

Responsible AI

From Research to Practice

Jeriek Van den Abeele

Research Scientist
Telenor Research & Innovation

Table 1: Overview of EU Legislations in the Digital Sector

Applicable law	Published in the Official Journal of the European Union.
In negotiation	Proposal by the European Commission entered the legislative process.
Planned initiative	Mentioned by the European Commission as potential legislative initiative.

Research & Innovation	Industrial Policy	Connectivity	Data & Privacy	IPR	Cybersecurity	Law Enforcement	Trust & Safety	E-commerce & Consumer Protection	Competition	Media	Finance
Digital Europe Programme Regulation, (EU) 2021/684	Recovery and Resilience Facility Regulation, (EU) 2021/241	Frequency Bands Directive, (EEC) 1987/372	ePrivacy Directive, (EC) 2002/58, 2017/0003(COD)	Database Directive, (EC) 1996/9	Regulation for a Cybersecurity Act, (EU) 2019/881, 2023/0108(COD)	Law Enforcement Directive, (EU) 2016/680	Product Liability Directive (PLD), (EEC) 1985/374, 2022/0302(COD)	Unfair Contract Terms Directive (UCTD), (EEC) 1993/13	EC Merger regulation, (EC) 2004/399	Satellite and Cable Directive, (EEC) 1993/83	Common VAT system, (EC) 2006/112, 2022/0407(CNS)
Horizon Europe Regulation, (EU) 2021/695, (EU) 2021/764	InvestEU Programme Regulation, (EU) 2021/523	Radio Spectrum Decision, (EC) 2002/678	European Statistics, (EC) 2009/223, 2023/0237(COD)	Community Design Directive, (EC) 2002/6, 2022/0391(COD)	Regulation to establish a European Cybersecurity Competence Centre, (EU) 2021/887	Directive on combating fraud and counterfeiting of non-cash means of payment, (EU) 2019/713	Toys Regulation, (EC) 2009/48, 2023/0290(COD)	Price Indication Directive, (EC) 1998/6	Technology Transfer Block Exemption, (EU) 2014/316	Information Society Directive, (EC) 2001/729	Administrative cooperation in the field of taxation, (EU) 2011/16
Regulation on a pilot regime for distributed ledger technology, (EU) 2022/658	Connecting Europe Facility Regulation, (EU) 2021/1153	Open Internet Access Regulation, (EU) 2015/2120	General Data Protection Regulation (GDPR), (EU) 2016/679	Enforcement Directive (IPR), (EC) 2004/48	NIS 2 Directive, (EU) 2022/2555	Regulation on interoperability between EU information systems in the field of borders and visa, (EU) 2019/817	European Standardization Regulation, (EU) 2012/1025	E-commerce Directive, (EC) 2000/31	Company Law Directive, (EU) 2017/1132, 2023/0089(COD)	Audio-visual Media Services Directive (AVMSD), (EU) 2010/13	Payment Service Directive 2 (PSD2), (EU) 2015/2366, 2023/0239(COD)
	Regulation on High Performance Computing Joint Undertaking, (EU) 2021/1173, 2024/0016(CNS)	European Electronic Communications Code Directive (EECC), (EU) 2018/1972	Regulation to protect personal data processed by EU institutions, bodies, offices and agencies, (EU) 2018/1726	Directive on the protection of trade secrets, (EU) 2016/943	Cybersecurity Regulation, (EU) 2023/2841	Regulation on terrorist content online, (EU) 2021/784	Radio Equipment Directive (RED), (EU) 2014/53	Unfair Commercial Practices Directive (UCPD), (EC) 2005/29	Market Surveillance Regulation, (EU) 2019/1020	Portability Regulation, (EU) 2017/1128	Digital Operational Resilience Act (DORA Regulation), (EU) 2022/2554
