EJLT European Journal of Law and Technology

How Distrust is Driving Artificial Intelligence Regulation in the European Union

Clement Guitton, Aurelia Tamò-Larrieux and Simon Mayer*

Abstract

The emergence of new technologies often brings with it a complex interplay between their acceptance by society, gauging their risks, and whether it is warranted for the state to be involved – typically via new or amended regulation. However, what drives regulators and decision-makers to even consider the question of whether there is a need for involvement has remained under-studied. In this article, we propose viewing regulation as a process with five distinct phases: laissez-faire, awareness, politicisation, regulation and cool-off. A critical phase is the transition between awareness and politicisation, as the latter commonly leads to regulatory action. We look at the emergence of regulation for artificial intelligence, aviation, genetically modified organisms, disinformation and retail self-checkouts to show that there is a correlation between distrust and politicisation. We further show the probable causal link specifically for regulating artificial intelligence in the EU, and derive possible policy implications from this conclusion.

Keywords: regulation, politicisation, agenda-setting, distrust, artificial intelligence.

* Clement Guitton and Simon Mayer are affiliated with the University of St Gallen; Aurelia Tamò-Larrieux is affiliated with the University of Lausanne.

1. Introduction

Legal scholars have long tried to explain what drives the regulation of new technologies.[1] Most of them have focused on answering the question: why regulate? We ask a slightly different question and examine what, upon the emergence of a new technology, drives regulators and decision-makers to even consider the need for regulation. We hypothesise that a key driver might be distrust of the new technology.

Proving this on a general basis would require a comprehensive review of all new technologies. Here, we focus on the special case of one technology currently very much in the headlines: artificial intelligence (AI). There is already a growing literature on over-trusting or under-trusting AI; for example, one possible application of AI is automatic decision-making,[2] and some scholars have advocated public policy workarounds as a solution.[3] Recognising key drivers of regulation would allow us to zero in on key, possibly non-explicit, motivations behind the will to regulate, and could lead to very different policy responses than have otherwise been considered. For instance, would risk assessment really address distrust? Or higher transparency requirements? Aligning the source of distrust with the policy response to avoid any mismatch – assuming that distrust can be rectified – would ensure higher effectiveness.

In other words, the reasons behind regulating are important, as they can influence the policy response;[4] questioning whether to regulate on the right ground is hence relevant. Is distrust a legitimate impetus for considering whether to regulate a new technology? It could be, but we argue that it mostly depends on the type of policy response being considered. There are also notably serious downsides that need to be kept in mind when distrust drives regulation: it may blindside opportunities that the new technology could bring; it can create further fragmentation in society by giving it prominence; and, most problematically, it leads to a focus on making the technology trustworthy. While this may appear a priori a positive development, it is a misguided one, as so many interconnected factors play a role in the complex concept of trust that it is impossible to try to influence any single one of them.[5] Limited resources could be better deployed in other areas.

---

[1] Lyria Bennett Moses, 'Why Have a Theory of Law and Technological Change?' 8 *Minnesota Journal of Law Science & Technology* 589.

[2] Alexander M. Aroyo and others, 'Overtrusting robots: Setting a research agenda to mitigate overtrust in automation' 12 *Journal of Behavioral Robotics* 423; P. Robinette and others, *Overtrust of robots in emergency evacuation scenarios* (IEEE 2016).

[3] Jason W. Burton, Mari-Klara Stein and Tina Blegind Jensen, 'A systematic review of algorithm aversion in augmented decision making' 33 *Journal of Behavioral Decision Making* 220.

[4] Ian Brown and Christopher T. Marsden, *Regulating code: Good governance and better regulation in the information age* (MIT Press 2013); Lawrence Lessig, 'Law regulating code regulating law' 35 *Loyola University Chicago Law Journal*.

[5] Aurelia Tamò-Larrieux and others, 'Regulating for trust: Can law establish trust in artificial intelligence?' (2023) 18 *Regulation & Governance*.

We approach our hypothesis in three steps. First, we designed a conceptual framework on the emergence of regulation for new technologies. A key part of this framework is the 'politicisation' phase, a pre-decision-making stage during which a new technology enters the political arena. Second, we use relevant case studies and discourse analysis to demonstrate a degree of correlation between distrust and the politicisation phase. Thirdly, we identify the probable causal link for how distrust led to politicisation in the case of AI regulation.

