploration of team-learning principles and designing of research plan.	4	1	4				**9**	**6**	Read and prepare to discuss the text: - Mahler, S. J. (2012). Teambased learning in social sciences research methods classes. In Team-Based Learning in the social sciences and humanities (pp. 113-128). Routledge.
Legal challenges in the digital society	2					2	**4**	**5**	Team project Presentation and Q&A, feedback and reflection.
Regulating AI Technology	2					2	**4**	**5**	Team project
Privacy of data used for AI training and copyright	2					2	**4**	**5**	
Ownership and usage of AIgenerated content	2					2	**4**	**5**	
Liability for damages caused by AI	2					2	**4**	**5**	
Introduction to Programming/Data Analysis	2				4		**6**	**5**	Classification task based on machine learning Team project
Basics of Machine Learning and deep learning	2				6		**8**	**5**	
Artificial intelligence model explainability	2				2		**4**	**5**	

Artificial intelligence applications with computer vision, large language models	2	1			4		**7**	**5**	
Total	**28**	**2**	**10**		**16**	**10**	**66**	**69**	

Assessment strategy	Weight %	Deadline	Assessment criteria
Active participation in lectures and seminars or laboratory work	15 %	Throughout the semester	Participation in discussions on the literature assigned, active engagement in seminars will be assessed. **Requirements for active participation in discussion:** - Clarity of contributions during discussions (0, 1.25, 2.5 points). - Relevance and depth of comments related to the literature (0, 1.25, 2.5 points); - Consistency of engagement throughout the seminar (0, 1.25, 2.5 points); - Ability to engage others and foster dialogue (0, 1.25, 2.5 points). **Evaluation Strategy:** - 0 points – if the student does not meet the requirement. - 1.25 points – if the student partially meets the requirement. - 2.5 points – if the student fully meets the requirement. Scores will range from 0 to 10 points.
AI-Driven Mapping	15 %	Week 5	During lectures, students are introduced to fundamental concepts. Subsequently, in seminars, they develop an understanding of the AI-driven software's operational principles, which is essential for mapping tasks. This is followed by a comprehensive testing of the program. Ultimately, the students independently execute the search of definitions of concepts and their mapping task using AIdriven program. The task is evaluated with a grade; **Requirements for AI-Driven Mapping:** - Variety and diversity of definitions (0, 1.25, 2.5 points). - Explanation of conceptual relationships (0, 1.25, 2.5 points). - Use of diverse online sources (0, 1.25, 2.5 points). - Accuracy and visual appeal of the map (0, 1.25, 2.5 points). **Evaluation Strategy:** - 0 points – if the student does not meet the requirement. - 1.25 points – if the student partially meets the requirement. - 2.5 points – if the student fully meets the requirement. Scores will range from 0 to 10 points.

Individual task about AI legal aspects	15 %	Week 10	Students will research and analyse a specific legal aspect of AI technology, such as data privacy, liability, intellectual property, or regulatory compliance. They will then produce a presentation detailing their findings and recommendations. **Requirements for presentation:** - Depth of Understanding demonstrates a comprehensive understanding of the chosen legal aspect of AI (0, 1, 2. points).
			- Research Skills utilizes a variety of credible sources, such as academic journals, legal texts, and reputable websites (0, 1, 2 points). - Analysis: presents a thoughtful analysis of the legal implications and challenges associated with AI technology, identifies key legal issues and discusses their significance (0, 1, 2 points). - Application: applies legal principles to real-world scenarios or case studies (0, 1, 2 points). - Critical Thinking: offers insights and perspectives that go beyond surface-level observations (0, 1, 2 points). **Evaluation Strategy:** - 0 points – if the student does not meet the requirement. - 1 point – if the student partially meets the requirement. - 2 points – if the student fully meets the requirement. Scores will range from 0 to 10 points.
Classification task based on machine learning	15 %	Week 15	Students will train a machine learning model for classification on selected dataset with a focus on sensitivity analysis to determine ethical issues. Requirements for classification task: - selected data and its preparation (0, 1.25, 2.5 points); - model selection, training, testing, sensitivity analysis (0, 1.25, 2.5 points); - defence by explanation of approach and process (0, 1.25, 2.5 points); - Ethical considerations and bias mitigation (0, 1.25, 2.5 points). **Evaluation Strategy:** - 0 points – if the student does not meet the requirement. - 1.25 points – if the student partially meets the requirement. - 2.5 points – if the student fully meets the requirement. Scores will range from 0 to 10 points.

Presenting team project	40 %	During the season	Students will assign to 2 – 4 groups, select specific industry and application area, and prepare product/service approach by considering ethical, legal and practical aspects. Project will require to provide program code for selected dataset. During last week of semester students will make presentation, and during season will provide written report.