	Regulation on Joint Undertakings under Horizon Europe, (EU) 2021/2085, 2022/0033(NLE)	.eu top-level domain Regulation, (EU) 2019/517	Regulation on the free flow of non-personal data, (EU) 2018/1807	Design Directive, 2022/0392(COD)	Information Security Regulation, 2022/0084(COD)	Temporary CSAM Regulation, (EU) 2021/1232, 2022/0135(COD)	eIDAS Regulation (European Digital Identity Framework), (EU) 2014/910	Directive on Consumer Rights (CRD), (EU) 2011/83	P2B Regulation, (EU) 2019/1160	Satellite and Cable II Directive, (EU) 2019/789	Crypto-assets Regulation (MiCA), (EU) 2023/1114
	Decision on a path to the Digital Decade, (EU) 2022/2481	Roaming Regulation, (EU) 2022/612	Open Data Directive (PSI), (EU) 2019/1024	Compulsory licensing of patents, 2023/0129(COD)	Cyber Resilience Act, 2022/0272(COD)	E-evidence Regulation, (EU) 2023/1543	Regulation for a Single Digital Gateway, (EU) 2018/1724	e-Invoicing Directive, (EU) 2014/55	Single Market Programme, (EU) 2021/690	Copyright Directive, (EU) 2019/790	Financial Data Access Regulation, 2023/0205(COD)
	European Chips Act (Regulation), (EU) 2023/1781	Union Secure Connectivity Programme, (EU) 2023/588	Data Governance Act (DGA Regulation), (EU) 2022/888	Standard essential patents, 2023/0133(COD)	Cyber Solidarity Act (Regulation), 2023/0102(COD)	Digitalisation of cross-border judicial cooperation, (EU) 2023/2844	General Product Safety Regulation, (EU) 2023/988	Regulation on cooperation for the enforcement of consumer protection laws, (EU) 2017/2394	Vertical Block Exemption Regulation (VBER), (EU) 2022/720	European Media Freedom Act, (EU) 2024/1083	Payment Services Regulation, 2023/0210(COD)
	Establishing the Strategic Technologies for Europe Platform (STEP), (EU) 2024/795	Gigabit Infrastructure Act, (EU) 2024/1309	European Data Act (Regulation), (EU) 2023/2854			Directive on combating violence against women, 2022/0066(COD)	Machinery Regulation, (EU) 2023/1230	Geo-Blocking Regulation, (EU) 2018/302	Digital Market Act (DMA Regulation), (EU) 2022/1925	Remuneration of musicians from third countries for recorded music played in the EU	Digital euro, 2023/0212(COD)
	European critical raw materials act (Regulation), (EU) 2024/1282	New radio spectrum policy programme (RSPP 2.0)	Interoperable Europe Act, (EU) 2024/993			Directive for combating sexual abuse and child sexual abuse material, 2024/0035(COD)	A) Act (Regulation), 2021/0106(COD)	Digital content Directive, (EU) 2019/770	Regulation on distortive foreign subsidies, (EU) 2022/2560		Regulation on combating late payment, 2023/0323(COD)
	Net Zero Industry Act, 2023/0081(COD)	Digital Networks Act	Regulation on data collection for short-term rental, (EU) 2024/1028			Digitalization of travel documents	Eco-design Regulation, 2022/0095(COD)	Directive on certain aspects concerning contracts for the sale of goods, (EU) 2019/771	Horizontal Block Exemption Regulations (HBER), (EU) 2023/1066, (EU) 2023/1067		
	EU Space Law		European Health Data Space (Regulation), 2022/0140(COD)				AI Liability Directive, 2022/0303(COD)	Digital Services Act (DSA Regulation), (EU) 2022/2065	Platform Work Directive, 2021/0414(COD)		
			Harmonisation of GDPR enforcement procedures, 2023/0202(COD)					Political Advertising Regulation, (EU) 2024/900	Single Market Emergency Instrument (SMEI), 2022/0278(COD)		
			Access to vehicle data, functions and resources					Right to repair Directive, 2023/0083(COD)			
			GreenData4all					Consumer protection, strengthened enforcement, cooperation			

