An AI Act-Driven Design for Detecting Brain Tumors through Reconfiguration

Antonio Curci[1,2,*], Andrea Esposito[1]

[1]Department of Computer Science, University of Bari Aldo Moro, Via E. Orabona 4, 70125 Bari, Italy
[2]Department of Computer Science, University of Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy

Abstract

Although Artificial Intelligence (AI) is permeating countless domains of application in modern society, it is important to design, develop, and deploy AI-based software that safeguards humans and their well-being. The AI Act, the European Union's legal framework to regulate AI, sets a new standard that must be met when creating such systems, which must protect human rights and emphasize human agency in decision-making processes. This research proposes the architecture of an interaction paradigm, designed starting from AI Act principles, aiming to support medical physicians in detecting brain tumors through a multi-modal model. The goal is to establish a symbiotic relationship between humans and AI in which the limitations of one can be compensated by the strengths of the other, while highlighting the importance of humans' judgment and expertise in making diagnoses.

Keywords

Symbiotic Artificial Intelligence, Multi-Modal Model, Medicine, Decision-Making, Human-AI Collaboration

1. Introduction

As scientific and technological progress advances at a very fast pace, Artificial Intelligence (AI) becomes more and more integrated in everyday activities. AI-based systems can vary depending on the domain in which they are deployed and used, being powered by different models, technologies, and interaction mechanisms [1].

Although AI can strongly support humans in performing repetitive and time-consuming tasks, there are several challenges that such systems can introduce regarding ethics, societal well-being and safety, and human agency [2]. In 2024, the European Union (EU) released a legal framework, *Artificial Intelligence Act* (AI Act), with the goal of regulating the creation, deployment, and use of AI. It undertakes a human-centric and risk-based approach that considers humans in all of their dimensions, regardless of their role of *users* that interact with a system [3]. The constraints and obligations that this legal framework introduces depend on the domain the system is intended for and the risks it could impose on humans and society. The AI Act strongly stresses the role of *Trustworthiness* of AI systems: it is obtained over time when using the system, while being a necessary precondition for regulatory compliance, as it allows increase the system's adoption and acceptance in humans' workflows.

Among the numerous fields in which AI is being introduced—e.g., education, industry, finance—medicine can be one of the most critical. AI is bringing substantial aid to physicians and patients, translating into faster diagnoses, more effective therapies, and significant steps forward in research. At the same time, several challenges must be taken into account: if these tools are misused or provide wrong suggestions to physicians, the consequences might be highly severe or, in some cases, irreversible [4]. This raises the need for creating AI systems that emphasize human agency while fostering effective collaboration with humans, making both parties work towards a common goal. The category of systems that are characterized by such features is called Symbiotic Artificial Intelligence, which encompasses a

HHAI-WS 2025: Workshops at the Fourth International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 09–13, 2025, Pisa, Italy

*Corresponding author

antonio.curci@uniba.it (A. Curci); andrea.esposito@uniba.it (A. Esposito)

0000-0001-6863-872X (A. Curci); 0000-0002-9536-3087 (A. Esposito)

subset of Human-Centered Artificial Intelligence (HCAI) systems that aim at enhancing humans skills, compensating for their limitations, and exploiting the interaction process to learn over time [5]. The factors that influence the establishment of a trustworthy human–AI relationship are multifaceted. For instance, ensuring that humans are *in-the-loop* can strongly affect the development of trust dynamics. In Symbiotic Artificial Intelligence (SAI), for example, keeping humans in control and informed about the processes that lie behind AI's output can be the gateway for enabling both parties to learn over time and exhibit adaptive behavior. There are several techniques that can be employed to implement interaction paradigms that support the integration of human feedback in the model's adaptation—for example, explainability. In this scenario, it can have a two-fold objective: first, it allows the system to provide users with explanations about its decision-making process [5] and, second, it serves as an instrument for humans to indicate where to intervene in the correction of the output [6, 7].