Author	Publishing year	Title	Issue of a periodical or volume of a publication; pages	Publishing house or internet site
Required reading				
1. Josh Cowls (ed.)	2021	Artificial Intelligence for Social Good	Philosophy and Technology, 34 vol., issue 1, pp. 1-6; 45-65;	Springer
2. Chui, M., Harryson, M., Valley, S., Manyika, J., & Roberts, R.	2018	Notes from the AI frontier applying AI for social good		McKinsey & Company
3. Cowls, J., King, T. C Taddeo, M., & Florid L.	2019	Designing AI for social good: Seven essential factors	pp. 1-26;	Available at SSRN: https://ssrn.com/abstract=3388669 or http://dx.doi.org/10.2139/ssrn.3388669
4. European Commission.	2020	AI Watch Assessing Technology Readiness Levels for Artificial Intelligence		Available at: https://publications.jrc.ec.europa.eu/repository/handle/JRC122014
5. Mahler, S. J. ().	2012	Team-based learning in social sciences research methods classes.	In Team-Based Learning in the social sciences and humanities, pp. 113-128;	Routledge
6. Bruyn, J. D., Vanleenhove, C.	2021	Artificial intelligence and the law		Available at: https://www.cambridge.org/core/books/artificial-intelligence-and-the-law/D7DF8D83C7A00A2A8D4904FC4B3B2899
7. Postema, G. J.	2022	AI in Law or AI in the Place of Law?		Available at: https://doi.org/10.1093/oso/9780190645342.003.0014
8. Gordon, J. S.	2021	AI and law: ethical, legal, and socio-political implications		Available at: https://link.springer.com/article/10.1007/s00146-021-01194-0
9. Howard, J., & Gugger, S.	2020	Deep Learning for Coders with fastai and PyTorch.		O'Reilly Media
10. McKinney, W.	2022	Python for data analysis		O'Reilly Media
11. Molnar, C	2020	Interpretable machine learning		Lulu
Recommended reading				

Ng, A.	2018	Machine learning yearning: Technical strategy for AI engineers, in the era of deep learning.		
Locascio, N. B. N.	2017	Fundamentals of deep learning: Designing next-generation machine intelligence algorithms		O'Reilly Media

Rubrics of the course

COMPETENCE	THRESHOLD LEVEL OF ACHIEVEMENT	TYPICAL LEVEL OF ACHIEVEMENT	EXCELLENT LEVEL OF ACHIEVEMENT
Recognition of Interdisciplinary Perspectives	· The complexity of the problem is not adequately defined, with insufficient justification for the need for an interdisciplinary approach within the context of ethical AI applications.	· The complex problem is presented adequately, but the relevance of the interdisciplinary approach could be better developed, especially regarding its impact on ethical AI solutions for social good.	· The complex problem is examined comprehensively, with a well-supported review of relevant literature, presenting insights that are aligned with current issues in AI and society.
	· Key insights or directional perspectives are described too superficially, with the omission of essential concepts related to AI ethics, law, or social good.	· The rationale for choosing specific interdisciplinary perspectives is explained, detailing why these perspectives are relevant and important for addressing the problem, though the exploration of other relevant fields might be limited.	· A thorough analysis is conducted, integrating insights from multiple disciplines, clearly defining the relevance of interdisciplinary concepts and why they are critical for understanding the ethical implications of AI.
	· There is uncertainty in choosing one or more perspectives for analyzing complex AI-related problems, with missing or underdeveloped insights.	· Some of the concepts related to AI ethics and law are expanded upon, though some aspects remain underdeveloped. Insights are generally balanced, but not all are clearly or consistently presented.	· The interdisciplinary problem is explored deeply, with clear definitions of key terms and concepts from various fields (e.g., AI, law, ethics, social science), explaining their interconnections and implications for AI's societal impact.
	· The provided insights are inconsistent or lack balance, with a failure to properly define interdisciplinary terms related to AI, ethics, and social impact.	· Scientific terms from different disciplines related to AI ethics, legal challenges, and societal impact are presented, though not all are fully explained. More effort should be made to present a more coherent set of insights.	· The critical perspectives and terminologies related to the interdisciplinary problem are highlighted and well-balanced, demonstrating a nuanced understanding of the ethical, legal, and social dimensions of AI. Relevant perspectives from other disciplines are acknowledged and critically analyzed.
Collaborative and Innovative Thinking	· Lacks openness to ideas and beliefs from other disciplines, resulting in a limited perspective on ethical AI issues.	· Other disciplines' ideas and beliefs are reviewed, but only superficially, without fully engaging with their implications for AI ethics and societal impact.	· Ideas and beliefs from other disciplines are critically evaluated, with a deep engagement in understanding their implications for ethical AI solutions and their societal impact.
	· The assumptions and premises of the chosen perspective are not critically assessed, leading to potential biases in analyzing AI's role in social good.	· The analysis is not sufficiently separated from personal views and preconceptions, affecting the objectivity of the problem-solving process in AI-related contexts.	· Personal biases are recognized, and the limitations of the chosen perspective are understood, with an effort to explore alternative views and integrate them into the problem-solving process.