Brussels, 29 November 2021
(OR. en)

14278/21

LIMITE

TELECOM 430
JAI 1288
COPEN 412
CYBER 307
DATAPROTECT 267
EJUSTICE 103
COSI 236
IXIM 262
ENFOPOL 465
FREMP 272
RELEX 1012
MI 879
COMPET 860
CODEC 1530

Interinstitutional File:
2021/0106(COD)

NOTE

From:	Presidency
To:	Delegations
No. Cion doc.:	8115/20
Subject:	Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text



GSMA-ETNO position paper on European Commission proposal for an Artificial Intelligence Act

GSMA and ETNO welcome the initiative of the European Commission to propose a Regulation on Artificial Intelligence, and notably the central role of the risk-based approach, which keeps citizens at the centre and is based on a robust ethical framework. Nevertheless, the risk-based approach requires more precision in order to achieve the intended results. It is encouraging to see that the legal focus is put on the use of AI and its impacts to society, rather than the technology itself.

AI applications will be a key driver of innovation for the European telecoms industry, notably as part of the shift towards 5G, virtualised networks. 5G and fibre connectivity will accelerate the digitisation of services and industrial processes, enabling the rapid expansion of the Internet of Things (IoT). The massive amounts of data generated by IoT connections and devices will open up new growth opportunities for data analytics and AI services in Europe. High-class, secure connectivity will then drive IoT, and IoT will in turn fuel European AI. Together, they can form a truly powerful virtuous circle that our industry is committed to nurturing. Digital network

AI Opportunities for the Telecoms Industry

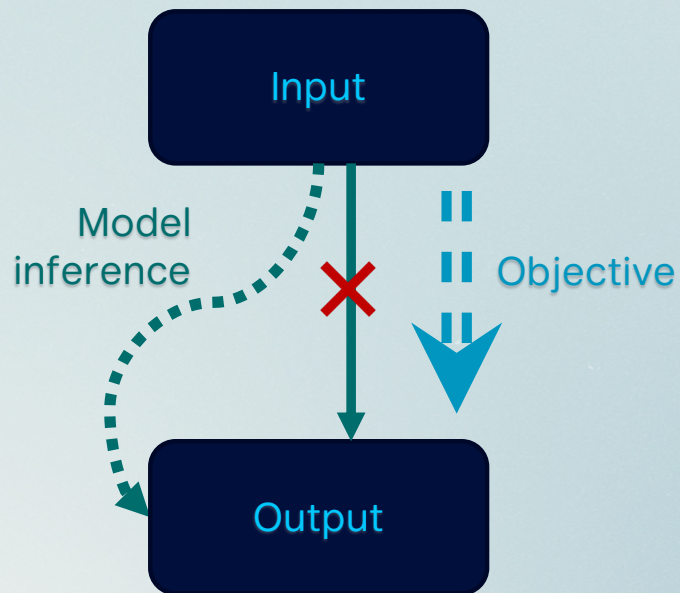
Core Business optimisation

Personalised and improved customer interaction

AI-Driven Mobile Data Insights

On 12 October 2021, the Presidency requested the delegations in the **WP TELECOM** to provide written comments and drafting suggestions on the first 29 articles of the proposed AIA by 26 October, with a view to start working on the first, partial compromise text of the proposal. Following this request, 18 Member States provided their written contributions.

OECD definition of AI (Nov 2023 update)



Unlike human-defined rule-based systems, AI systems have a **degree of autonomy** regarding how to accomplish their task.

An **AI system** is a **machine-based** system that,

- for explicit or implicit objectives, **infers**,
 - from the **input** it receives,
- how to generate **outputs** such as
 - predictions,
 - content,
 - recommendations, or
 - decisions

that can influence physical or virtual **environments**.

Different AI systems vary in their levels of **autonomy** and **adaptiveness** after deployment.