This research work introduces the proposal of a new AI-based system, called *BrainDetect*, that aims to detect brain tumors based on gray-scale 2-D Magnetic Resonance Imaging (MRI) scans and tabular data concerning the image. It is powered by a multi-modal model presented in [8], for which a User Interface (UI) is being created along with an interaction mechanism that exploits Gradient-weighted Class Activation Mapping (GradCAM) [9] explanations output to retrain the model based on human expertise.

The article is structured as follows: section 2 discusses the importance of keeping humans *in-the-loop* and at the center of the decision-making process, exploring an interaction paradigm and the AI Act; section 3 illustrates the proposed architecture with the explanation-based intervention mechanism and presents a prototype of the UI; section 4 reports the conclusions and the future work of the research.

2. Keeping Humans in the Loop

AI is strongly contributing to the diagnosis of diseases and illnesses thanks to its ability to process large amounts of data in short amounts of time, supporting physicians in detection and recognition activities [10]. The case of tumors, represent an exemplary case in which AI can be substantially helpful. This research work focuses on brain tumors, which are abnormal growths of cells within the brain or its surrounding structures, which can be either benign or malignant [11]. These tumors pose a significant health concern due to their complex and heterogeneous nature, rapid progression, and high mortality rates. Early and accurate detection is critical, as it can improve the effectiveness of therapies, reducing the risk of irreversible neurological damage [12].

The models that power the AI-based solutions for tumor detection are progressively improving, providing more support to humans. At the same time, the level of sophistication comes with the cost of complexity, which is proportionally increasing over time. In this regard, a technique that has been gaining more interest in the last few years is the use of more than one modality of data to train an AI model. Multi-modal approaches can increase accuracy, taking into account multiple and heterogeneous aspects and contributing to more reliable outcomes [13].

2.1. Interactive Machine Learning for Reconfiguration

When it comes to creating AI systems that support physicians in performing such delicate tasks (e.g., tumor detection, tumor treatment), designers and developers might face several challenges in letting users be properly aware of the processes that lie behind the systems' output and the motivations that led to the outcomes. Transparency, which is the the intelligibility of the algorithm itself and its inner workings [5], plays a crucial role in this context, as it enables users to obtain insights about the model, its structure, and processes. Explainability, on the other hand, indicates the property of the model to generate human-understandable explanations of its outcomes and decision-making processes [5]. Although black-box models should be avoided [14, 15], their high performance often justifies their use, thus making *post-hoc* explainability techniques useful as a workaround [5, 16]. In the case of convolutional processes for images, one of the most widely used methods is Class Activation Mapping (CAM), specifically, GradCAM, which highlights the spatial regions in the input image that most

influence the model's reasoning by leveraging the gradients of the target class with respect to the feature maps [9]. These methods can have a double-sided function, representing both the explanation of the reasoning process and the instrument to modify or correct the outcome reached by the model. Exploiting such explanations for reconfiguring the model can be particularly useful for implementing Interactive Machine Learning (IML), which allows human expertise to be integrated in the model, adjusting its performance based on their judgement and experience [17]. The integration of IML into workflows can represent a significant step towards the establishment of a symbiotic relationship between humans and AI, improving collaboration [18].

In this regard, the interaction paradigm presented by Desolda et al. is introduced, which highlights that humans must be provided with the necessary instruments to make informed decisions when using an AI system, especially in medicine, while being enabled to iteratively be part of the model's reasoning [5]. The paradigm has three building blocks at its core: *Clarification*, *Reconfiguration*, and *Iterative Exploration*. More specifically, Clarification concerns providing users with usable explanations concerning how the system reached its output, Reconfiguration enables physicians to revise and check the outputs, correcting the system's response when necessary, and Iterative Exploration represents the strategy that allows users to perform decision-making step-by-step and iteratively [19].

