1. Record Nr. UNINA9910743694703321

Autore	Calvaresi Davide
Titolo	Explainable and Transparent AI and Multi-Agent Systems : 5th International Workshop, EXTRAAMAS 2023, London, UK, May 29, 2023, Revised Selected Papers / / edited by Davide Calvaresi, Amro Najjar, Andrea Omicini, Reyhan Aydogan, Rachele Carli, Giovanni Ciatto, Yazan Mualla, Kary Främling
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2023
ISBN	9783031408786 3031408780
Edizione	[1st ed. 2023.]
Descrizione fisica	1 online resource (289 pages)
Collana	Lecture Notes in Artificial Intelligence, , 2945-9141 ; ; 14127
Altri autori (Persone)	NajjarAmro OmiciniAndrea AydoganReyhan CarliRachele CiattoGiovanni MuallaYazan ämlingKary
Disciplina	006.30285436
Soggetti	Multiagent systems Machine learning Compilers (Computer programs) Natural language processing (Computer science) Computer science Computers, Special purpose Multiagent Systems Machine Learning Compilers and Interpreters Natural Language Processing (NLP) Theory of Computation Special Purpose and Application-Based Systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia

Nota di contenuto	Explainable Agents and multi-agent systems -- Mining and Validating Belief-based Agent Explanations -- Evaluating a mechanism for explaining BDI agent behaviour -- A General-Purpose Protocol for Multi-Agent based Explanations -- Dialogue Explanations for Rules-based AI Systems -- Estimating Causal Responsibility for Explaining Autonomous Behavior -- Explainable Machine Learning -- The Quarrel of Local Post-hoc Explainers for Moral Values Classification in Natural Language Processing -- Bottom-Up and Top-Down Workflows for Hypercube- and Clustering-based Knowledge Extractors -- Imperative Action Masking for Safe Exploration in Reinforcement Learning -- Reinforcement Learning in Cyclic Environmental Change for Non-Communicative Agents: A Theoretical Approach -- Inherently Interpretable Deep Reinforcement Learning through Online Mimicking -- Counterfactual, Contrastive, and Hierarchical Explanations with Contextual Importance and Utility -- Cross-domain applied XAI -- Explanation Generation via Decompositional Rules Extraction for Head and Neck Cancer Classification -- Metrics for Evaluating Explainable Recommender Systems -- Leveraging Imperfect Explanations for Plan Recognition Problems -- Reinterpreting Vulnerability to Tackle Deception in Principles-Based XAI for Human-Computer Interaction -- Using Cognitive Models and Wearables to Diagnose and Predict Dementia Patient Behaviour.
Sommario/riassunto	This volume LNCS 14127 constitutes the refereed proceedings of the 5th International Workshop, EXTRAAMAS 2023, held in London, UK, in May 2023. The 15 full papers presented together with 1 short paper were carefully reviewed and selected from 26 submissions. The workshop focuses on Explainable Agents and multi-agent systems; Explainable Machine Learning; and Cross-domain applied XAI. .