Other key issues: general-purpose AI, high-risk classification, open source, responsibilities along value chain



Researchers have a critical role to make AI more responsible

- By weighing in on regulation and policy
- By guiding practices at our organisations
- By advancing AI research in more responsible directions

Researchers have a critical role to make AI more responsible

- By weighing in on regulation and policy
- By guiding practices at our organisations
- By advancing AI research in more responsible directions

What could go wrong?



Biases in automated CV screening



Discrimination in personalised offers



Unfair call centre performance management



Chatbot leaking sensitive information

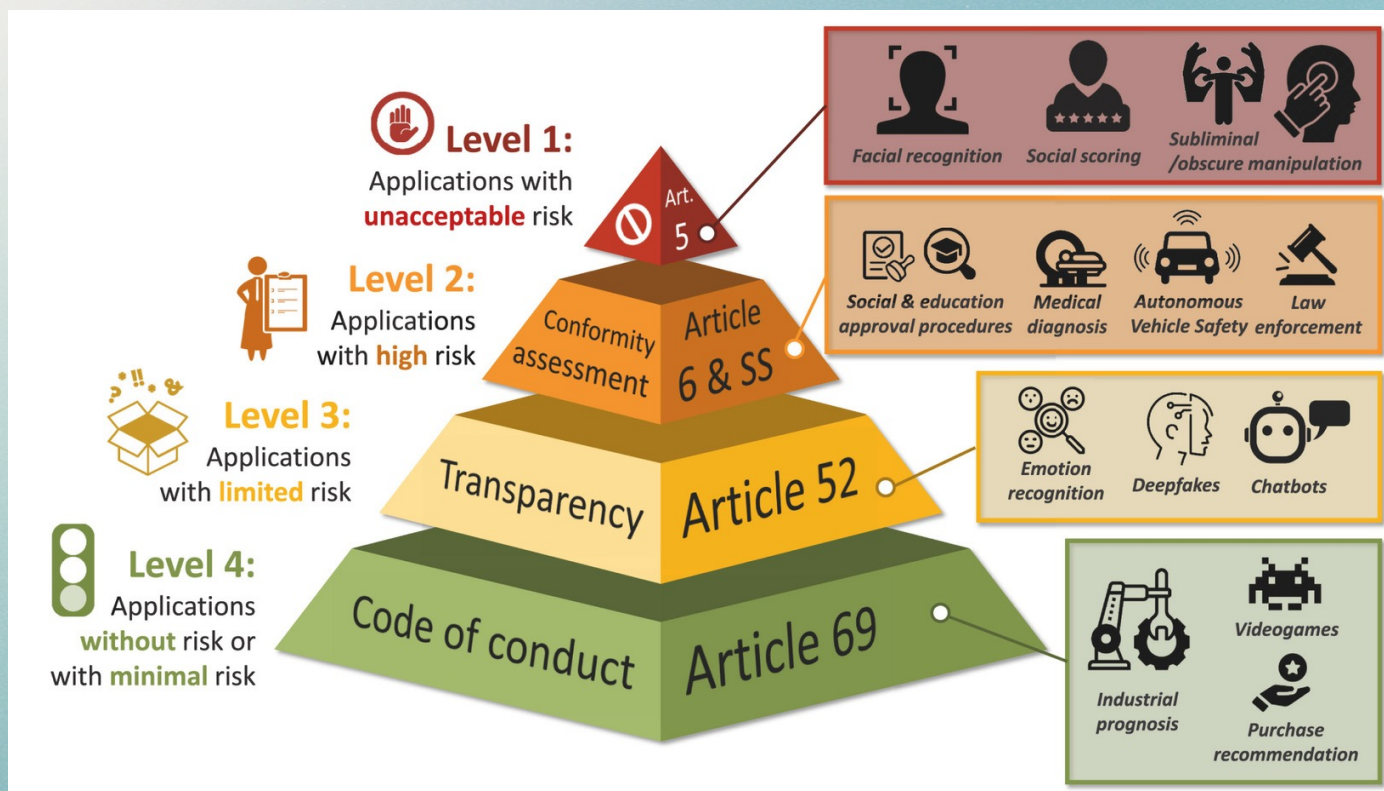


Privacy-violating tracking of employee wellbeing



Automated vulnerability discovery and exploitation

The AI Act will not magically save us ...





