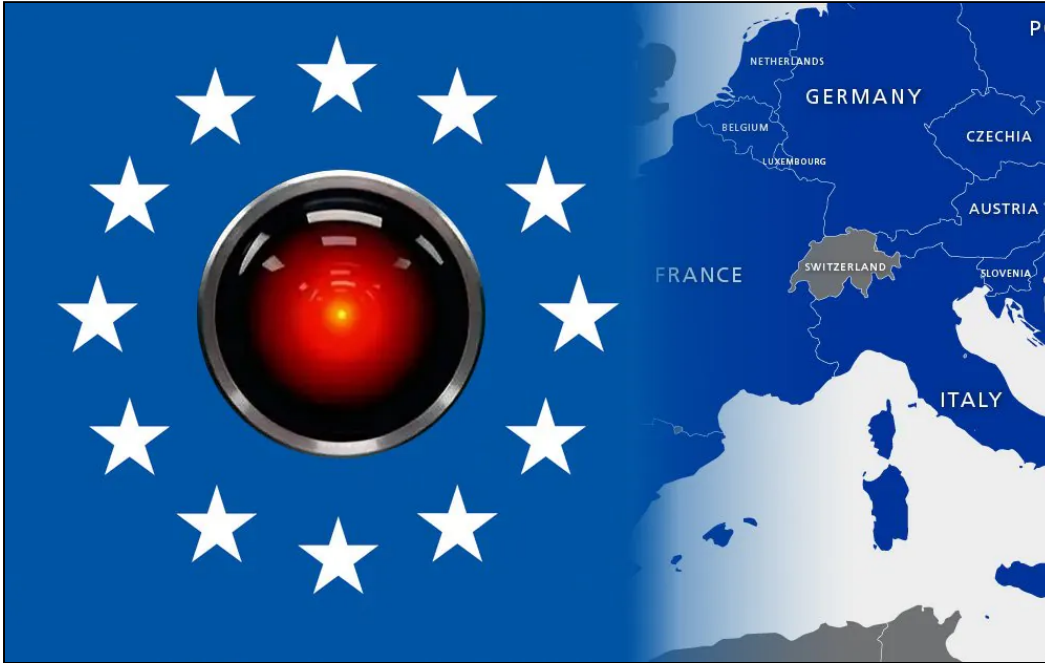


The Failings of the Draft EU Artificial Intelligence Act

Originally published at <https://www.unite.ai/the-failings-of-the-draft-eu-artificial-intelligence-act/>



A new legal critique of the European Union's draft 'AI Act' levels a wide array of criticisms at the proposed regulations [released in April](#), concluding that much of the document is 'stitched together' from scarcely applicable 1980s consumer regulation; that it actually promotes a deregulated AI environment in Europe, rather than bringing the sector under coherent regulation; and – among a slew of other criticisms – that the proposals map out a future regulatory AI framework that has 'little sense and impact'.

Entitled *Demystifying the Draft EU Artificial Intelligence Act*, the [pre-print](#) is a collaboration between researchers from UCL London and Radboud University in Nijmegen.

The paper adds to a growing body of negative opinion about the proposed implementation (rather than the much-admired intent) of a regulatory AI framework, including the contention [in April](#) of one of the draft regulation's own contributors that the proposed guidelines are 'lukewarm, short-sighted and deliberately vague', which characterized the European Commission's document as a proponent of 'fake ethics'.

Manipulative AI Systems

The new paper contends that the AI Act's proposed restrictions on 'manipulative systems' is hamstrung by a vague and even contradictory definition of 'harm', commenting that '[a] cynic might feel the Commission is more interested in prohibitions' rhetorical value than practical effect'.

The [draft regulations](#) outline two putative prohibited practices:

(a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological

harm;

(b) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;

The researchers argue that these restrictions do not address whether or not an AI provider's services or software is successful in furthering their own aims, but only if the end user suffers 'harm' in the process. They add that the draft's definition of harm is fatally limited to individual users, rather than the kind of collective or societal harm that can be reasonably inferred from a number of AI-based controversies of recent years, such as the Cambridge Analytica debacle.

The paper observes that 'In real life, harm can accumulate without a single event tripping a threshold of seriousness, leaving it difficult to prove'.

