Authors	Title	Publication	Year	Publisher	I1	I2	E1	E2	E3	E4	Included?
Lisegow, Lucas; Barnes, Ethan; Pennington, Ava; Thackeray, John;	Enhancing explainability in large language models through belief change: A simulation-based approach		2024					x			No
Fenoglio, Enzo; Kazim, Emre;	AI explainability, interpretability, and transparency	The Elgar Companion to Applied AI Ethics	2024	Edward Elgar Publishing							Yes
Ivanov, Yuri;	Understanding the inner workings of large language models: Interpretability and explainability	MZ Journal of Artificial Intelligence	2024								Yes
Yang, Zhou; Sun, Zhensu; Yue, Terry Zhuo; Devanbu, Premkumar; Lo, David;	Robustness, security, privacy, explainability, efficiency, and usability of large language models for code	arXiv preprint arXiv:2403.07506	2024					x			No
Fujiwara, Kaito; Sasaki, Miyu; Nakamura, Akira; Watanabe, Natsumi;	Measuring the interpretability and explainability of model decisions of five large language models	Open Science Framework: Charlottesville, VA, USA	2024					x			No
Quttainah, Majdi; Mishra, Vinaytosh; Madakam, Somayya; Lurie, Yotam; Mark, Shlomo;	Cost, usability, credibility, fairness, accountability, transparency, and explainability framework for safe and effective large language models in medical education: Narrative review and qualitative study	Jmir Ai	2024	JMIR Publications Inc., Toronto, Canada							Yes
Yin, Fen; Zhong, Mu; Ru, Zhihao;	Exploring Explainability in Large Language Models		2025	Preprints				x			No
Zhao, Haiyan; Chen, Hanjie; Yang, Fan; Liu, Ninghao; Deng, Huiqi; Cai, Hengyi; Wang, Shuaiqiang; Yin, Dawei; Du, Mengnan;	Explainability for large language models: A survey	ACM Transactions on Intelligent Systems and Technology	2024	ACM New York, NY							Yes
Mumuni, Fuseini; Mumuni, Alhassan;	Explainable artificial intelligence (XAI): from inherent explainability to large language models	arXiv preprint arXiv:2501.09967	2025					x			No
Barnes, Emily; Hutson, James;	Navigating the complexities of ai: The critical role of interpretability and explainability in ensuring transparency and trust	International Journal of Multidisciplinary and Current Educational Research	2024								Yes
Balasubramaniam, Nagadivya; Kauppinen, Marjo; Hiekkanen, Kari; Kujala, Sari;	Transparency and explainability of AI systems: ethical guidelines in practice	International working conference on requirements engineering: foundation for software quality	2022	Springer							Yes
Fernández-Becerra, Laura; González-Santamarta, Miguel Ángel; Guerrero-Higueras, Ángel Manuel; Rodríguez-Lera, Francisco Javier; Olivera, Vicente Matellán;	Enhancing trust in autonomous agents: An architecture for accountability and explainability through blockchain and large language models	arXiv preprint arXiv:2403.09567	2024					x			No
Gaur, Manas; Sheth, Amit;	Building trustworthy NeuroSymbolic AI Systems: Consistency, reliability, explainability, and safety	AI Magazine	2024	Wiley Online Library							Yes
Eke, Christopher Ifeanyi; Shuib, Liyana;	The role of explainability and transparency in fostering trust in AI healthcare systems: a systematic literature review, open issues and potential solutions	Neural Computing and Applications	2024	Springer							Yes
Luo, Haoyan; Specia, Lucia;	From understanding to utilization: A survey on explainability for large language models	arXiv preprint arXiv:2401.12874	2024					x			No
Kim, Sunnie SY;	Establishing Appropriate Trust in AI through Transparency and Explainability	Extended Abstracts of the CHI Conference on Human Factors in Computing Systems	2024								Yes
Balasubramaniam, Nagadivya; Kauppinen, Marjo; Rannisto, Antti; Hiekkanen, Kari; Kujala, Sari;	Transparency and explainability of AI systems: From ethical guidelines to requirements	Information and Software Technology	2023	Elsevier							Yes
Petkovic, Dragutin;	It is not "Accuracy vs. Explainability"—we need both for trustworthy AI systems	IEEE Transactions on Technology and Society	2023	IEEE							Yes
Bilokon, Paul;	Explainability and Semantics-Bridging Natural Language Flexibility and Formal Precision: Toward a Semantic Framework for Large Language Models		2024					x			No
Morandini, Sofia; Fraboni, Federico; Balatti, Enzo; Hackmann, Aranka; Brendel, Hannah; Puzzo, Gabriele; Volpi, Lucia; Giusino, Davide; De Angelis, Marco; Pietrantoni, Luca;	Assessing the transparency and explainability of AI algorithms in planning and scheduling tools: a review of the literature	Human Interaction and Emerging Technologies (IHIET 2023): Artificial Intelligence and Future Applications. AHFE	2023								Yes
Goktas, Polat;	Ethics, transparency, and explainability in generative ai decision-making systems: A comprehensive bibliometric study	Journal of Decision Systems	2024	Taylor & Francis							Yes
San Miguel, Gabriela Sanchez; Griffith, Henry; Silva, Jacob; Rathore, Heena;	Evaluating the Explainability of Large Language Models for Ethical Decision Making	2025 IEEE International Conference on Consumer Electronics (ICCE)	2025	IEEE							Yes
Chundru, Swathi;	Ensuring Data Integrity Through Robustness and Explainability in AI Models	Transactions on Latest Trends in Artificial Intelligence	2020								Yes
Beamish, David; Exarchakis, Georgios;	Large Language Models for Explainability in Machine Learning							x			No
Silva Mercado, Jonathan;	AI Assisted Domain Modeling Explainability and Traceability	Proceedings of the ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems	2024								Yes
Penicig, James; Chen, Benjamin; Wilson, Scott; Garcia, Sophia;	Assessing explainability in large language models through soft counterfactual analysis: A comparative study of google gemini and openai chatgpt		2024					x			No
Ali, Ibad;	AI Transparency and Explainability							x			No
McNamara Jr, Keith;	Simplifying AI Explanations for the General User: Investigating the Efficacy of Plain Language for Explainability and Interpretability		2023	University of Florida						x	No
Ehsan, Upol; Watkins, Elizabeth A; Wintersberger, Philipp; Manger, Carina; Kim, Sunnie SY; Van Berkel, Niels; Riener, Andreas; Riedl, Mark O;	Human-centered explainable AI (HCXAI): Reloading explainability in the era of large language models (LLMs)	Extended Abstracts of the CHI Conference on Human Factors in Computing Systems	2024								Yes
Spinner, Thilo;	Visual, Interactive Deep Model Debugging: Supporting AI Development and Explainability		2024							x	No
Luz, Ayuns;	Enhancing the interpretability and explainability of ai-driven risk models using llm capabilities		2024	Technical Report				x			No
Spitzer, Philipp; Celis, Sebastian; Martin, Dominik; Kühl, Niklas; Satzger, Gerhard;	Looking Through the Deep Glasses: How Large Language Models Enhance Explainability of Deep Learning Models	Proceedings of Mensch und Computer 2024	2024								Yes
Mohammadi, Seyedali; Raff, Edward; Malekar, Jinendra; Palit, Vedant; Ferraro, Francis; Gaur, Manas;	WellDunn: On the Robustness and Explainability of Language Models and Large Language Models in Identifying Wellness Dimensions	arXiv preprint arXiv:2406.12058	2024					x			No
Roundtree, Aimee Kendall;	AI explainability, interpretability, fairness, and privacy: an integrative review of reviews	International Conference on Human-Computer Interaction	2023	Springer							Yes
Axelsson, Victoria; Wettersten, Linn;	Designing AI for decision making: The importance of trust, control and explainability within a government agency		2024							x	No
Ehsan, Upol; Liao, Q Vera; Muller, Michael; Riedl, Mark O; Weisz, Justin D;	Expanding explainability: Towards social transparency in ai systems	Proceedings of the 2021 CHI conference on human factors in computing systems	2021								Yes
Oz, Burak; Mussaferada, Aliatta; Coursaris, Constantinos, Senecal, Sylvain; Léger, Pierre-majorique;	The Impact of Generative AI Transparency and Explainability on Decision Performance		2025								Yes
Figênio, Mateus R; Santanché, André; Gomes-Jr, Luiz;	The Impact of Activation Patterns in the Explainability of Large Language Models–A Survey of recent advances	Escola Regional de Banco de Dados (ERBD)	2024	SBC							Yes
Wu, Shi; Zhou, Jianlong; Dong, Yifei; Chen, Fang;	Enhancing Explainability of Deep Learning-Based ECG Diagnosis Using Large Language Models	Proceedings of the 2024 8th International Conference on Advances in Artificial Intelligence	2024								Yes
Makridis, Georgios; Fatouros, Georgios; Kiourtis, Athanasios; Kotios, Dimitrios; Koukos, Vasileios; Kyriazis, Dimosthenis; Soldatos, Jonh;	Towards a unified multidimensional explainability metric: evaluating trustworthiness in AI models	2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)	2023	IEEE							Yes
Hu, Yuanyuan; Giacaman, Nasser; Donald, Claire;	Enhancing Trust in Generative AI: Investigating Explainability of LLMs to Analyse Confusion in MOOC Discussions.	LAK Workshops	2024					x			No
Ferrario, Andrea; Loi, Michele;	How explainability contributes to trust in AI	Proceedings of the 2022 ACM conference on fairness, accountability, and transparency	2022								Yes
Freyer, Nils; Groß, Dominik; Lipprandt, Myriam;	The ethical requirement of explainability for AI-DSS in healthcare: a systematic review of reasons	BMC Medical Ethics	2024	Springer							Yes
Chen, Guanxu; Liu, Dongrui; Luo, Tao; Shao, Jing;	SEER: Self-Explainability Enhancement of Large Language Models' Representations	arXiv preprint arXiv:2502.05242	2025					x			No
Amara, Jihen; König-Ries, Birgitta; Samuel, Sheeba;	Enhancing explainability in multimodal large language models using ontological context	arXiv preprint arXiv:2409.18753	2024					x			No
Elia, Miriam; Lopez, Alba Maria; Corrador, Katherin Alexandra; Bauer, Bernhard; Garcia-Cuesta, Esteban;	MQG4AI Towards Responsible High-risk AI–Illustrated for Transparency Focusing on Explainability Techniques	arXiv preprint arXiv:2502.11889	2025					x			No
Tan, Ting Fang; Elangovan, Kabilan; Ong, Jasmine; Shah, Nigam; Sung, Joseph; Wong, Tien Yin; Xue, Lan; Liu, Nan; Wang, Haibo; Kuo, Chang Fu;	A proposed score evaluation framework for large language models: Safety, consensus, objectivity, reproducibility and explainability	arXiv preprint arXiv:2407.07666	2024					x			No
Marqas, Ridwan Boya; Almufti, Saman M; Yusif, Rezhna Azad;	Unveiling explainability in artificial intelligence: a step to-wards transparent AI										Yes
Mojsilovic, Aleksandra;	Introducing AI explainability 360	Au- gust 8, 2019)[2021-05-15]. https://www.ibm.com/blogs/resear	2019								Yes
Pace, Teresa; Ranes, Bryan;	Bias, explainability, transparency, and trust for AI-enabled military systems	Assurance and Security for AI-enabled Systems	2024	SPIE							Yes
Baker, Stephanie; Xiang, Wei;	Explainable ai is responsible ai: How explainability creates trustworthy and socially responsible artificial intelligence	arXiv preprint arXiv:2312.01555	2023					x			No
Shin, Donghee;	The effects of explainability and causality on perception, trust, and acceptance: Implications for explainable AI	International journal of human-computer studies	2021	Elsevier							Yes
Costa, Renato;	Form Trust to Transparency: Understanding the Influence of Explainability on AI Systems	Journal of Innovative Technologies	2023								Yes
Wala, Jens;	AI in Critical Infrastructures–Explainability and Models	35. Forum Bauinformatik, fbi 2024	2024	Technische Universität Hamburg, Institut für Digitales und Autonomes Bauen							Yes
Pandian, A; Mohankumar, K Pradeep; Padmini, S; Amaran, Sibi; Sreekumar, K;	Enhancing Trust and Collaboration Using Explainability in Natural Language Processing for AI-Driven Healthcare	Explainable Artificial Intelligence in the Healthcare Industry	2025	Wiley Online Library							Yes
Famiglini, Lorenzo;	Enhancing the Explainability and Reliability of AI support for Informed Healthcare Decisions		2025	Italy						x	No
Wysocki, Oskar; Davies, Jessica Katharine; Vigo, Markel; Armstrong, Anne Caroline; Landers, Dónal; Lee, Rebecca; Freitas, André;	Assessing the communication gap between AI models and healthcare professionals: Explainability, utility and trust in AI-driven clinical decision-making	Artificial Intelligence	2023	Elsevier							Yes
Blake, Harrison;	Transparency and Explainability in AI: Ensuring Trust and Understanding in Autonomous Decision-Making Systems							x			No
Mendes, Bruno Miguel Lopes;	Integration of AI and Law Viewed Through Explainability: A Framework to Facilitate AI Integration in the Judiciary		2023	Universidade NOVA de Lisboa (Portugal)				x			No

LEGEND

X	not included because of fail constraint
I1	Papers that are peer reviewed.