Human values,
rights & privacy



Fairness & non-
discrimination



Data governance



Safety & security



Transparency &
explainability



Human oversight
& control



Responsible AI
principles



Playbooks
& tools



Raising standards
with partners



Safeguarding
customers

Researchers have a critical role to make AI more responsible

- By weighing in on regulation and policy
- By guiding practices at our organisations
- By advancing AI research in more responsible directions

Green AI

Human-
centric AI

ENFIELD

Trustworthy AI

Adaptive AI

Healthcare

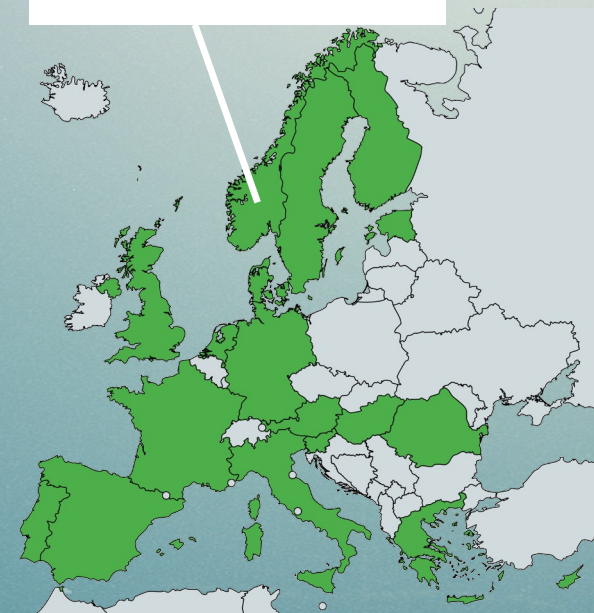
Energy

Manufacturing

Space

Join us through open calls: enfield-project.eu

Horizon Europe project
30 partners in 18 countries





ENFIELD projects at Telenor

Status
November 2024

Green AI

Review of environmental benefits and harms of AI

- Tertiary review on sustainability throughout AI lifecycle
- *With consortium partners*

Framework for Green AI practices (Open Call TES1)

- Development of AI sustainability metrics framework for ASR
- *María Ulan (RISE), with SINTEF*

Green generative language models [Open Call TES3?]

- Looking into energy efficiency of fine-tuning, prompting and in-context learning techniques

Human-centric AI

User trust in AI

- Broad survey study of Telenor employees' current work practices, perceptions and attitudes towards AI
- Focus on understanding trust drivers, how trust levels impact the adoption and use of AI, and the potential of explainable AI

User perspectives on XAI [Open Call TES2]

- Empirical investigation of trust calibration and user preferences for explainability features in impactful AI-assisted decision-making tasks in a corporate environment with technical workers

Trustworthy AI

Security and robustness of AI systems [Open Call TES2]