	· The problem-solving process nearly excludes alternative perspectives, focusing solely on traditional or familiar approaches.	· Perspectives from other disciplines are acknowledged but are not deemed important enough, leading to a narrow understanding of AI's role in solving complex social issues.	· An understanding of the biases inherent in personal viewpoints and the limitations of the primary perspective is demonstrated, ensuring a well-rounded approach to analyzing AI's role in social good.
	· Only standard ideas and conventional solutions are proposed, without considering innovative or interdisciplinary approaches to AI applications.	· While some non-standard ideas are introduced, they are not fully developed or integrated, resulting in a solution that lacks innovative or interdisciplinary depth.	· Innovative and interdisciplinary approaches are not only proposed but are also well-developed, demonstrating a clear understanding of how different disciplines can collaborate to create impactful AI solutions.
Collaboration	· Collaboration occurs sporadically, with minimal initiative or effort to engage with others in an interdisciplinary team within the context of AI for social good.	· The opinions of other interdisciplinary team members are listened to, but the information is not fully integrated or utilized in the context of ethical AI discussions.	· The opinions of other interdisciplinary team members are actively listened to and objectively assessed, with a focus on integrating perspectives from different disciplines to address AI's ethical and societal challenges effectively.
	· The rules of effective team collaboration are known, but trust is placed only in the competencies of a few team members, which limits the potential for interdisciplinary learning.	· The rules of effective teamwork are generally followed, and trust is placed in most team members' competencies, enabling a more cohesive team dynamic.	· The rules of effective teamwork are fully adhered to, with trust in all team members' competencies, fostering a collaborative environment that maximizes interdisciplinary input.
	· Feedback is provided to others, but team members are hesitant to accept or act upon it, hindering the collaborative process.	· Feedback is provided appropriately, and other team members' suggestions are received with some consideration and reflection.	· Constructive feedback is provided and received openly, leading to an environment where interdisciplinary team members' suggestions are valued and acted upon.
	· There is some ability to seek compromise with other interdisciplinary team members, but psychological tension and stress are generally avoided rather than addressed.	· Compromise is sought effectively with interdisciplinary team members, and tasks are shared more evenly, contributing to a balanced workload.	· When tackling complex problems, there is a strong ability to seek and achieve compromise with other interdisciplinary team members, setting clear goals, establishing priorities, and emphasizing core values while supporting others in the team.
	· Task distribution within the team is uneven, with limited sharing of responsibilities and workload.	· There is a more objective assessment of the opinions of other interdisciplinary team members, but not all are regarded as equally important in achieving the group's goals.	· The opinions of other interdisciplinary team members are objectively evaluated, and any signs of bias are quickly identified and addressed to prevent hindrances to the problem-solving process. Effective responses are made to any emerging issues, ensuring that the team works towards the most suitable solutions.

Communication in an Interdisciplinary Team	· Knowledge from different disciplines about the complex problem being addressed is not adequately communicated to the audience, leading to misunderstandings or gaps in comprehension, particularly in the context of ethical AI.	· Knowledge gathered from different disciplines is communicated to the audience, but the level of understanding varies, depending on the audience's preparedness or background in AI and ethics.	· Knowledge from different disciplines is effectively communicated to the audience, with careful consideration of the audience's background and preparedness, ensuring that complex ethical AI issues are understood by all.
	· Different opinions are heard but are often judged too critically, which may hinder open dialogue and collaboration within the interdisciplinary team.	· Different opinions are listened to, and the ideas of others are considered, though there may still be some bias in how these ideas are evaluated.	· Different opinions are not only heard but are objectively evaluated, fostering an open and constructive dialogue within the interdisciplinary team.
	· Insights are difficult to explain clearly to other team members, causing confusion and slowing down the problem-solving process.	· Insights are communicated more clearly to team members, enabling a better understanding and more effective collaboration.	· Insights are communicated clearly and effectively to all team members, enhancing the team's ability to collaborate and solve problems related to AI and its societal impact.
	· Personal assumptions and conclusions are not always clearly formulated, leading to potential miscommunication within the team.	· Personal assumptions and conclusions are usually clearly formulated, allowing for clearer communication and fewer misunderstandings within the interdisciplinary team.	· Personal assumptions and conclusions are clearly and effectively formulated, ensuring that all team members have a shared understanding of the issues and proposed solutions.
Critical Reflection	· Reflection is limited to the learner's existing experiences, with little effort made to incorporate new insights or perspectives related to ethical AI.	· Reflection allows for the identification of fragmented assumptions and challenges encountered during different stages of the problem-solving process in AI-related contexts.	· Reflection provides detailed insights into the challenges and learning outcomes encountered during the various stages of the problem-solving process, with a strong focus on ethical AI implications.
	· During reflection, the broader societal impact of the complex problem-solving process is almost entirely overlooked, leading to a narrow understanding of the issue.	· Reflection on complex problem-solving is more thorough but still lacks depth, with some aspects of the problem not fully explored or understood.	· Reflection thoroughly describes the problem-solving process, considering the complexity of the issue and the potential societal impact of the proposed solution.
	· Potential limitations of the proposed solution to the problem are rarely considered, resulting in an incomplete analysis and understanding of ethical AI challenges.	· While the societal impact is considered, the analysis may miss some critical dimensions, and the reflection does not fully capture the complexity of the issue.	· The reflection is comprehensive, addressing all relevant aspects of the problem, including limitations and potential unintended consequences, ensuring a well-rounded understanding of the ethical and societal dimensions of AI.