2.2. The AI Act and Decision Making

The AI Act is reshaping the way that AI systems are being created and deployed, introducing new obligations that aim at safeguarding the well-being of humans and society [3, 20]. The legal framework introduces a risk-based classification of AI systems: unacceptable risk, high risk, limited risk, and minimal risk. Depending on this classification, these systems must comply with various obligations and standards concerning multiple aspects, ranging from ensuring human oversight and control to requiring high quality documentation from deployers [21]. For instance, Article 10 sets a standard concerning training data which must be fair, representative, and free from bias [3]. With respect to decision-making and activities that can have an impact on other individuals, Article 14 emphasizes that AI systems must be designed to allow human intervention or override, ensuring humans remain in control over critical decisions [3].

This research work relies on the main principles that the AI Act is based on, highlighting the importance of its application in decision-making scenarios. If properly implemented, the legal framework can contribute to the achievement of a symbiotic relationship between humans and AI, which finds an almost natural application in scenarios in which humans are required to make choices. Decision-making is a very delicate and intricate process influenced by various factors that touch on cognition, emotions, expertise, and personal experience [22]. Any external input, such as AI's responses, can alter physicians' traditional way of carrying out tasks like creating diagnoses or therapies. Thus, it must ensure that its users are provided with the proper instruments and conditions to reach outcomes that are not harmful to society or other individuals [23]. For example, in the case of brain tumor detection, a wrong diagnosis, blindly accepted by a physician, can lead to unnecessary treatments, which could seriously damage patients' health. This implies that humans should trust AI only if they are put in the conditions to use their judgment to distinguish the appropriateness of the outputs, even if they are not AI specialists or computer scientists [24, 25].

The application of the interaction paradigm and the research presented in [26] provided the instruments to build the proposals of the interaction paradigm for BrainDetect, as well as the initial wire-frame prototypes of its UI, as described in the next section.

3. Explanation-Based Intervention

The multi-modal model that BrainDetect features is composed of two main channels that are merged into one through a concatenation layer. The two inputs that it supports are 2-D grayscale MRI scan images of human brain and their relative tabular data [8]. Although the current model exhibits high

levels of accuracy (99%), it is important to ensure that end-users are provided with the right instruments to determine the correctness of its outputs and intervene when necessary.

This research proposes an architecture of an IML system [17] based on GradCAM explainability output generated upon the classification of human brain MRI scans to detect tumors. The interaction paradigm in question is illustrated in fig. 1. The goal is to keep humans always *in-the-loop*, enabling them to adjust the AI model's reasoning process based on their expertise and knowledge, ensuring a suitable level of automation of the system for carrying out their task properly. At the same time, transparency is also strongly considered by integrating an explainability, GradCAM, to ensure that physicians can grasp the areas of interest of the model.

After receiving the MRI scan and tabular data as input to the system, the model processes them and provides a binary classification output: *ill* or *healthy*. The physician can either agree (case 1 in fig. 1) with the classification or disagree with it. In the latter case, the physician either has not detected a tumor at all (case 2 in fig. 1), or has detected a tumor elsewhere with respect to the areas highlighted by the system (case 3 in fig. 1). In both cases, human feedback is provided to the model, affecting its future decisions by reinforcing or inhibiting its behavior. This can be implemented, for example, through *Reinforcement Learning (RL) from human feedback* [27].

In case 1, the feedback is sent to the model with no further details. In cases 2 and 3, the user is led to the *Reconfiguration Screen*, illustrated in fig. 2. Here, the MRI appears subdivided into n patches of equal size, each clickable and available for selection. By selecting one patch (or multiple adjacent ones), physicians indicate to the system an area that may contain a tumor.

If the physician disagrees with the AI system, they are asked to express their confidence in their decision—for example, through a simple semantic differential scale (see fig. 2. This allows the model to weigh human feedback during its adaptation. Although further investigation is needed, this design decision was made since corrections can be noisy or, at times, wrong. It represents a way of "letting AI know" that the user is in disagreement with its output but still uncertain. To reach high-quality outcomes, it is important to avoid fitting the model to potentially incorrect corrections, which could hinder the human–AI trust dynamic [28, 1].