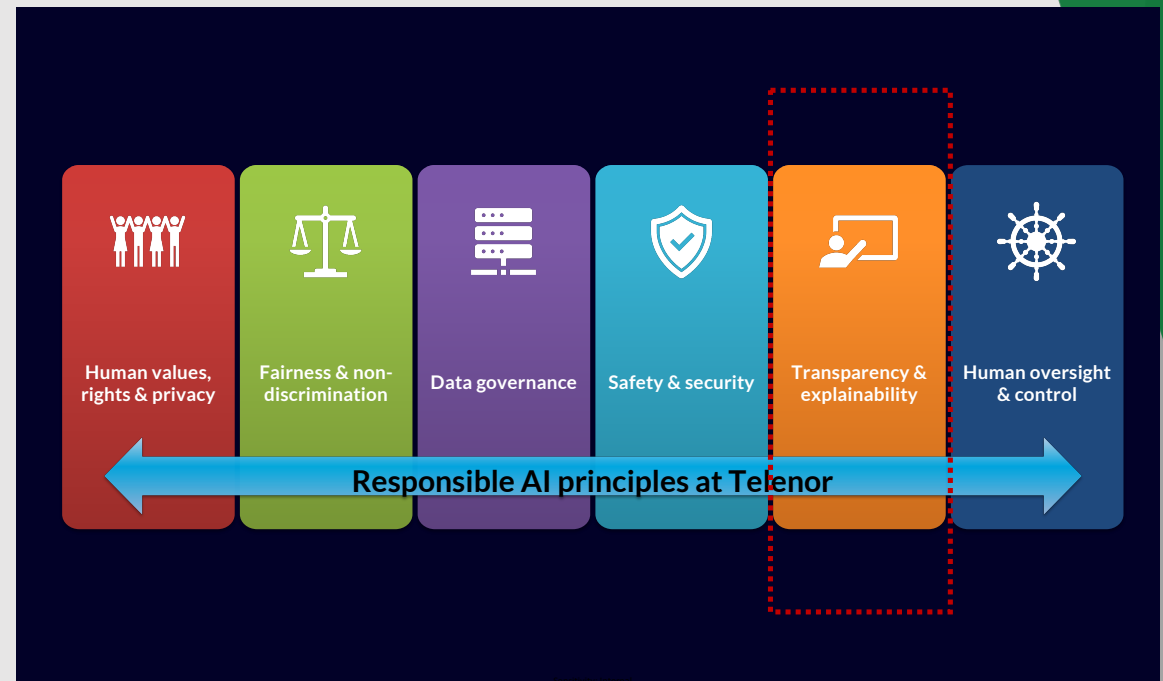
- Interest in improving LLM reliability and security, particularly defences against adversarial attacks
- Also exploring uncertainty quantification techniques (including conformal prediction) to make timeseries modelling more robust

ENFIELD Human-centric AI at Telenor

Motivations

Telenor launched a broad **AI-first strategy** in 2023.

There is a **lack of concrete empirical studies** focusing on the specific trust dynamics in AI adoption and the impact of explainable AI on impactful tasks in corporate environments.



User trust in AI

- Broad survey study of Telenor employees' current work practices, perceptions and attitudes towards AI
- Focus on understanding trust drivers, how trust levels impact the adoption and use of AI, and the potential of explainable AI

Prior research: Trust in AI-enabled systems

Paper: *A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective* (Bach et al., 2024)

- Synthesises 23 empirical studies to provide insights into future AI design strategies

Main findings:

- The **definition of user trust** needs to be adapted to the specific context
 - Often building on ABI model of organisational trust: key trustee characteristics are **ability** (knowledge and skills), **benevolence** (concern for others), and **integrity** (ethical/moral principles)
 - *“The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party.”* (Mayer et al., 1995)
- **Key factors** influencing user trust:
 - **Socio-ethical considerations** (e.g., fairness, data protection, accountability)
 - **Technical and design features** (e.g., transparency, explainability, reliability)
 - **User characteristics** (e.g., prior experience, cognitive load, involvement in development)

User perspectives on XAI [Open Call TES2]

- Empirical investigation of trust calibration and user preferences for explainability features in impactful AI-assisted decision-making tasks in a corporate environment with technical workers

Prior research: Impact of explainability

AI in decision making

- Ubiquity in tasks requiring accuracy; yet, limited by stochastic nature and biases
- Enhances human decisions in high-stakes fields (e.g., healthcare, critical infrastructure)

Trust and understanding: human-centric AI

- Essential to know AI system's limitations and mentally model error boundaries (*Bansal et al., 2019*)
- Trust in AI needs calibration, e.g., supported by transparency and explanations (*Ribeiro et al., 2016*)