I2	Papers that target the explainability phenomenon in the context of LLMs or AI.
E1	Papers that are not written in the English language.
E2	Papers that are preprints, white papers, blog posts, or workshop papers without rigorous peer review.
E3	Papers that purely summarize previous work without providing new contributions or methodologies.
E4	Papers that represent patents, bachelor's, master's, or doctoral theses.

Total Count:	237

Authors	Title	Source	Year	Publisher	Included
Vermesan, Ovidiu; Piuri, Vincenzo; Scotti, Fabio; Genovese, Angelo; Labati, Ruggero Donida; Coscia, Pasquale;	Explainability and interpretability concepts for edge ai systems	Advancing Edge Artificial Intelligence	2024	River Publishers	Yes
Banerjee, Gautam; Dhar, Subhankar; Roy, Satyaki; Syed, Riddhiman; Das, Anurita;	Explainability and Transparency in Designing Responsible AI Applications in the Enterprise	The International Conference on Computing, Communication, Cybersecurity & AI	2024	Springer	Yes
Pawlicka, Aleksandra; Pawlicki, Marek; Jaroszewska-Choraś, Dagmara; Kozik, Rafał; Choraś, Michał;	Enhancing Clinical Trust: The Role of AI Explainability in Transforming Healthcare	2024 IEEE International Conference on Data Mining Workshops (ICDMW)	2024	IEEE	Yes
Balan, Dragomir;	Advancing the Trustworthiness of AI: An Integrated Approach to Explainability				No
Alshar'e, Marwan; Abualkishik, Abdallah; Abuhmaidan, Khaled; Kayed, Ahmad;	Enhancing Project Security: Unveiling Trust, Interpretability and Explainability in the Age of AI	5G Enabled Technology for Smart City and Urbanization System	2025	Chapman and Hall/CRC	Yes
Thurzo, Andrej;	Provable AI Ethics and Explainability in Medical and Educational AI Agents: Trustworthy Ethical Firewall	Electronics	2025	MDPI	Yes
Würfel, Jakob; Papenfuß, Anne; Wies, Matthias;	Operationalizing AI Explainability Using Interpretability Cues in the Cockpit: Insights from User-Centered Development of the Intelligent Pilot Advisory System (IPAS)	International Conference on Human-Computer Interaction	2024	Springer	Yes
Stieler, Fabian;	The role of active learning in developing trustworthy AI-Approaches for enhancing transparency and explainability in processes and systems		2025		No
Schoenherr, Jordan Richard; Abbas, Roba; Michael, Katina; Rivas, Pablo; Anderson, Theresa Dirndorfer;	Designing AI using a human-centered approach: Explainability and accuracy toward trustworthiness	IEEE Transactions on Technology and Society	2023	IEEE	Yes
Amin, Rifat Mehreen;	Designing AI-driven Interfaces for Improved Interaction, Decision-Making, and Explainability	Companion Proceedings of the 30th International Conference on Intelligent User Interfaces	2025		Yes
Werz, Johanna M; Borowski, Esther; Isenhardt, Ingrid;	Explainability as a Means for Transparency? Lay Users' Requirements Towards Transparent AI	Cognitive Computing and Internet of Things	2024	AHFE Open Acces	Yes
Kakolu, Sridevi; Faheem, Muhammad Ashraf;	Building trust with generative AI chatbots: Exploring explainability, privacy, and user acceptance	Iconic Research And Engineering Journals	2024		Yes
Hois, Joana; Theofanou-Fuelbier, Dimitra; Junk, Alischa Janine;	How to achieve explainability and transparency in human AI interaction	HCI International 2019-Posters: 21st International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21	2019	Springer	Yes
Kemmerzell, Nils; Schreiner, Annika;	Quantifying the Trade-Offs Between Dimensions of Trustworthy AI-An Empirical Study on Fairness, Explainability, Privacy, and Robustness	German Conference on Artificial Intelligence (Künstliche Intelligenz)	2024	Springer	Yes
Andrulis, Jonas; Meyer, Ole; Schott, Grégory; Weinbach, Samuel; Gruhn, Volker;	Domain-Level Explainability--A Challenge for Creating Trust in Superhuman AI Strategies	arXiv preprint arXiv:2011.06665	2020		No
Shin, Donghee;	User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability	Journal of Broadcasting & Electronic Media	2020	Taylor & Francis	Yes
Bárcena, José Luis Corcuera; Ducange, Pietro; Marcelloni, Francesco; Renda, Alessandro;	Increasing trust in AI through privacy preservation and model explainability: Federated Learning of Fuzzy Regression Trees	Information Fusion	2025	Elsevier	Yes
Kaur, Harmanpreet; Adar, Eytan; Gilbert, Eric; Lampe, Cliff;	Sensible AI: Re-imagining interpretability and explainability using sensemaking theory	Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency	2022		Yes
Sheth, Amit; Gaur, Manas; Roy, Kaushik; Venkataraman, Revathy; Khandelwal, Vedant;	Process knowledge-infused ai: Toward user-level explainability, interpretability, and safety	IEEE Internet Computing	2022	IEEE	Yes
Yu, Shuren;	Towards trustworthy and understandable ai: unraveling explainability strategies on simplifying algorithms, appropriate information disclosure, and high-level collaboration	Proceedings of the 26th International Academic Mindtrek Conference	2023		Yes
Marey, Ahmed; Arjmand, Parisa; Alerab, Ameerh Dana Sabe; Eslami, Mohammad Javad; Saad, Abdelrahman M; Sanchez, Nicole; Umair, Muhammad;	Explainability, transparency and black box challenges of AI in radiology: Impact on patient care in cardiovascular radiology	Egyptian Journal of Radiology and Nuclear Medicine	2024	Springer	Yes
An, Songyang; Teo, Kelvin; McConnell, Michael V; Marshall, John; Galloway, Christopher; Squirrell, David;	AI explainability in oculomics: how it works, its role in establishing trust, and what still needs to be addressed	Progress in Retinal and Eye Research	2025	Elsevier	Yes
Govea, Jaime; Gutierrez, Rommel; Villegas-Ch, William;	Transparency and precision in the age of AI: evaluation of explainability-enhanced recommendation systems	Frontiers in Artificial Intelligence	2024	Frontiers Media SA	Yes
Atf, Zahra; Lewis, Peter R;	Human centricity in the relationship between explainability and trust in ai	IEEE Technology and Society Magazine	2024	IEEE	Yes
Aranovich, Tatiana de Campos; Matulionyte, Rita;	Ensuring AI explainability in healthcare: problems and possible policy solutions	Information & Communications Technology Law	2023	Taylor & Francis	Yes
Radcliffe, Thomas;	Explainability in Generative AI: Strategies to Open the Black Box and Increase Trust		2024		No
Falconmore, Rhiannon;	On the role of explainability and uncertainty in ensuring safety of AI applications		2022	University of Oxford	No
Hamon, Ronan; Junklewitz, Henrik; Sanchez, Ignacio; Malgieri, Gianclaudio; De Hert, Paul;	Bridging the gap between AI and explainability in the GDPR: towards trustworthiness-by-design in automated decision-making	IEEE Computational Intelligence Magazine	2022	IEEE	Yes
Speith, Timo; Xu, Jing;	Explainability and Transparency in Practice: A Comparison Between Corporate and National AI Ethics Guidelines in Germany and China	International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems	2024	Springer	Yes
Vemula, Srikanth; DuFresne-To, Kainen;	Building Intelligent and Intelligible AI: A Framework for Human-like Autonomy and Explainability in Critical Infrastructure	2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)	2024	IEEE	Yes
Frank, Edwin;	Explainability and interpretability in robust and secure AI algorithms		2024		No
Kim, Sunnie SY; Watkins, Elizabeth Anne; Russakovsky, Olga; Fong, Ruth; Monroy-Hernández, Andrés;	" help me help the ai": Understanding how explainability can support human-ai interaction	Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems	2023		Yes
Baron, Sam;	Trust, explainability and AI	Philosophy & Technology	2025	Springer	Yes
Nannini, Luca; Balayn, Agathe; Smith, Adam Leon;	Explainability in AI policies: A critical review of communications, reports, regulations, and standards in the EU, US, and UK	Proceedings of the 2023 ACM conference on fairness, accountability, and transparency	2023		Yes
Kalakoti, Rajesh; Bahsi, Hayretdin; Nõmm, Sven;	Improving IoT security with explainable AI: Quantitative evaluation of explainability for IoT botnet detection	IEEE Internet of Things Journal	2024	IEEE	Yes
Reimer, Ulrich; Maier, Edith; Tödtli, Beat;	Going Beyond Explainability in Medical AI Systems.	Modellierung (Companion)	2020		No
Agostinho, Carlos; Dikopoulou, Zoumpolia; Lavasa, Eleni; Perakis, Konstantinos; Pitsios, Stamatis; Branco, Rui; Reji, Sangeetha; Hetterich, Jonas; Biliri, Evmorfia; Lampathaki, Fenareti;	Explainability as the key ingredient for AI adoption in Industry 5.0 settings	Frontiers in artificial intelligence	2023	Frontiers Media SA	Yes
Lane, Jacqueline M; Boussioux, Leonard; Ayoubi, Charles; Chen, Ying Hao; Len, Camila; Spens, Rebecca; Wagh, Pooja; Wang, Pei-Hsin;	The Explainability Edge? A Field Experiment on Generative AI-Augmented Evaluations of Early-Stage Innovations	Harvard Business School Technology & Operations Mgt. Unit Working Paper	2024		No
Huang, Shiyuan;	Exploring Explainability and Interpretability in Generative AI	Proceedings of the 20th Workshop of Young Researchers' Roundtable on Spoken Dialogue Systems	2024		No
Batic, Djordje; Stankovic, Vladimir; Stankovic, Lina;	Toward transparent load disaggregation—A framework for quantitative evaluation of explainability using explainable AI	IEEE Transactions on Consumer Electronics	2023	IEEE	Yes
Joshi, Satyadhar;	AI and Financial Model Risk Management: Applications, Challenges, Explainability, and Future Directions		2025		No
De, Paulami; Munshi, Shiladitya;	Towards Trustworthy AI: The Confluence of Ethics, Transparency, and Explainability				Yes
Sacharidis, Dimitris;	Fairness and Explainability for Enabling Trust in AI Systems	A Human-Centered Perspective of Intelligent Personalized Environments and Systems	2024	Springer	Yes