The final objective is to enable continuous learning on the system's behalf, with humans guiding the process by correcting mistakes or highlighting important features.

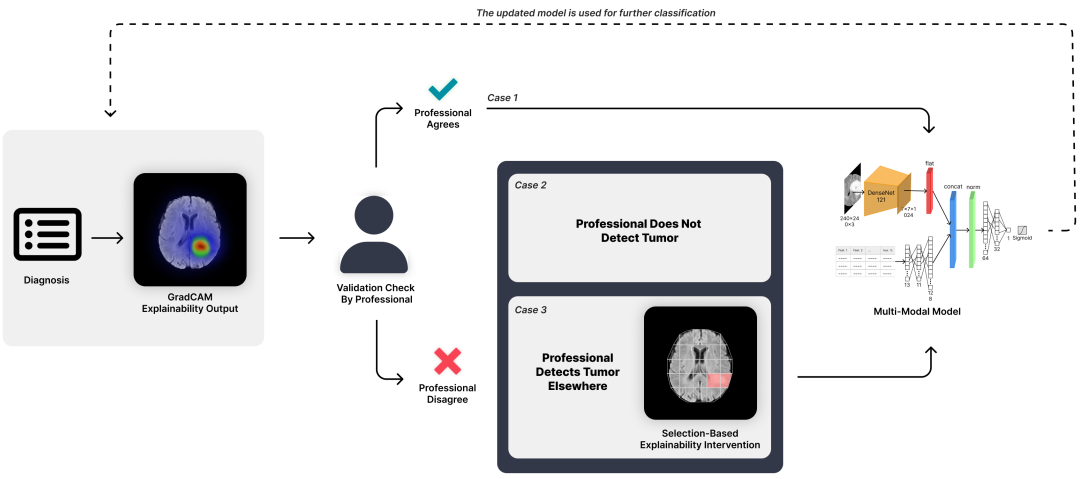

Figure 1: Proposed architecture of the interaction paradigm for the AI model reconfiguration based on explanations. When validating the AI output, three cases can occur: the physician confirms the AI decision (case 1), the physician overturns the AI decision since no tumor is present (case 2), or the physician confirms the AI decision but recognizes a tumor in a different area of the MRI scan (case 3).

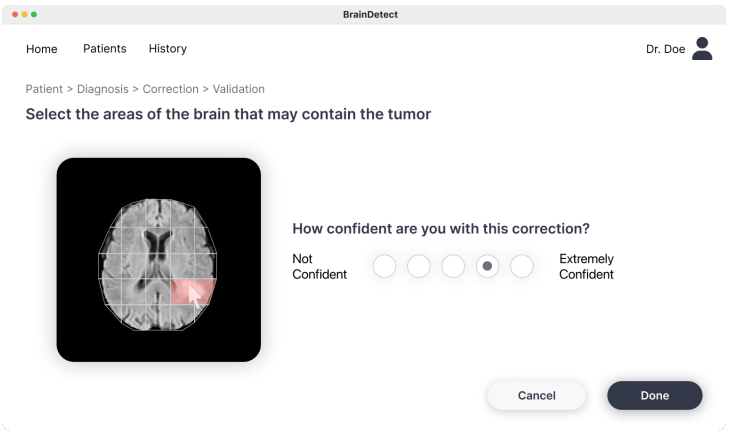

Figure 2: Example of human intervention on BrainDetect for the validation screen. physicians are able to indicate their confidence in their own decision, for example, through a simple semantic differential scale.