Empirical studies and trust calibration

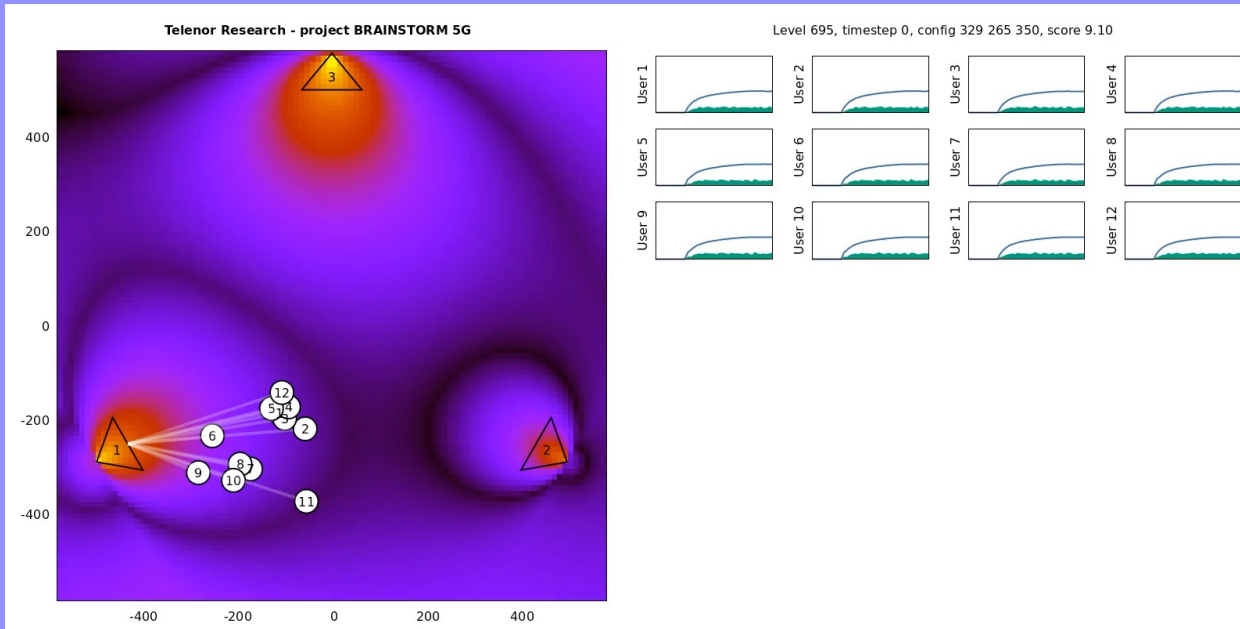
- Case-specific information like confidence scores and local explanations can theoretically support better decision outcomes, but require empirical testing to validate effectiveness
- Studies reveal mixed effects of transparency/explainability on trust in AI, depending strongly on the setup
 - Trust calibration is insufficient to improve accuracy of AI-assisted decisions, it also depends on complementarity of the user's knowledge (*Zhang et al., 2020*)
 - What happens in a scenario where domain experts need to make impactful decisions?

Reinforcement learning for network energy optimisation

MARIO
007100

WORLD
1-1

TIME
272



Morphing between situations demonstrates
dynamic capabilities of the agent

Reinforcement learning for network energy optimisation

Telenor and Ericsson join forces to pioneer the usage of AI and Machine Learning Research for a Sustainable and Smarter Future

In a step towards advancing the realm of artificial intelligence (AI) and machine learning (ML) technologies within the telecommunications industry, Telenor and Ericsson (NASDAQ: ERIC) have signed a Memorandum of Understanding (MoU) for a three-year partnership that aims to explore, develop, and test cutting-edge AI/ML solutions towards enhancing energy efficiency without compromising on the quality of connectivity in mobile networks.

PRESS RELEASE | NOV 30, 2023

MARIO
007100

WORLD
1-1

TIME
272



Researchers have a critical role to make AI more responsible!

- By weighing in on regulation and policy
- By guiding practices at our organisations
- By advancing AI research in more responsible directions