Shamsuddin, Rittika; Tabrizi, Habib B; Gottimukkula, Pavan R;	Towards responsible AI: an implementable blueprint for integrating explainability and social-cognitive frameworks in AI systems	AI Perspectives & Advances	2025	Springer	Yes
Bogaert, Jeremie; Standaert, François-Xavier;	A question on the explainability of large language models and the word-level univariate first-order plausibility assumption	arXiv preprint arXiv:2403.10275	2024		No
Zhou, Jianlong; Chen, Fang; Holzinger, Andreas;	Towards explainability for AI fairness	International workshop on extending explainable AI beyond deep models and classifiers	2020	Springer	Yes
Beaudouin, Valérie; Bloch, Isabelle; Bounie, David; Clémençon, Stéphan; d'Alché-Buc, Florence; Eagan, James; Maxwell, Winston; Mozharovskyi, Pavlo; Parekh, Jayneel;	Flexible and context-specific AI explainability: a multidisciplinary approach	arXiv preprint arXiv:2003.07703	2020		No
Arya, Vijay; Bellamy, Rachel KE; Chen, Pin-Yu; Dhurandhar, Amit; Hind, Michael; Hoffman, Samuel C; Houde, Stephanie; Liao, Q Vera; Luss, Ronny; Mojsilović, Aleksandra;	Ai explainability 360 toolkit	Proceedings of the 3rd ACM India joint international conference on data science & management of data (8th ACM IKDD CODS & 26th COMAD)	2021		Yes
Ok, Emmanuel;	Transparency and Explainability in AI Chatbots: How They Shape User Confidence		2025		No
Pavlidis, Georgios;	Unlocking the black box: analysing the EU artificial intelligence act's framework for explainability in AI	Law, Innovation and Technology	2024	Taylor & Francis	Yes
Held, Jan; Itani, Hani; Cioppa, Anthony; Giancola, Silvio; Ghanem, Bernard; Van Droogenbroeck, Marc;	X-vars: Introducing explainability in football refereeing with multi-modal large language models	Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition	2024		Yes
Yang, Lingxue; Wang, Hongrun; Deleris, Léa A;	What does it mean to explain? a user-centered study on ai explainability	International Conference on Human-Computer Interaction	2021	Springer	Yes
Pavlin, Gregor; de Villiers, Johan Pieter; Ziegler, Jürgen; Jousselme, A-L; Costa, P; Laskey, K; de Waal, Alta; Blasch, Erik; Jansen, Lennard;	Relations between explainability, evaluation and trust in ai-based information fusion systems	2021 IEEE 24th International Conference on Information Fusion (FUSION)	2021	IEEE	Yes
Liu, Chung-Feng; Chen, Zhih-Cherng; Kuo, Szu-Chen; Lin, Tzu-Chi;	Does AI explainability affect physicians' intention to use AI?	International journal of medical informatics	2022	Elsevier	Yes
Adamyk, Bogdan; Benson, Vladlena; Adamyk, Oksana; Fraczek, Bożena; Chinnaswamy, Anitha;	Barriers to Trust in AI: a Study of the Explainability Technologies Adoption in Banks	2024 14th International Conference on Advanced Computer Information Technologies (ACIT)	2024	IEEE	Yes
Hou, Tsung-Yu; Tseng, Yu-Chia; Yuan, Chien Wen Tina;	Is this AI sexist? The effects of a biased AI's anthropomorphic appearance and explainability on users' bias perceptions and trust	International Journal of Information Management	2024	Elsevier	Yes
Kempt, Hendrik; Heilinger, Jan-Christoph; Nagel, Saskia K;	Relative explainability and double standards in medical decision-making: Should medical AI be subjected to higher standards in medical decision-making than doctors?	Ethics and Information Technology	2022	Springer	Yes

Authors	Title	Source	Year	Publisher	Excl.	Decision
Stepin, Ilia; Suffian, Muhammad; Catala, Alejandro; Alonso-Moral, Jose M;	How to build self-explaining fuzzy systems: from interpretability to explainability [AI-eXplained]	IEEE Computational Intelligence Magazine	2024	IEEE		Yes
Rizzo, Matteo; Veneri, Alberto; Albarelli, Andrea; Lucchese, Claudio; Nobile, Marco; Conati, Cristina;	A theoretical framework for ai models explainability with application in biomedicine	2023 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)	2023	IEEE		Yes
Islam, Sheikh Rabiul; Eberle, William; Bundy, Sid; Ghafoor, Sheikh Khaled;	Infusing domain knowledge in ai-based" black box" models for better explainability with application in bankruptcy prediction	arXiv preprint arXiv:1905.11474	2019		x	No
Meuwissen, Mariska; Bollen, Laury;	Transparency versus Explainability in AI	Unpublished. https://doi. org/10. 13140/RG	2021		x	No
Pasquale, Frank; Malgieri, Gianclaudio;	Generative AI, explainability, and score-based natural language processing in benefits administration	Journal of Cross-disciplinary Research in Computational Law	2024			Yes
Giorgetti, Claudia; Contissa, Giuseppe; Basile, Giuseppe;	Healthcare AI, explainability, and the human-machine relationship: a (not so) novel practical challenge	Frontiers in Medicine	2025	Frontiers Media SA		Yes
Kaur, Davinder; Uslu, Suleyman; Durresi, Arjan; Badve, Sunil; Dundar, Murat;	Trustworthy explainability acceptance: A new metric to measure the trustworthiness of interpretable ai medical diagnostic systems	Complex, intelligent and software intensive systems: Proceedings of the 15th international conference on complex, intelligent and software intensive systems (CISIS-2021)	2021	Springer		Yes
Cortese, João Figueiredo Nobre Brito; Cozman, Fabio Gagliardi; Lucca-Silveira, Marcos Paulo; Bechara, Adriano Figueiredo;	Should explainability be a fifth ethical principle in AI ethics?	AI and Ethics	2023	Springer		Yes
Onitiu, Daria;	The limits of explainability & human oversight in the EU Commission's proposal for the Regulation on AI-a critical approach focusing on medical diagnostic systems	Information & Communications Technology Law	2023	Taylor & Francis		Yes
Morandini, Sofia; Fraboni, Federico; Puzzo, Gabriele; Giusino, Davide; Volpi, Lucia; Brendel, Hannah; Pietrantoni, L;	Examining the Nexus between Explainability of AI Systems and User's Trust: A Preliminary Scoping Review	Proceedings http://ceur-ws. org, ISSN	2023		x	No
Gebreyesus, Yibrah; Dalton, Damian; De Chiara, Davide; Chinnici, Marta; Chinnici, Andrea;	AI for Automating Data Center Operations: Model Explainability in the Data Centre Context Using Shapley Additive Explanations (SHAP)	Electronics	2024	MDPI		Yes
Kalusivalingam, Aravind Kumar; Sharma, Amit; Patel, Neha; Singh, Vikram;	Leveraging SHAP and LIME for Enhanced Explainability in AI-Driven Diagnostic Systems	International Journal of AI and ML	2021			Yes
Lee, ChangHyun; Cha, KyungJin;	FAT-CAT—Explainability and augmentation for an AI system: A case study on AI recruitment-system adoption	International Journal of Human-Computer Studies	2023	Elsevier		Yes
Ehsan, Upol; Riedl, Mark O;	Explainability pitfalls: Beyond dark patterns in explainable AI	Patterns	2024	Elsevier		Yes
Wang, Zheng; Wang, Jiachen; Tian, Chengyu; Ali, Ahsan; Yin, Xicheng;	Adopting AI teammates in knowledge-intensive crowdsourcing contests: the roles of transparency and explainability	Kybernetes	2024	Emerald Publishing Limited		Yes
Sharma, Yukut; Verma, Abhinav; Rao, Krisstina; Eluri, Vivek;	Reasonable explainability for regulating AI in health	ORF Issue Brief	2020		x	No
Nannini, Luca;	Habemus a Right to an Explanation: so What?–A Framework on Transparency-Explainability Functionality and Tensions in the EU AI Act	Proceedings of the AAAI/ACM conference on AI, ethics, and society	2024			Yes
Ferreira, Juliana J; Monteiro, Mateus S;	What are people doing about XAI user experience? A survey on AI explainability research and practice	Design, User Experience, and Usability. Design for Contemporary Interactive Environments: 9th International Conference, DUXU 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22	2020	Springer		Yes
Borrelli, Michael; Musch, Sean; Khan, Shabbi;	Applying the EU AI Act to the Automotive Industry: Ensuring Explainability and Transparency	Available at SSRN 4819226	2024		x	No
Schmitt, Vera; Villa-Arenas, Luis-Felipe; Feldhus, Nils; Meyer, Joachim; Spang, Robert P; Möller, Sebastian;	The role of explainability in collaborative human-ai disinformation detection	Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency	2024			Yes
Parkar, Erum; Gite, Shilpa; Mishra, Sashikala; Pradhan, Biswajeet; Alamri, Abdullah;	Comparative study of deep learning explainability and causal ai for fraud detection	International Journal on Smart Sensing and Intelligent Systems	2024	De Gruyter Poland		Yes
Raz, Aviad; Heinrichs, Bert; Avnoon, Netta; Eyal, Gil; Inbar, Yael;	Prediction and explainability in AI: Striking a new balance?	Big Data & Society	2024	SAGE Publications Sage UK: London, England		Yes
Dessain, Jean; Bentaleb, Nora; Vinas, Fabien;	Cost of explainability in ai: An example with credit scoring models	World Conference on Explainable Artificial Intelligence	2023	Springer		Yes
Cheteni, Eddwin; Vambe, William Tichaona;	Explainability in Machine Learning & AI Models for Complex Data Structures on Scorecards Development in Retail Banking	2024 4th International Multidisciplinary Information Technology and Engineering Conference (IMITEC)	2024	IEEE		Yes
Rasaq, Sodiq Oyetunji;	Transparency in AI-Driven Cloud Systems: Enhancing Trust Through Explainability				x	No
Villegas-Ch, William; García-Ortiz, Joselin; Jaramillo-Alcazar, Angel;	An Approach Based on Recurrent Neural Networks and Interactive Visualization to Improve Explainability in AI Systems	Big Data and Cognitive Computing	2023	MDPI		Yes
Sovrano, Francesco; Vitali, Fabio;	An objective metric for explainable AI: how and why to estimate the degree of explainability	Knowledge-Based Systems	2023	Elsevier		Yes
Sun, Jiao; Liao, Q Vera; Muller, Michael; Agarwal, Mayank; Houde, Stephanie; Talamadupula, Kartik; Weisz, Justin D;	Investigating explainability of generative AI for code through scenario-based design	Proceedings of the 27th International Conference on Intelligent User Interfaces	2022		x	No