4. Conclusions

The strong impact that AI has on modern society is being regulated by legal bodies working towards a more ethical and safe creation and deployment of such systems, especially in application domains requiring decisions that can impact other individuals. Medicine is the domain analyzed in this research, which proposes an architecture for a multi-modal model that detects brain tumors. The interaction paradigm focuses on complying with the AI Act, keeping humans in-the-loop by ensuring that they can revise and check the predictions made by the system, and correcting potential mistakes made by the model. The ultimate goal, as mentioned in the sections above, is to reach symbiosis between humans and AI, where both can learn from each other, improving over time, and compensating for the limitations with the other's strength [5]. Making BrainDetect fall under the category of SAI is an objective that is being undertaken from the beginning of the project, which is serving as a case study for the investigation of the necessary instruments to pursue *Symbiosis-by-Design*.

Currently, the work presented here is mostly a proposal: although the actual AI model for classification exists (see [8]), ongoing research efforts aim at introducing human feedback in the AI model training and in implementing the interaction loop presented in fig. 1. Therefore, future work of this research regards implementing and refining BrainDetect by adhering to Human-Centred Design principles [29].

Through user studies, the interaction loop presented in fig. 1 could be further refined by exploring additional implicit factors (e.g., decision-making time) that could provide indications on the evolution of human–AI trust. Such factors could be instrumental to the model adaptation.

It is also intended to investigate the integration of the selection of non-adjacent areas of the brain in the reconfiguration step, specifically for critical patients with a brain that has multiple ill regions. An additional user study is required to assess the effectiveness of this proposal in accomplishing human–AI symbiosis by analyzing how the interaction mechanism proposed in fig. 1 impacts users' performance and their trust in AI.

Acknowledgments

The research of Antonio Curci is supported by the co-funding of the European Union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centers, companies, and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000013 – "Future Artificial Intelligence Research – FAIR" - CUP: H97G22000210007). The research of Andrea Esposito is funded by a Ph.D. fellowship within the framework of the Italian "D.M. n. 352, April 9, 2022"- under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 3.3 – Ph.D. Project "Human-Centred Artificial Intelligence (HCAI) techniques for supporting end users interacting with AI systems," co-supported by "Eusoft S.r.l." (CUP H91I22000410007).

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] B. Shneiderman, Human-Centered AI, 1 ed., Oxford University PressOxford, 2022. URL: https://academic.oup.com/book/41126. doi:10.1093/oso/9780192845290.001.0001.

[2] F. Paternò, M. Burnett, G. Fischer, M. Matera, B. Myers, A. Schmidt, Artificial Intelligence versus End-User Development: A Panel on What Are the Tradeoffs in Daily Automations?, in: C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, K. Inkpen (Eds.), Human-Computer Interaction – INTERACT 2021, volume 12936, Springer International Publishing, Cham, 2021, pp. 340–343. URL: https://link.springer.com/10.1007/978-3-030-85607-6_33. doi:10.1007/978-3-030-85607-6_33, series Title: Lecture Notes in Computer Science.

[3] European Parliament, Council of the European Union, Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024.

[4] W. Xiong, H. Fan, L. Ma, C. Wang, Challenges of human—machine collaboration in risky decision-making, Frontiers of Engineering Management 9 (2022) 89–103. URL: https://link.springer.com/10.1007/s42524-021-0182-0. doi:10.1007/s42524-021-0182-0.

[5] G. Desolda, A. Esposito, R. Lanzilotti, A. Piccinno, M. F. Costabile, From human-centered to symbiotic artificial intelligence: a focus on medical applications, Multimedia Tools and Applications (2024). URL: https://link.springer.com/10.1007/s11042-024-20414-5. doi:10.1007/s11042-024-20414-5.

[6] G. Desolda, G. Dimauro, A. Esposito, R. Lanzilotti, M. Matera, M. Zancanaro, A Human–AI interaction paradigm and its application to rhinocytology, Artificial Intelligence in Medicine 155 (2024) 102933. URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365724001751. doi:10.1016/j.artmed.2024.102933.