Harmful AI Systems Allowed, but Not for EU Consumption

The AI Act proposes to implement a ban on 'real-time' biometric systems in public spaces by law enforcement. Though some public skepticism has been aimed at the exceptions the proposals make for anti-terrorism, child trafficking and the pursuit of a European arrest warrant, the researchers note also that nothing would prevent suppliers from selling contravening biometric systems to oppressive regimes.

The paper observes that this is already historical practice, as revealed in 2020 report from Amnesty international.

It further states that the AI Act's specification of 'real time' biometric system is arbitrary, and excludes off-line analytical systems, such as later processing of video footage from protest events.

Additionally, it's noted that the proposals offer no mechanism to restrict biometric systems that are *not* related to law enforcement, which instead are lazily deferred to the GDPR; and that the GDPR itself 'places a requirement of high-quality, individual consent for each scanned person which is effectively impossible to fulfil'.

The wording of this section of the AI Act also comes in for criticism from the researchers. The draft stipulates that pre-authorization will be required for the deployment of biometric systems for competent authorities' 'individual use' of such systems – but does not clarify what 'individual use' means in this context. The paper notes that controversial warrants can be *thematic*, and relate to broad organizations, purposes and places.

Further, the draft regulations do not stipulate a transparency mechanism for the number and type of authorizations issued, making public scrutiny problematic.

Outsourcing Regulation to 'Harmonized Standards'

The research states that the most important entities in the AI Act are actually not mentioned even once in the draft regulations: CEN (European Committee for Standardisation) and CENELEC (European Committee for Electrotechnical Standardisation) – two of three European Standardisation Organizations

(ESOs) that the European Commission can mandate to formulate harmonized standards, which in many cases would remain the governing regulatory frameworks for certain types of AI services and deployments.

This effectively means that AI producers can choose to follow the standards of what are in effect competing rather than complementary regulations, instead of meeting the essential requirements outlined in the AI Act. This allows providers to more loosely interpret the proposed regulations when they come into force in 2024-5.

The paper's researchers also opine that intervening years of industrial lobbying among standards bodies are likely to redefine these 'essential standards' considerably, and suggest that 'ideal' regulations should start out at a higher ethical level and legislative clarity, if only to account for this inevitable process of attrition.

Legitimizing the Fallacy of Emotion Recognition Systems

The AI Act features provision against the deployment of emotion recognition and categorization systems – frameworks that may not necessarily identify an individual, but either claim to understand what they are feeling or to be able to categorize them in terms of gender, ethnicity, and various other economic and social signifiers.

The researchers argue that this clause is pointless, since GDPR already obliges the purveyors of such systems to provide users with clear information about the use of such systems, so that users may opt out (which may involve not using an online service, or not entering an area where such systems are announced to exist).

More importantly, the paper claims that this clause legitimizes a [debunked](#) technology, and goes on to characterise [FACS](#)-style emotion recognition systems in the light of the shameful history of phrenology and other near-Shamanistic approaches to social categorization from the early industrial age.

'Those claiming to detect emotion use oversimplified, questionable taxonomies; incorrectly assume universality across cultures and contexts; and risk '[taking] us back to the phrenological past' of analysing character traits [from facial structures](#). The Act's provisions on emotion recognition and biometric categorisation seem insufficient to mitigate the risks.'

A Too-Modest Proposal

Besides these, the researchers address other perceived shortcomings in the AI Act in regard to regulation of deepfakes, a lack of oversight for carbon emissions of AI systems, duplication of regulatory oversight with other frameworks, and inadequate definition of prosecutable legal entities.

They urge legislators and civil activists to take action to redress the problems identified, and further note than even their extensive deconstruction of the draft regulations has had to omit many other areas of concern, due to lack of space.

Nonetheless, the paper applauds the Act's vanguard attempt to introduce a system of horizontal regulation of AI systems, citing its many 'sensible elements', such as creating a hierarchy of risk assessment levels, having the commitment to introduce prohibitions, and proposing a public database of systems to which purveyors would need to contribute in order to gain European legitimacy, though noting the legal quandaries that this later requirement is likely to bring up.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More