Pawlicka, Aleksandra; Pawlicki, Marek; Kozik, Rafał; Kurek, Wiktor; Choraś, Michał;	How explainable is explainability? Towards better metrics for explainable AI	The International Research & Innovation Forum	2023	Springer		Yes
Shajalal, Md; Boden, Alexander; Stevens, Gunnar; Du, Delong; Kern, Dean-Robin;	Explaining AI Decisions: Towards Achieving Human-Centered Explainability in Smart Home Environments	World Conference on Explainable Artificial Intelligence	2024	Springer		Yes
Fiandrino, Claudio; Gómez, Eloy Pérez; Pérez, Pablo Fernández; Mohammadalizadeh, Hossein; Fiore, Marco; Widmer, Joerg;	AIChronoLens: advancing explainability for time series AI forecasting in mobile networks	IEEE INFOCOM 2024-IEEE Conference on Computer Communications	2024	IEEE		Yes
Hauptman, Allyson I; Schelble, Beau G; Duan, Wen; Flathmann, Christopher; McNeese, Nathan J;	Understanding the influence of AI autonomy on AI explainability levels in human-AI teams using a mixed methods approach	Cognition, Technology & Work	2024	Springer		Yes
Alsaigh, Roba; Mehmood, Rashid; Katib, Iyad;	AI explainability and governance in smart energy systems: a review	Frontiers in Energy Research	2023	Frontiers Media SA		Yes
Labati, Ruggero Donida; Coscia, Pasquale;	Explainability and Interpretability Concepts for Edge AI Systems	Advancing Edge Artificial Intelligence				Yes
Santosh, KC; Wall, Casey;	AI, Ethical Issues and Explainability-Applied Biometrics		2022	Springer		Yes
Guleria, Pratiyush; Sood, Manu;	Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling	Education and Information Technologies	2023	Springer		Yes
Gertych, Arkadiusz; Faust, Oliver;	AI explainability and bias propagation in medical decision support	Computer Methods and Programs in Biomedicine	2024	Elsevier		Yes
Shah, Arsh;	Frameworks For Improving AI Explainability Using Accountability Through Regulation and Design	Available at SSRN 3617349	2020		x	No
von Zahn, Moritz; Zacharias, Jan; Lowin, Maximilian; Chen, Johannes; Hinz, Oliver;	Navigating AI conformity: A design framework to assess fairness, explainability, and performance	Electronic Markets	2025	Springer		Yes
Morandini, Sofia; Fraboni, Federico; Hall, Mark; Quintana-Amate, Santiago; Pietrantoni, Luca;	User perspectives on AI explainability in aerospace manufacturing: a Card-Sorting study	Frontiers in Organizational Psychology	2025	Frontiers Media SA		Yes
Herm, Lukas-Valentin; Wanner, Jonas; Seubert, Franz; Janiesch, Christian;	I Don't Get IT, but IT seems Valid! The Connection between Explainability and Comprehensibility in (X) AI Research.	ECIS	2021			Yes
Karangara, Rajath; Devineni, Siva Karthik; Challa, Narayana;	Enhancing Explainability in AI Fraud Detection	Journal of Artificial Intelligence Research	2023			Yes
Maclure, Jocelyn;	AI, explainability and public reason: The argument from the limitations of the human mind	Minds and Machines	2021	Springer		Yes
Urcelay, Lucia; Hinjos, Daniel; Martin-Torres, Pablo A; González, Marta; Méndez, Marta; Civico, Salva; Alvarez-Napagao, Sergio; Garcia-Gasulla, Dario;	Exploring the Role of Explainability in AI-Assisted Embryo Selection.	CCIA	2023		x	No
Rueda, Jon; Rodríguez, Janet Delgado; Jounou, Iris Parra; Hortal-Carmona, Joaquín; Ausín, Txetxu; Rodríguez-Arias, David;	"Just" accuracy? Procedural fairness demands explainability in AI-based medical resource allocations	AI & society	2024	Springer		Yes
Grobrügge, Arne; Mishra, Nidhi; Jakubik, Johannes; Satzger, Gerhard;	Explainability in AI-Based Applications–A Framework for Comparing Different Techniques	2024 26th International Conference on Business Informatics (CBI)	2024	IEEE		Yes
Islam, Shareeful; Basheer, Nihala; Silvestri, Stefano; Papastergiou, Spyridon; Ciampi, Mario;	Intelligent Dynamic Cybersecurity Risk Management Framework with Explainability and Interpretability of AI models for Enhancing Security and Resilience of Digital Infrastructure		2024		x	No
Rasaq, Sodiq Oyetunji;	Building Trust in Cloud-Based AI: The Need for Explainability in Data Science					Yes
Li, Zhaobin;	Enhancing Human-AI Collaboration through Adaptive Interaction and Explainability	Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society	2024			Yes
Gerdes, Anne;	The role of explainability in AI-supported medical decision-making	Discover Artificial Intelligence	2024	Springer		Yes
Walke, Fabian; Bennek, Lars; Winkler, Till J;	Artificial intelligence explainability requirements of the AI act and metrics for measuring compliance	International Conference on Wirtschaftsinformatik	2023	Springer		Yes
Moltubakk Kempton, Alexander; Vassilakopoulou, Polyxeni;	Accountability, Transparency and Explainability in AI for Healthcare	European Society for Socially Embedded Technologies (EUSSET)	2021			Yes
Walter, Holmes;	Trustworthy AI: Ensuring Explainability and Bias Mitigation in Secure Deep Learning Frameworks	Journal of network and systems management	2024		x	No
Liu, Hong; Zhong, Chen; Alnusair, Awny; Islam, Sheikh Rabiul;	FAIXID: A framework for enhancing AI explainability of intrusion detection results using data cleaning techniques	Journal of network and systems management	2021	Springer		Yes
Groen, Arjan M; Kraan, Rik; Amirkhan, Shahira F; Daams, Joost G; Maas, Mario;	A systematic review on the use of explainability in deep learning systems for computer aided diagnosis in radiology: Limited use of explainable AI?	European Journal of Radiology	2022	Elsevier		Yes
Avuthu, Yogeswara Reddy;	Trustworthy AI in Cloud MLOps: Ensuring Explainability, Fairness, and Security in AI-Driven Applications	Journal of Scientific and Engineering Research	2021			Yes
Broniatowski, David; Reyna, Valerie; Vosoughi, Armin; Shearn, Ms Alaysha;	Evaluating the Explainability and Interpretability of AI Systems				x	No
Woodruff, Allison; Anderson, Yasmin Asare; Armstrong, Katherine Jameson; Gkiza, Marina; Jennings, Jay; Moessner, Christopher; Viegas, Fernanda; Wattenberg, Martin; Wrede, Fabian; Kelley, Patrick Gage;	" A cold, technical decision-maker": Can AI provide explainability, negotiability, and humanity?	arXiv preprint arXiv:2012.00874	2020		x	No
Kotsiopoulos, Thanasis; Papakostas, Gerasimos; Vafeiadis, Thanasis; Dimitriadis, Vasileios; Nizamis, Alexandros; Bolzoni, Andrea; Bellinati, Davide; Ioannidis, Dimosthenis; Votis, Konstantinos; Tzovaras, Dimitrios;	Revolutionizing defect recognition in hard metal industry through AI explainability, human-in-the-loop approaches and cognitive mechanisms	Expert Systems with Applications	2024	Elsevier		Yes

Authors	Title	Source	Year	Publisher	Decision
Elayaraja, C; Rahila, J; Velavan, P; Rajest, S Suman; Shynu, T; Rahman, Md Mehdi;	Depth Sensing in AI on Exploring the Nuances of Decision Maps for Explainability	Optimizing Intelligent Systems for Cross-Industry Application	2024	IGI Global	Yes
Arya, Vijay; Bellamy, Rachel KE; Chen, Pin-Yu; Dhurandhar, Amit; Hind, Michael; Hoffman, Samuel C; Houde, Stephanie; Liao, Q Vera; Luss, Ronny; Mojsilović, Aleksandra;	One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques	arXiv preprint arXiv:1909.03012	2019		No
Hooper, Kerrie; Lunn, Stephanie;	A Scoping Review of Transparency and Explainability in AI Ethics Guidelines	The International FLAIRS Conference Proceedings	2024		Yes
Ontika, Nazmun Nisat; Syed, Hussain Abid; Saßmannshausen, Sheree May; Harper, Richard HR; Chen, Yunan; Park, Sun Young; Grisot, Miria; Chow, Astrid; Blaumer, Nils; Pinatti de Carvalho, Aparecido Fabiano;	Exploring human-centered AI in healthcare: diagnosis, explainability, and trust	European Society for Socially Embedded Technologies (EUSSET)	2022		Yes
Zhang, Tongze; Chung, Tammy; Dey, Anind; Bae, Sang Won;	Exploring Algorithmic Explainability: Generating Explainable AI Insights for Personalized Clinical Decision Support Focused on Cannabis Intoxication in Young Adults	2024 International Conference on Activity and Behavior Computing (ABC)	2024	IEEE	Yes
Zhang, Yuhao; McAreavey, Kevin; Liu, Weiru;	Developing and experimenting on approaches to explainability in AI systems	ICAART 2022: 14th International Conference on Agents and Artificial Intelligence	2022		Yes
Lee, Alvina Hui Shan; Shankararaman, Venky; Ouh, Eng Lieh;	Vision Paper: Advancing of AI Explainability for the Use of ChatGPT in Government Agencies–Proposal of A 4-Step Framework	2023 IEEE International Conference on Big Data (BigData)	2023	IEEE	Yes
Pierce, Robin L; Van Biesen, Wim; Van Cauwenberge, Daan; Decruyenaere, Johan; Sterckx, Sigrid;	Explainability in medicine in an era of AI-based clinical decision support systems	Frontiers in genetics	2022	Frontiers Media SA	Yes
Ahmadi, Neda; Nilashi, Mehrbakhsh;	Does Explainability Enhance the Effectiveness of AI Models in Public Health? The COVID-19 Context	Journal of Soft Computing and Decision Support Systems	2024		Yes
Mensah, George Benneh;	Ensuring AI Explainability in Clinical Decision Support Systems				No
Fraser, Henry; Simcock, Rhyle; Snoswell, Aaron J;	Ai opacity and explainability in tort litigation	Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency	2022		Yes
Charalampakos, Foivos; Tsouparopoulos, Thomas; Koutsopoulos, Iordanis;	Joint Explainability-Performance Optimization With Surrogate Models for AI-Driven Edge Services	arXiv preprint arXiv:2503.07784	2025		No
Iyer, Vish; Manshad, Moe; Brannon, Daniel;	A Value-Based Approach to AI Ethics: Accountability, Transparency, Explainability, and Usability	Mercados y Negocios	2025	Universidad de Guadalajara	Yes
Oussalah, Mourad;	AI explainability. A bridge between machine vision and natural language processing	Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10–15, 2021, Proceedings, Part III	2021	Springer	No