[7] A. Esposito, M. Calvano, A. Curci, F. Greco, R. Lanzilotti, A. Piccinno, Explanation-Driven Interventions for Artificial Intelligence Model Customization: Empowering End-Users to Tailor Black-Box AI in Rhinocytology, in: C. Santoro, A. Schmidt, M. Matera, A. Bellucci (Eds.), End-User Development, volume 15713, Springer Nature Switzerland, Cham, 2025, pp. 161–170. URL: https://link.springer.com/10.1007/978-3-031-95452-8_10. doi:10.1007/978-3-031-95452-8_10.

[8] A. Curci, A. Esposito, Detecting Brain Tumors Through Multimodal Neural Networks, in: 13th International Conference on Pattern Recognition Applications and Methods, SCITEPRESS – Science and Technology Publications, Lda., Rome, Italy, 2024, pp. 995–1000. URL: https://arxiv.org/abs/2402.00038. doi:10.5220/0012608600003654.

[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, International Journal of Computer Vision 128 (2020) 336–359. URL: http://link.springer.com/10.1007/s11263-019-01228-7. doi:10.1007/s11263-019-01228-7.

[10] D. Göndöcs, V. Dörfler, AI in medical diagnosis: AI prediction & human judgment, Artificial Intelligence in Medicine 149 (2024) 102769. URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365724000113. doi:10.1016/j.artmed.2024.102769.

[11] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, D. W. Ellison, The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary, Acta Neuropathologica 131 (2016) 803–820. URL: http://link.springer.com/10.1007/s00401-016-1545-1. doi:10.1007/s00401-016-1545-1.

[12] R. Stupp, M. Weller, K. Belanger, U. Bogdahn, S. K. Ludwin, D. Lacombe, R. O. Mirimanoff, Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma, n engl j med (2005).

[13] C. Shang, H. Zhang, H. Wen, Y. Yang, Understanding Multimodal Deep Neural Networks: A Concept Selection View, 2024. URL: http://arxiv.org/abs/2404.08964. doi:10.48550/arXiv.2404.08964, arXiv:2404.08964 [cs].

[14] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence 1 (2019) 206–215. URL: https://www.nature.com/articles/s42256-019-0048-x. doi:10.1038/s42256-019-0048-x.

[15] R. O. Weber, A. J. Johs, P. Goel, J. M. Silva, XAI is in trouble, AI Magazine 45 (2024) 300–316. URL: https://onlinelibrary.wiley.com/doi/10.1002/aaai.12184. doi:10.1002/aaai.12184.

[16] C. O. Retzlaff, A. Angerschmid, A. Saranti, D. Schneeberger, R. Röttger, H. Müller, A. Holzinger, Post-hoc vs ante-hoc explanations: xAI design guidelines for data scientists, Cognitive Systems Research 86 (2024) 101243. URL: https://linkinghub.elsevier.com/retrieve/pii/S1389041724000378. doi:10.1016/j.cogsys.2024.101243.

[17] N. A. Wondimu, C. Buche, U. Visser, Interactive Machine Learning: A State of the Art Review, 2022. URL: http://arxiv.org/abs/2207.06196. doi:10.48550/arXiv.2207.06196, arXiv:2207.06196 [cs].

[18] J. Lee, Is Artificial Intelligence Better Than Human Clinicians in Predicting Patient Outcomes?, Journal of Medical Internet Research 22 (2020) e19918. URL: http://www.jmir.org/2020/8/e19918/. doi:10.2196/19918.

[19] G. Desolda, G. Dimauro, A. Esposito, R. Lanzilotti, M. Matera, M. Zancanaro, A Human–AI interaction paradigm and its application to rhinocytology, Artificial Intelligence in Medicine 155 (2024) 102933. URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365724001751. doi:10.1016/j.artmed.2024.102933.