Kempton, Alexander Moltubakk; Vassilakopoulou, Polyxeni;	Accountability, transparency and explainability in AI for healthcare	8th International Conference on Infrastructures in Healthcare	2021		Yes
Ghosh, Promila; Mondal, Amit Kumar; Chatterjee, Sajib; Masud, Mehedi; Meshref, Hossam; Bairagi, Anupam Kumar;	Recognition of sunflower diseases using hybrid deep learning and its explainability with AI	Mathematics	2023	MDPI	Yes
Ferreira, Juliana Jansen; Monteiro, Mateus de Souza;	Do ML Experts Discuss Explainability for AI Systems? A discussion case in the industry for a domain-specific solution	arXiv preprint arXiv:2002.12450	2020		No
Ahmed, Aminu Jansen; Abdullahi, Ali Usman; Gital, Abdulsalam Ya'u; Dutse, Abubakar Yusuf;	The Role of Metadata in Promoting Explainability and Interoperability of AI-based Prediction Models	Journal of Exceptional Multidisciplinary Research	2024		Yes
Jin, Weina; Li, Xiaoxiao; Hamarneh, Ghassan;	The XAI alignment problem: Rethinking how should we evaluate human-centered AI explainability techniques	arXiv preprint arXiv:2303.17707	2023		No
Kollias, Dimitrios; Arsenos, Anastasios; Kollias, Stefanos;	Domain adaptation explainability & fairness in ai for medical image analysis: Diagnosis of covid-19 based on 3-d chest ct-scans	Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition	2024		Yes
Thurzo, Andrej;	Provable AI Ethics and Explainability in Next-Generation Medical and Educational AI agents: Trustworthy Ethical Firewall		2025	Preprints	No
Rezaeian, Olya; Bayrak, Alparslan Emrah; Asan, Onur;	Explainability and AI Confidence in Clinical Decision Support Systems: Effects on Trust, Diagnostic Performance, and Cognitive Load in Breast Cancer Care	arXiv preprint arXiv:2501.16693	2025		No
Arazo, Eric; Stoev, Hristo; Bosch, Cristian; Suárez-Cetrulo, Andrés L; Simón-Carbajo, Ricardo;	Xpression: A Unifying Metric to Optimize Compression and Explainability Robustness of AI Models	World Conference on Explainable Artificial Intelligence	2024	Springer	Yes
Kaur, Davinder;	Trustworthy AI: Ensuring Explainability and Acceptance		2023	Purdue University	No
Suh, Ashley; Hurley, Isabelle; Smith, Nora; Siu, Ho Chit;	Fewer Than 1% of Explainable AI Papers Validate Explainability with Humans	arXiv preprint arXiv:2503.16507	2025		No
Nalepa, Grzegorz; Araszkiewicz, Michał; Nowaczyk, Sławomir; Bobek, Szymon;	Building trust to ai systems through explainability: technical and legal perspectives		2019		No
Sovrano, Francesco; Sapienza, Salvatore; Palmirani, Monica; Vitali, Fabio;	Metrics, explainability and the European AI act proposal	J	2022	MDPI	Yes
Santosh, KC; Wall, Casey;	AI, Ethical Issues, and Explainability for Biometrics—Summary and What Else?	AI, ethical issues and explainability—applied biometrics	2022	Springer	Yes
Solaiman, Barry; Bloom, Mark G;	AI, explainability, and safeguarding patient safety in Europe	The Future of Medical Device Regulation	2022	Cambridge University Press Cambridge	Yes
Tsolakis, Nikos; Maga-Nteve, Christoniki; Vrochidis, Stefanos; Bassiliades, Nick; Meditskos, Georgios;	Enhancing Cardiac AI Explainability Through Ontology-Based Evaluation	2024 15th International Conference on Information, Intelligence, Systems & Applications (IISA)	2024	IEEE	Yes
Kibria, Mohammad Golam; Kucirka, Lauren; Mostafa, Javed;	Assessing AI Explainability: A Usability Study Using a Novel Framework Involving Clinicians	arXiv preprint arXiv:2503.16920	2025		No
Balasubramaniam, Nagadivya; Kauppinen, Marjo; Truong, Hong-Linh; Kujala, Sari;	Candidate Solutions for Defining Explainability Requirements of AI Systems	International Working Conference on Requirements Engineering: Foundation for Software Quality	2024	Springer	Yes
Wajahat, Iram; Singh, Amritpal; Keshtkar, Fazel; Bukhari, Syed Ahmad Chan;	A Symbolic AI Framework for Enhanced Diabetes Prognosis Accuracy and Explainability	2024 IEEE International Conference on Future Machine Learning and Data Science (FMLDS)	2024	IEEE	Yes
Droste, Jakob; Deters, Hannah; Fuchs, Ronja; Schneider, Kurt;	Peeking Outside the Black-Box: AI Explainability Requirements beyond Interpretability.	REFSQ Workshops	2024		No
Tan, Myles Joshua Toledo; Benos, Panayiotis V;	Addressing Intersectionality, Explainability, and Ethics in AI-Driven Diagnostics: A Rebuttal and Call for Transdisciplinary Action	arXiv preprint arXiv:2501.08497	2025		No
Sarker, Iqbal H;	AI-driven cybersecurity and threat intelligence: cyber automation, intelligent decision-making and explainability		2024	Springer Nature	Yes
Sheu, Ruey-Kai; Pardeshi, Mayuresh Sunil;	A survey on medical explainable AI (XAI): recent progress, explainability approach, human interaction and scoring system	Sensors	2022	MDPI	Yes
Tecks, Austin; Peschlow, Thomas; Vigliensoni, Gabriel;	Explainability paths for sustained artistic practice with AI	arXiv preprint arXiv:2407.15216	2024		No
Torda, Tommaso;	Interpretability and explainability AI in advanced neural Networks for topological quantum field Theory and medical application	Università degli Studi di Roma" La Sapienza"	2024		No
Border, Samuel P; Sarder, Pinaki;	From what to why, the growing need for a focus shift toward explainability of AI in digital pathology	Frontiers in Physiology	2022	Frontiers Media SA	Yes
Shin, Donghee;	Explainable AI: how explainability impacts the human interaction with AI	Social Science Asia	2022		Yes
Noah, Oliver;	Evaluating Explainability in AI Models: Comparing SHAP, LIME, and Other Techniques		2024		No
McNamara, Stephanie L; Yi, Paul H; Lotter, William;	The clinician-AI interface: intended use and explainability in FDA-cleared AI devices for medical image interpretation	NPJ Digital Medicine	2024	Nature Publishing Group UK London	Yes
Luk, Chiu-Ying; Chung, Hoi-Lam; Yim, Wai-Kuen; Leung, Ching-Wah;	Regulating generative AI: Ethical considerations and explainability benchmarks		2024	OSF	No
Quinn, Mary; Piper, Blake; Bliss, James P; Keever, David;	Recommended methods for using the 2020 NIST principles for AI explainability	2020 IEEE International Conference on Big Data (Big Data)	2020	IEEE	Yes
El Oualidi, Taoufik; Assar, Saïd;	Bridging explainability and interpretability in AI-driven SCM projects to enhance decision-making	ITM Web of Conferences	2024	EDP Sciences	Yes
Melsión, Gaspar Isaac; Torre, Ilaria; Vidal, Eva; Leite, Iolanda;	Using explainability to help children understand gender bias in AI	Proceedings of the 20th Annual ACM Interaction Design and Children Conference	2021		Yes
Encarnação, L Miguel; Kohlhammer, Jörn; Steed, Chad A;	Visualization for ai explainability	IEEE Computer Graphics and Applications	2022	IEEE	Yes
Rizzo, Matteo; Veneri, Alberto; Albarelli, Andrea; Lucchese, Claudio; Conati, Cristina;	A theoretical framework for ai models explainability	arXiv preprint arXiv:2212.14447	2022		No
Umbrello, Steven; Yampolskiy, Roman V;	Designing AI for explainability and verifiability: a value sensitive design approach to avoid artificial stupidity in autonomous vehicles	International Journal of Social Robotics	2022	Springer	Yes
Etori, Naome A; Gini, Maria;	WisCompanion: Integrating the Socratic Method with ChatGPT-Based AI for Enhanced Explainability in Emotional Support for Older Adults	International Conference on Human-Computer Interaction	2024	Springer	Yes
Yu, Yunxia;	Hybrid models for accuracy and explainability in AI systems	Third International Conference on Electronic Information Engineering, Big Data, and Computer Technology (EIBDCT 2024)	2024	SPIE	Yes
Bidkar, Deepti Vinayak;	A literature review to underline necessity of explainability in AI and discuss existing explainable AI techniques		2021		Yes
Herrmann, Markus; Wabro, Andreas; Winkler, Eva;	Percentages and reasons: AI explainability and ultimate human responsibility within the medical field	Ethics and Information Technology	2024	Springer	Yes
Lu, Yuhang;	Enhancing Explainability and Performance in AI-Driven Face Recognition and Deepfake Detection		2024	EPFL	No
Rajuroy, Adam; Emmanuel, Mr;	Enhancing AI Autonomy and Explainability in Real-Time Decision-Making.				No
Hempel, Philip; Bender, Theresa; Spicher, Nicolai;	Enhancing explainability in ECG analysis through evidence-based AI interpretability	2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)	2024	IEEE	Yes
Kühl, Niklas; Meske, Christian; Nitsche, Maximilian; Lobana, Jodie;	Investigating the Role of Explainability and AI Literacy in User Compliance	arXiv preprint arXiv:2406.12660	2024		No
Li, Ron C; Muthu, Naveen; Hernandez-Boussard, Tina; Dash, Dev; Shah, Nigam H;	Explainability in Medical AI	Intelligent Systems in Medicine and Health: The Role of AI	2022	Springer	Yes
Norkute, Milda; Herger, Nadja; Michalak, Leszek; Mulder, Andrew; Gao, Sally;	Towards explainable ai: Assessing the usefulness and impact of added explainability features in legal document summarization	Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems	2021		Yes
Stadler, Alexandra; Müller, Werner G; Harman, Radoslav;	Green LIME: Improving AI Explainability through Design of Experiments	arXiv preprint arXiv:2502.12753	2025		No
Alexandropoulou, Christina-Athanasia; Panagiotopoulos, Ilias; Kleanthous, Styliani; Dimitrakopoulos, George; Constantinou, Ioannis; Politi, Eleni; Ntalaperas, Dimitrios; Papageorgiou, Xanthi; Stylianides, Charithea; Ioannides, Nikos;	AI-enabled solutions, explainability and ethical concerns for predicting sepsis in ICUs: a systematic review	2023 IEEE 19th International Conference on e-Science (e-Science)	2023	IEEE	Yes

Authors	Title	Venue	Year	Publisher	Mark	Incl.
Li, Xinze; Lin, Fanfan; Mantooth, Homer Alan; Rodríguez-Andina, Juan José;	AI Explainability for Power Electronics: From a Lipschitz Continuity Perspective	arXiv preprint arXiv:2501.09948	2025		x	No
Prama, Tabia Tanzin; Amrin, Jannatul Ferdaws; Anwar, Md Mushfique; Sarker, Iqbal H;	AI Enabled User-Specific Cyberbullying Severity Detection with Explainability	arXiv preprint arXiv:2503.10650	2025		x	No
Denzel, Philipp; Brunner, Stefan; Luley, Paul-Philipp; Frischknecht-Gruber, Carmen; Reif, Monika Ulrike; Schilling, Frank-Peter; Amini, Amin; Repetto, Marco; Iranfar, Arman; Weng, Joanna;	A framework for assessing and certifying explainability of health-oriented ai systems	Explainable AI in Medicine Workshop, Lugano, Switzerland, 2-3 November 2023	2023		x	No
Fiandrino, Claudio; Bonati, Leonardo; D'Oro, Salvatore; Polese, Michele; Melodia, Tommaso; Widmer, Joerg;	Explora: AI/ml explainability for the open ran	Proceedings of the ACM on Networking	2023	ACM New York, NY, USA		Yes
Galli, Antonio; La Gatta, Valerio; Moscato, Vincenzo; Postiglione, Marco; Sperlì, Giancarlo;	Explainability in AI-based behavioral malware detection systems	Computers & Security	2024	Elsevier		Yes
Walke, Fabian; Bennek, Lars; Winkler, Till J;	AI in Government: A Study on Explainability of High-Risk AI-Systems in Law Enforcement and Police Service	International Conference on Wirtschaftsinformatik	2023	Springer		Yes
Adenekan, Tobiloba Kollawole;	Adaptive Prompt Learning for Generative AI: Ensuring Explainability and Fairness in Financial and Real Estate Applications		2024		x	No
Wasserman-Rozen, Hofit; Gilad-Bachrach, Ran; Elkin-Koren, Niva;	Lost in translation: the limits of explainability in AI	Cardozo Arts & Ent. LJ	2024	HeinOnline		Yes
Koul, Amrita; Gupta, Sachin; Gupta, Bhoomi;	AI with Consciousness: Advancing Explainability in the Healthcare Sector	2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)	2023	IEEE		Yes
Rasaq, Sodiq Oyetunji;	Ensuring Fairness in Cloud-Based AI: The Role of Explainability in Ethical Decision-Making					Yes
Hemment, Drew; Murray-Rust, Dave; Belle, Vaishak; Aylett, Ruth; Vidmar, Matjaz; Broz, Frank;	Experiential AI: Enhancing explainability in artificial intelligence through artistic practice		2022		x	No
Talypova, Dinara; Wintersberger, Philipp;	Explainability Challenges in Continuous Invisible AI for Self-Augmentation		2023			Yes
Duque Anton, Simon D; Schneider, Daniel; Schotten, Hans D;	On Explainability in AI-Solutions: A Cross-Domain Survey	International Conference on Computer Safety, Reliability, and Security	2022	Springer		Yes
Patibandla, Rama Seetha Maha Lakshmi; Yenduri, Gokul; Maddikunta, Praveen Kumar Reddy; Murugan, Ramalingam; Gadekallu, Thippa Reddy;	Explainability AI: A Key Driver for Enhanced Quality of Service in Actuarial Practices	Artificial Intelligence and Actuarial Science	2024	Chapman and Hall/CRC		Yes
Zhang, Jiajin; Chao, Hanqing; Dasegowda, Giridhar; Wang, Ge; Kalra, Mannudeep; Yan, Pingkun;	Quantifying Trustworthiness of Explainability in Medical AI		2022		x	No
Raposo, Vera Lúcia;	The fifty shades of black: about black box AI and explainability in healthcare	Medical Law Review	2025	Oxford University Press		Yes
Zhang, Tianyue; He, Xin;	Enhancing Explainability in Medical AI: Developing Human-Centered Participatory Design Cards	International Conference on Human-Computer Interaction	2024	Springer		Yes
Ewerz, Benjamin; Hoefler, Margit; Marx, Cyril; Moertl, Peter;	How to facilitate explainability of AI for Increased User Trust: Results of a Study with a COVID-19 Risk Calculator	2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)	2021	IEEE		Yes
Tandon, Sara; Wang, Jennifer;	Surfacing AI Explainability in Enterprise Product Visual Design to Address User Tech Proficiency Differences	Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems	2023			Yes
Berg, TWAN;	A Tool Combining Explainability and Benefits of Artificial Intelligence to Improve Trust in AI in Agriculture		2024	University of Twente	x	No
Gomes, Hugo; Ferreira, João; Rodrigues, Manuel;	Insights into Quantitative Evaluation Techniques for AI Explainability	International Symposium on Distributed Computing and Artificial Intelligence	2024	Springer		Yes
Arya, Vijay; Bellamy, Rachel KE; Chen, Pin-Yu; Dhurandhar, Amit; Hind, Michael; Hoffman, Samuel C; Houde, Stephanie; Liao, Q Vera; Luss, Ronny; Mojsilović, Aleksandra;	Ai explainability 360: Impact and design	Proceedings of the AAAI conference on artificial intelligence	2022			Yes
Martinez, Samantha Escobar;	Subjective evaluation of AI explainability methods and their applicability to chest x-rays				x	No
Mars, Clodéric; Dès, Rémi; Boussard, Matthieu;	Explainable AI: The three stages of Explainable AI: How explainability facilitates real world deployment of AI	Conference Humains et IA, Travailler en Intelligence	2020			Yes
Griesmayer, Max;	Explainable AI: a comparison on explainability methods in the field of facial recognition		2024	Technische Universität Wien	x	No
Smyth, Dr Sandra;	Systems Integration Model For AI Solutions with Explainability Requirements	Available at SSRN 5169810	2025		x	No
Holzinger, Andreas; Müller, Heimo;	Toward human–AI interfaces to support explainability and causability in medical AI	Computer	2021	IEEE		Yes
Iqbal, Junaid; Eldred, Addison;	Symbolic AI Meets Deep Learning: A Hybrid Approach to Improving Explainability and Predictive Accuracy		2025		x	No
Matulionyte, Rita; Aranovich, Tatiana;	Trade secrets versus the AI explainability principle	Research handbook on intellectual property and artificial intelligence	2022	Edward Elgar Publishing	x	No
Zhang, Xin; Sheng, Victor S;	Neuro-Symbolic AI: Explainability, Challenges, and Future Trends	arXiv preprint arXiv:2411.04383	2024		x	No
Zhu, Ruiqi; Christensen, Cecilie; Zarrin, Bahram; Bækgaard, Per; Alstrøm, Tommy Sonne;	Towards Trustworthy AI in Demand Planning: Defining Explainability for Supply Chain Management					Yes
Simón-Carbajo, Ricardo;	Xpression: A unifying metric to optimize Compression and Explainability robustness of AI models					Yes
Gagnon, Elisa; Deregt, Anouk; Lapointe, Liette;	Clarity in Complexity: Advancing AI Explainability through Sensemaking		2025			Yes
Zhang, Lufan; Scifleet, Paul;	Charting the Transformation of Enterprise Information Management: AI Explainability and Transparency in EIM Practice					Yes
Pérez, Pablo Fernández; Fiandrino, Claudio; Gómez, Eloy Pérez; Mohammadalizadeh, Hossein; Fiore, Marco; Widmer, Joerg;	AIChronoLens: AI/ML Explainability for Time Series Forecasting in Mobile Networks	IEEE Transactions on Mobile Computing	2025	IEEE		Yes
van der Veer, Sabine N; Riste, Lisa; Cheraghi-Sohi, Sudeh; Phipps, Denham L; Tully, Mary P; Bozentko, Kyle; Atwood, Sarah; Hubbard, Alex; Wiper, Carl; Oswald, Malcolm;	Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries	Journal of the American Medical Informatics Association	2021	Oxford University Press		Yes
Sydorenko, Tetiana;	The Impact of AI Explainability on Cognitive Dissonance and Trust in Human-AI Recruitment Teams					Yes
Pramanik, Vishal; Maliha, Maisha; Jha, Sumit Kumar;	Enhancing Integrated Gradients Using Emphasis Factors and Attention for Effective Explainability of Large Language Models				x	No
Man, Siu Shing; Wen, Huiying; Huang, Jingyu; Zhang, Tingru;	Enhancing Public Trust and Acceptance of AI-Assisted Medical Diagnosis Systems Through Explainability	Available at SSRN 5090771			x	No
Sure, Tharun Anand Reddy; Chittibala, Dinesh Reddy;	LEVERAGING USER FEEDBACK FOR ENHANCING EXPLAINABILITY IN AI APPLICATIONS					Yes
Samudrala, Vamshi Krishna; Rao, Vallu Visrutatma; Pulakhandam, Winner;	IoMT Platforms for Advanced AI-Powered Skin Lesion Identification: Enhancing Model Interpretability, Explainability, and Diagnostic Accuracy with CNN and Score-CAM to Significantly Improve Healthcare Outcomes				x	No
Tan, Ting Fang; Elangovan, Kabilan; Ong, Jasmine Chiat Ling; Lee, Aaron; Shah, Nigam H; Sung, Joseph JY; Wong, Tien Yin; Lan, Xue; Liu, Nan; Wang, Haibo;	A Proposed SCORE Evaluation Framework for Large Language Models–Safety, Consensus & Context, Objectivity, Reproducibility and Explainability	Consensus & Context, Objectivity, Reproducibility and Explainability			x	No
Filogna, Silvia; Malizia, Alessio; Mazzei, Daniele; Prencipe, Giuseppe; Sgandurra, Giuseppina; Turchi, Tommaso;	Telemedicine and AI: From Co-Design to Explainability	2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI)	2024	IEEE		Yes
Smirnov, Alexander; Ponomarev, Andrew; Agafonov, Anton;	Ontology-Based Neuro-Symbolic AI: Effects on Prediction Quality and Explainability	IEEE Access	2024	IEEE		Yes
Pawlicki, Marek; Pawlicka, Aleksandra; Srutek, Mscislaw; Kozik, Rafal; Choras, Michal;	Interpreting Intrusions-The Role of Explainability in AI-Based Intrusion Detection Systems	International Conference on Computer Recognition Systems	2023	Springer		Yes
Chen, Yanfang; Chen, Ding; Song, Shichao; Niu, Simin; Wang, Hanyu; Tang, Zeyun; Xiong, Feiyu; Li, Zhiyu;	Hrde: Retrieval-augmented large language models for chinese health rumor detection and explainability	arXiv preprint arXiv:2407.00668	2024		x	No
Tanwar, Anita; Aggarwal, Rashmi; Kaur, Arvinder;	Explainability of AI in Improving the Quality of Service in the Actuarial Industry	Artificial Intelligence and Actuarial Science	2024	Chapman and Hall/CRC		Yes
Spreeuwenberg, Silvie;	Choose for AI and for explainability	OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"	2019	Springer		Yes
Holm, Sune;	On the justified use of AI decision support in evidence-based medicine: validity, explainability, and responsibility	Cambridge Quarterly of Healthcare Ethics	2023	Cambridge University Press		Yes