[20] R. J. Neuwirth, Prohibited artificial intelligence practices in the proposed EU artificial intelligence act (AIA), Computer Law & Security Review 48 (2023) 105798. URL: https://linkinghub.elsevier.com/retrieve/pii/S0267364923000092. doi:10.1016/j.clsr.2023.105798.

[21] B. Gyevnar, N. Ferguson, B. Schafer, Bridging the Transparency Gap: What Can Explainable AI Learn from the AI Act?, in: K. Gal, A. Nowé, G. J. Nalepa, R. Fairstein, R. Rădulescu (Eds.), Frontiers in Artificial Intelligence and Applications, IOS Press, 2023. URL: https://ebooks.iospress.nl/doi/10.3233/FAIA230367. doi:10.3233/FAIA230367.

[22] C. Giachino, M. Cepel, E. Truant, A. Bargoni, Artificial intelligence-driven decision making and firm performance: a quantitative approach, Management Decision (2024). URL: https://www.emerald.com/insight/content/doi/10.1108/MD-10-2023-1966/full/html. doi:10.1108/MD-10-2023-1966.

[23] D. Niraula, K. C. Cuneo, I. D. Dinov, B. D. Gonzalez, J. B. Jamaluddin, J. J. Jin, Y. Luo, M. M. Matuszak, R. K. Ten Haken, A. K. Bryant, T. J. Dilling, M. P. Dykstra, J. M. Frakes, C. L. Liveringhouse, S. R. Miller, M. N. Mills, R. F. Palm, S. N. Regan, A. Rishi, J. F. Torres-Roca, H.-H. M. Yu, I. El Naqa, Intricacies of human–AI interaction in dynamic decision-making for precision oncology, Nature Communications 16 (2025) 1138. URL: https://www.nature.com/articles/s41467-024-55259-x. doi:10.1038/s41467-024-55259-x.

[24] L. D. Urquhart, G. McGarry, A. Crabtree, Legal Provocations for HCI in the Design and Development of Trustworthy Autonomous Systems, in: Nordic Human-Computer Interaction Conference, ACM, Aarhus Denmark, 2022, pp. 1–12. URL: https://dl.acm.org/doi/10.1145/3546155.3546690. doi:10.1145/3546155.3546690.

[25] D. G. Widder, L. Dabbish, J. D. Herbsleb, A. Holloway, S. Davidoff, Trust in Collaborative Automation in High Stakes Software Engineering Work: A Case Study at NASA, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 2021, pp. 1–13. URL: https://dl.acm.org/doi/10.1145/3411764.3445650. doi:10.1145/3411764.3445650.

[26] M. Calvano, A. Curci, G. Desolda, A. Esposito, R. Lanzilotti, A. Piccinno, Building Symbiotic AI: Reviewing the AI Act for a Human-Centred, Principle-Based Framework, 2025. URL: http://arxiv.org/abs/2501.08046. doi:10.48550/arXiv.2501.08046, arXiv:2501.08046 [cs].

[27] T. Kaufmann, P. Weng, V. Bengs, E. Hüllermeier, A Survey of Reinforcement Learning from Human Feedback, 2023. URL: https://arxiv.org/abs/2312.14925. doi:10.48550/ARXIV.2312.14925.

[28] P. Kieseberg, E. Weippl, A. M. Tjoa, F. Cabitza, A. Campagner, A. Holzinger, Controllable AI - An Alternative to Trustworthiness in Complex AI Systems?, in: A. Holzinger, P. Kieseberg, F. Cabitza, A. Campagner, A. M. Tjoa, E. Weippl (Eds.), Machine Learning and Knowledge Extraction, volume 14065, Springer Nature Switzerland, Cham, 2023, pp. 1–12. URL: https://link.springer.com/10.1007/978-3-031-40837-3_1. doi:10.1007/978-3-031-40837-3_1, series Title: Lecture Notes in Computer Science.

[29] ISO, 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems, 2019. URL: https://www.iso.org/standard/77520.html.