Sovrano, Francesco; Sapienza, Salvatore; Palmirani, Monica; Vitali, Fabio;	A survey on methods and metrics for the assessment of explainability under the proposed AI act	Legal Knowledge and Information Systems	2021	IOS Press	x	No
Alexander, Francis J; Borders, Tammie; Sheffield, Angie; Wonders, Marc;	Workshop report for next-gen AI for proliferation detection: Accelerating the development and use of explainability methods to design AI systems suitable for nonproliferation mission applications		2020	Brookhaven National Lab. (BNL), Upton, NY (United States); Idaho National Lab ...	x	No
Lognoul, Michael;	Explainability of AI Tools in Private Sector: An Attempt for Systemization	Available at SSRN 3685906	2020		x	No
Rasaq, Sodiq Oyetunji;	Regulating AI Ethics: The Importance of Explainability in Governance and Compliance				x	No
EL OUALIDI, Taoufik; Assar, Saïd;	Does AI Explainability Meet End-Users' Requirements? Insights From A Supply Chain Management Case Study		2024			Yes
Giavina-Bianchi, Mara; Vitor, William Gois; Fornasiero de Paiva, Victor; Okita, Aline Lissa; Sousa, Raquel Machado; Machado, Birajara;	Explainability agreement between dermatologists and five visual explanations techniques in deep neural networks for melanoma AI classification	Frontiers in Medicine	2023	Frontiers Media SA		Yes
Narkhede, Jeet;	Comparative Evaluation of Post-Hoc Explainability Methods in AI: LIME, SHAP, and Grad-CAM	2024 4th International Conference on Sustainable Expert Systems (ICSES)	2024	IEEE		Yes
Matulionyte, Rita;	Reconciling trade secrets and AI explainability: face recognition technologies as a case study	R Matulionyte,'Reconciling Trade Secrets and Explainable AI: face recognition technology as a case study'(2022)	2021			Yes
Müftüoğlu, Z; Kızrak, MA; Yıldırım, T;	Privacy-preserving mechanisms with explainability in assistive AI technologies	Advances in Assistive Technologies: Selected Papers in Honour of Professor Nikolaos G. Bourbakis–Vol. 3	2021	Springer		Yes
Reinhold, Samira Christine;	Explainability in Medical AI-Legal Regulation of Explainability in Clinical Decision Support Systems		2021		x	No
San Segundo-Gonzalo, Alberto; Mohino-Herranz, Inma; Alonso-Diaz, Sagrario; G Gallegos, Jaime; Perez-Sanz, Fernando Jose; Cornejo-Bueno, Laura Maria;	AI-Based Classification of Eeg Signals for Inner and Pronounced Speech and Visualized Cognitive Tasks: A Machine Learning and Explainability Approach				x	No
Khosravi, Pasha;	Applications of Probabilistic Reasoning in Trustworthy AI: From Handling Missing Data to Explainability		2023	University of California, Los Angeles	x	No

Authors	Title	Source	Year	Publisher	Notes	Included
Richard, Guy-Junior; Habonneau, Jérôme; Guériot, Didier; Le Caillec, Jean-Marc;	AI Explainability and Acceptance: A Case Study for Underwater Mine Hunting	ACM Journal of Data and Information Quality	2024	ACM New York, NY		Yes
Xu, Bo; Bhatti, Uzair Aslam; Tang, Hao; Yan, Jialin; Wu, Shulei; Sarhan, Nadia; Awwad, Emad Mahrous; MS, Syam; Ghadi, Yazeed Yasin;	Towards explainability for AI-based edge wireless signal automatic modulation classification	Journal of Cloud Computing	2024	Springer		Yes
Sacramed, Michael T;	Reviewing the Philippines Legal Landscape of Artificial Intelligence (AI) in Business: Addressing Bias, Explainability, and Algorithmic Accountability	International Journal of Research and Innovation in Social Science	2024	International Journal of Research and Innovation in Social Science (IJRISS)		Yes
Leventi-Peetz, A-M; Östreich, T; Lennartz, Werner; Weber, Kai;	Scope and sense of explainability for ai-systems	Intelligent Systems and Applications: Proceedings of the 2021 Intelligent Systems Conference (IntelliSys) Volume 1	2022	Springer		Yes
Okolo, Chinasa T;	AI Explainability in the Global South: Towards an Inclusive Praxis for Emerging Technology Users		2023	Cornell University		Yes
Al Noman, Abdullah; Arif, Abu Shamim Mohammad;	Recognition of brain tumors using deep learning with data balancing techniques and its explainability with AI	2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)	2023	IEEE		Yes
Murray, David; Stankovic, Lina; Stankovic, Vladimir;	Transparent AI: Explainability of deep learning based load disaggregation	Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation	2021			Yes
Quach, Luyl-Da; Quoc, Khang Nguyen; Nguyen, Chi-Ngon; Thai-Nghe, Nguyen;	Explainable AI for Plant Disease Detection: Assessing Explainability in Classifying Maize Leaves Diseases with Focus Score and Ablation-CAM	International Conference on Intelligent Systems and Data Science	2024	Springer		Yes
Quattrocchio, Letizia;	Integration of Uncertainty into Explainability Methods to Enhance AI Transparency in Brain MRI Classification		2024	Politecnico di Torino	x	No
Kritikos, Mihalis; Iphofen, Ron;	The Explainability and Transparency of AI Monitoring Tools	Ethical AI Surveillance in the Workplace	2023	Emerald Publishing Limited		Yes
Gjona, Ermela;	Human-Centered Explainability Attributes In AI-Powered Eco-Driving: Understanding Truck Drivers' Perspective		2023		x	No
Lee, Changhyun; Kwon, Hun Yeong; Cha, Kyung Jin;	Human Cognition for Mitigating the Paradox of AI Explainability: A Pilot Study on Human Gaze-based Text Highlighting	The International FLAIRS Conference Proceedings	2024			Yes
Christian, Emedo Chinyere;	Explainability Imperative of Generative Artificial Intelligence Navigating the Moral Dilemma of AI in Nigeria and Charting a Path for the Future	Universal Library of Arts and Humanities	2024		x	No
PINNA, Emanuele;	Applications of Generative AI via Latent Diffusion for Explainability in Geolocalization		2024		x	No
Pithayarungsarit, Pawinee; Rieger, Tobias; Onnasch, Linda; Roesler, Eileen;	The pop-out effect of rarer occurring stimuli shapes the effectiveness of AI explainability	Proceedings of the Human Factors and Ergonomics Society Annual Meeting	2024	SAGE Publications Sage CA: Los Angeles, CA		Yes
Mathur, Niharika; Zubatiy, Tamara; Rozga, Agata; Mynatt, Elizabeth;	"Why Did You Say That?": Understanding Explainability in Conversational AI Systems for Older Adults with Mild Cognitive Impairment (MCI)	International Conference on Ubiquitous Computing and Ambient Intelligence	2023	Springer		Yes
Simkute, Auste;	Reimagining explainability in expert contexts: setting foundations for effective expert-AI interactions through design and collaboration		2024	The University of Edinburgh	x	No
Xi, Jianing; Wang, Dan; Yang, Xuebing; Zhang, Wensheng; Huang, Qinghua;	Cancer omic data based explainable AI drug recommendation inference: A traceability perspective for explainability	Biomedical Signal Processing and Control	2023	Elsevier		Yes
He, Gaole; Balayn, Agathe; Buijsman, Stefan; Yang, Jie; Gadiraju, Ujwal;	Opening the Analogical Portal to Explainability: Can Analogies Help Laypeople in AI-assisted Decision Making?	Journal of Artificial Intelligence Research	2024			Yes
Mesin, Ante;	How Distributed Ledger Technology (DLT) Can Support Explainability in AI-based Algorithmic Trading				x	No
Debnath, Anjan; Hasan, Md Mahedi; Raihan, M; Samrat, Nadim; Alsulami, Mashael M; Masud, Mehedi; Bairagi, Anupam Kumar;	A smartphone-based detection system for tomato leaf disease using efficientNetV2B2 and its explainability with artificial intelligence (AI)	Sensors	2023	MDPI	x	No
Hasan, Md Mahedi; Fahim Habib, Md Sharzul Mostafa;	Bangla Handwritten Character Recognition Using Deep Learning Approaches and its Explainability with AI				x	No
AGAFONOV, ANTON;	Ontology-Based Neuro-Symbolic AI: Effects on Prediction Quality and Explainability		2024		x	No
Novak, Jasminko; Maljur, Tina; Drenska, Kalina;	Transferring ai explainability to user-centered explanations of complex Covid-19 information	International Conference on Human-Computer Interaction	2022	Springer		Yes
Sewnath, Ghanshaam; Crijnen, Jur;	How much is too much? Levels of AI Explainability within Decision Support System's User Interfaces for improved decision-making performance	Proceedings MSG-184 Towards Training and Decision Support for Complex Multi-Domain Operations	2021			Yes
Sporsem, Tor; Finserås, Stine Rasdal; Strümke, Inga;	Clinicians don't know what explanations they need: A case study on eliciting AI software explainability requirements	arXiv preprint arXiv:2501.09592	2025		x	No
Rueda Etxebarria, Jon;	"Just" accuracy? Procedural fairness demands explainability in AI-based medical resource allocations		2022	Springer Nature		Yes
Farahmand, Fariborz;	Commonsense for AI: an interventional approach to explainability and personalization	AI & SOCIETY	2024	Springer		Yes
Alsafran, Ahmed Sulaiman; Hassan, Abdelrahman; Abusada, Faisal; Alaraj, Moath M;	Challenges and Solutions for AI Explainability in Smart Grid Literature Review	Available at SSRN 4807531	2024		x	No
Ganapavarapu, Giridhar; Mukherjee, Sumanta; Martinez Gil, Natalia; Sarpatwar, Kanthi; Rajasekharan, Amaresh; Dhurandhar, Amit; Arya, Vijay; Vaculin, Roman;	AI Explainability 360 toolkit for time-series and industrial use cases	Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining	2023		x x	No
Ji, Jiazhou; Li, Ruizhe; Li, Shujun; Guo, Jie; Qiu, Weidong; Huang, Zheng; Chen, Chiyu; Jiang, Xiaoyu; Lu, Xinru;	Detecting machine-generated texts: Not just" ai vs humans" and explainability is complicated	arXiv preprint arXiv:2406.18259	2024		x	No
Park, Minjung; Chai, Sangmi;	AI model for predicting legal judgments to improve accuracy and explainability of online privacy invasion cases	Applied Sciences	2021	MDPI		Yes
Serrano, Carlos Entrena;	The risks of opening the car bonnet: A semi-systematic review of explainability in AI systems		2021	Aalborg University Copenhagen	x	No
Srinivasan, Srikanth Kaniyanoor; Krishna, V; Sahoo, Harsha Vardhan;	A methodology to use ai for automotive safety using model explainability		2021	Technical report, EasyChair	x	No
Vallayil, Manju; Nand, Parma; Yan, Wei Qi;	Explainable AI through Thematic Clustering and Contextual Visualization: Advancing Macro-Level Explainability in AFV Systems	Indo-American Journal of Pharma and Bio Sciences	2024		x	No
Sitaraman, Surendar Rama; Alagarsundaram, Poovendran; Kumar, Veerandra;	AI-Driven Skin Lesion Detection with CNN and Score-CAM: Enhancing Explainability in IoMT Platforms		2024		x	No
Ericstam, Sonia Bastigkeit;	AI in the Workplace: Regulating Explainability and Consent in Algorithmic Management	Digital Governance: Confronting the Challenges Posed by Artificial Intelligence	2024	Springer		Yes
Rinta-Kahila, Tapani; Someh, Ida; Indulska, Marta; Ryan, Ian; van Leent, Ryan;	Building explainability into public sector artificial intelligence: helping stakeholders understand the thinking behind AI decision-making	The University of Queensland; SAP SE	2021		x	No
Bastigkeit Ericstam, Sonia;	AI in the Workplace: Regulating Explainability and Consent in Algorithmic Management	Faculty of Law, Stockholm University Research Paper	2024		x	No
Erdoğanyılmaz, Cihan; Mengünoğlul, Berkay; Balci, Muhammet;	Unveiling the black box: Investigating the interplay between ai technologies, explainability, and legal implications	2023 8th International Conference on Computer Science and Engineering (UBMK)	2023	IEEE		Yes
Fiorino, Davide;	Using Deep Generative Models for Representation Learning with Applications to AI Explainability		2021	Politecnico di Torino	x	No
Okolo, Chinasa T; Lin, Hongjin;	Practitioner perspectives on model explainability in ai for social good	Preprint	2023		x	No
Hammer, Alexander; Goettling, Marc; Malberg, Hagen; Linke, Axel; Richter, Sergio; Mangner, Norman; Schmidt, Martin;	Fusion of automatically learned rhythm and morphology features matches diagnostic criteria and enhances AI explainability		2024		x	No
Cecil, Julia; Lermer, Eva; Hudecek, Matthias FC; Sauer, Jan; Gaube, Susanne;	Explainability does not mitigate the negative impact of incorrect AI advice in a personnel selection task	Scientific reports	2024	Nature Publishing Group UK London		Yes
Wulff, Kristin; Finnestrand, Hanne;	Creating meaningful work in the age of AI: explainable AI, explainability, and why it matters to organizational designers	AI & SOCIETY	2024	Springer		Yes
Beger, Jan;	The crucial role of explainability in healthcare AI	European Journal of Radiology	2024	Elsevier		Yes
van Diemen, Timon Raoul;	Explainability of a black-box system in AI-systems				x	No
Jin, Tian; Wang, Yufeng; Ming, Wenlong; Lüdtke, Ingo; Lewis, Adam; Wang, Sheng;	AI-driven design approach for dual active bridge converters with increased explainability	IET Conference Proceedings CP878	2024	IET		Yes
Xu, Alex; Panczyk, Nataly; Radaideh, Majdi I;	Towards Explainable AI in Nuclear: Introducing Ad Hoc Model Explainability					Yes
Klein, Georgia; Kavassalis, Sarah C;	Exploring the Explainability Power of AI in Identifying Causal Factors of Air Pollution: An Ozone in Los Angeles Case-Study	AGU Fall Meeting Abstracts	2023		x	No
Haid, Charlotte; Tran, Gia-phong; Fottner, Johannes;	Explainability in AI-based shift scheduling	Intelligent Human Systems Integration (IHSI 2025): Integrating People and Intelligent Systems	2025	AHFE Open Acces		Yes
Koch, Janin; Fortes, Vitor;	Combining Human-centred Explainability and Explainable AI	CHI 2022-Conference on Human Factors in Computing Systems	2022		x	No
Maglianella, Leandro; Nicoletti, Lorenzo; Giagu, Stefano; Napoli, Christian; Scardapane, Simone;	Convergent Approaches to AI Explainability for HEP Muonic Particles Pattern Recognition	Computing and Software for Big Science	2023	Springer		Yes
Kreuzer, Tim; Papapetrou, Panagiotis; Zdravkovic, Jelena;	AI Explainability Methods in Digital Twins: A Model and a Use Case	International Conference on Enterprise Design, Operations, and Computing	2024	Springer		Yes
Saidov, Arslonbek; Sharipov, Daler;	THE IMPORTANCE OF EXPLAINABILITY IN MEDICAL AI AND PREDICTOR REDUCTION IN CARDIOVASCULAR RISK ASSESSMENT	DIGITAL TRANSFORMATION AND ARTIFICIAL INTELLIGENCE	2025		x	No
Saldiran, Emre; Hasanzade, Mehmet; Inalhan, Gokhan; Tsourdos, Antonios;	Explainability of AI-driven air combat agent	2023 IEEE Conference on Artificial Intelligence (CAI)	2023	IEEE		Yes
Allison, Jesse; Farrar, Drew; Nash, Treya; Román, Carlos; Weeks, Morgan; Ju, Fiona Xue;	Play Me Something Icy: Practical Challenges, Explainability and the Semantic Gap in Generative AI Music	arXiv preprint arXiv:2408.07224	2024			Yes
Liapis, Antonios; Zhu, Jichen;	The Need for Explainability in AI-Based Creativity Support Tools	Proceedings of the Human Centered AI workshop at NeurIPS 2022	2022		x	No
Waefler, Toni; Schmid, Ute;	Explainability is not enough: requirements for human-AI-partnership in complex socio-technical systems		2020	Otto-Friedrich-Universität		Yes
Boyaci, Tamer; de Véricourt, Francis; Canyakmaz, Caner;	Beyond the Black Box: Unraveling the Role of Explainability in Human-AI Collaboration	Available at SSRN 5045415	2024		x	No
Ellis, Sam; Gomes, Sandra; Trumble, Matthew; Halling-Brown, Mark D; Young, Kenneth C; Warren, Lucy M;	Explainability of an AI-based breast cancer risk prediction tool	17th International Workshop on Breast Imaging (IWBI 2024)	2024	SPIE	x	No
Aiongi, Francesco;	Towards multi-granular explainable AI: increasing the explainability level for deepfake detectors		2024		x	No
Kelley, Patrick Gage; Woodruff, Allison;	Advancing explainability through AI literacy and design resources	Interactions	2023	ACM New York, NY, USA		Yes

Authors	Title	Source	Year	Publisher	Mark	Decision
You, Yuzhe; Zhao, Jian;	Gamifying XAI: Enhancing AI Explainability for Non-technical Users through LLM-Powered Narrative Gamifications	arXiv preprint arXiv:2410.04035	2024		x	No
Rezazade Mehrizi, Mohammad H; Mol, Ferdinand; Peter, Marcel; Ranschaert, Erik; Dos Santos, Daniel Pinto; Shahidi, Ramin; Fatehi, Mansoor; Dratsch, Thomas;	The impact of AI suggestions on radiologists' decisions: a pilot study of explainability and attitudinal priming interventions in mammography examination	Scientific Reports	2023	Nature Publishing Group UK London		Yes
Molé, Michele;	The quest for effective fundamental labour rights in the European post-pandemic scenario: Introducing principles of explainability and understanding for surveillance through AI algorithms and IoT devices.	19th International Conference in Commemoration of Marco Biagi: Work Beyond the Pandemic. Towards a Human-Centred Recovery	2022		x	No
Vuarin, Louis; Steyer, Véronique;	Marketing faced with the organizational challenge of AI explainability	Décisions Marketing	2023	EMS Editions	x	No
Karaoğlu, Mustafa Sina;	Explainability of Generative AI for Architecture: Investigating the Alignment With Design Intentions		2023	Middle East Technical University (Turkey)	x	No
Bobek, Szymon; Krakowska, Monika; Mozolewski, Maciej; Rak, Dorota; Zych, Magdalena; Wójcik, Magdalena; Nalepa, Grzegorz J;	User-centric evaluation of explainability of AI with and for humans: a comprehensive empirical study	arXiv preprint arXiv:2410.15952	2024		x	No
Fleiß, Jürgen; Bäck, Elisabeth; Thalmann, Stefan;	Explainability and the intention to use AI-based conversational agents	Proceedings of the First International Workshop on Explainable and Interpretable Machine Learning (XI-ML 2020)	2020			Yes
Michael, Loizos;	Explainability and the fourth AI revolution	Handbook of Research on Artificial Intelligence, Innovation and Entrepreneurship	2023	Edward Elgar Publishing	x	No
Carvalho, Claudio Alexandre S;	Ethical challenges of AI-based psychotherapy: the case of explainability	Scenari: rivista semestrale di filosofia contemporanea: 17, 2, 2022	2022	Mimesis Edizioni	x	No
Cheng, Hao-Fei;	Advancing Explainability and Fairness in AI with Human-Algorithm Collaborations		2022	University of Minnesota		Yes
Al Noman, Abdullah; Arif, Abu Shamim Mohammad;	Brain Tumor Recognition from MRI Using Deep Learning with Data Balancing Methods and Its Explainability with AI	International Conference on Image Processing and Capsule Networks	2023	Springer		Yes
Abdelhalim, Esraa;	The Willingness to Collaborate with Artificial Intelligence (AI) in the Workplace: The Role of AI Autonomy and Explainability		2023		x	No
Holm, Thomas Ploug & Søren;	The Limits of Explainability in Health AI-Why Current Concepts of AI Explainability Cannot Accommodate Patient Interests	Journal of Applied Ethics and Philosophy	2025			Yes
Yang, Na-Eun; Han, Eun-Bin; Choi, Jun-Ho;	Explainability of Image Generative AI for Novice and Expert Users: A Comparative Study of Static and Dynamic Explanations	디지털콘텐츠학회논문지 (J. DCS)	2024		x	No
Makhlouf, Karima;	Advancing Ethical and Responsible AI: Exploring Fairness, Privacy, and Explainability through Causal Perspectives		2024	École polytechnique	x	No
Garrett, Zack;	Explaining Go: Challenges in Achieving Explainability in AI Go Programs		2023		x	No
Mehrizi, Mohammad H Rezazade; Mol, Ferdinand; Peter, Marcel; Ranschaert, Erik; Dos Santos, Daniel Pinto; Shahidi, Ramin; Fatehi, Mansoor; Dratsch, Thomas;	How are radiologists' decisions impacted by AI suggestions? Moderating effect of explainability inputs and attitudinal priming in examining mammograms		2022			Yes
Pinna, Emanuele;	Applicazioni dell'Intelligenza Artificiale Generativa tramite Latent Diffusion per l'Explainability nella Geolocalizzazione= Applications of Generative AI via Latent Diffusion for Explainability in Geolocalization		2024	Politecnico di Torino	x	No