2. Theory and Methodology

2.1 Regulation and Technology

Society's quest to regulate emerging technologies is nothing new, with examples of regulating the use (and importantly misuse) of technology by humans dating back to the Hammurabi Code of 1754 BC.[6] In fact, looking back at how society and policymakers have approached the regulation of technology illustrates nicely the (vocal) fears of society towards new technology and the resulting regulatory compromises (taking into account the different interest groups within the ecosystem). We build upon existing research on regulatory models in political sciences and law to propose a regulatory model based around five stages:

1. An intentional or unintentional laissez-faire phase, during which the technology emerges, is deployed, and starts interacting with its ecosystem and society.
2. An awareness phase, during which the technology and its impact become visible and discussed within society, and existing regulatory tools start to be applied to it.
3. A politicisation phase, during which members of the political class realise that the technology poses (new) risks, requiring (new) regulation.
4. A regulation phase, during which debates and lobbying effort gradually peak to shape how the (future) structure of the market for the new technology, as well as society's usage of it, will be enabled and constrained by the regulation.
5. A cool-off phase, during which regulation is applied and enforced. There may be a (constantly changing) gap between the regulation and its enforcement, between the ideal case and how reality is, as is often the case with evolving social norms – this is beyond the scope of this article.

Of particular interest is the transition between the phase of awareness towards the phase of politicisation, where there is a circular relationship between the triggers leading to a (politicised) discourse and the discourse amplifying the perception of those triggers.[7] Because of this focus, a narrow definition of regulation – as opposed

[6] Thibault Schrepel, *Blockchain + Antitrust: The Decentralization Formula* (Edward Elgar Publishing 2022).
[7] Rebecca Crootof and BJ Ard, 'Structuring TechLaw' 34 *Harvard Journal of Law & Technology* 347.

to those detached from the role of the state – is warranted.[8] We hence adopt Selznick's definition of regulation (1985), i.e. a 'sustained and focused control exercised by a public agency over activities that are valued by a community'.[9] Both deliberate state influence and the onus on 'valued by a community' fit well with the etymological meaning of 'political' as the affairs of the polis, the community – hence making this an appropriate definition for the study of politicisation.[10]

The focus on the transition between politicisation and regulation is further warranted as extremely few issues which become politicised return to a state of laissez-faire.[11] Permissive rules may emerge and, within our model, this will still fall within a way of regulating the technology. Or the issue might linger on without decisions being made on them – but even in this case, the issue will remain a political one because it is divisive and a solution through a political decision will be deemed, at least by some, as necessary. The timeframe from politicisation to new regulation – lingering or not – may well be protracted because of the nature of democratic processes, or because of the complexity of the topic, or because a new technical development arises in the course of the regulatory process that alters the needs (as seen with the inclusion of foundation models in the AI Act). Such non-linear processes should not be taken as evidence of regulation not happening. There is, however, one key factor that could trump many others in explaining a position within this process: distrust. We posit that it is at the core of our perception of risks, of the requirement to 'manage' risk, and of our perception that the risk is adequately being 'managed'.

Our regulatory model draws from established research that has investigated how regulatory agendas are set. Notably, Black and Murray (2019) describe how technologies have in the past disrupted regulatory regimes.[12] In their analysis they distinguish among six other phases that occur before a 'concept' becomes regulated: the proof of a theoretical concept; the development of a prototype; the development of a commercial manufacturing and distribution system; the licensing or approval for development; the commercial marketing and exploitation; and reactive regulation and control. While their first three phases are situated within what we call the laissez-faire phase, the fourth phase of licensing (and corresponding fifth phase of commercial exploitation) overlaps with the question of whether the new technology has to be regulated (a question that legal scholars have elaborated upon).[13] Black and

[8] Julia Black and Dimitry Kingsford Smith, 'Critical Reflections on Regulation [Plus a reply by Dimity Kingsford Smith]' 27 *Australasian Journal of Legal Philosophy* 1.
[9] Philip Selznick, 'Focusing Organisational Research on Regulation' in R. Noll (ed), *Regulatory Policy and the Social Sciences* (University of California Press 1985).
[10] Robert Baldwin, Martin Cave and Martin Lodge, *Understanding Regulation: Theory, Strategy, and Practice* (Oxford University Press 2012); A. Heywood, *Political theory: an introduction* (Palgrave Mcmillan 2004).
[11] Crootof and Ard (n 7).
[12] Julia Black and Andrew Murray, 'Regulating AI and Machine Learning: Setting the Regulatory Agenda' 10 *European Journal of Law and Technology* 1.
[13] Crootof and Ard (n 7); Moses (n 1).

Murray's final phase, reactive regulation and control, is, in our regulation-focused model, expanded across phases 3–5.[14]

Other legal scholars have provided insights into the transition between the awareness phase (2) and politicisation (3); however, these analyses are tied to the question of whether new technology has to be regulated (through new forms of regulation, or by the extension of current legal frameworks). This is also why the analysis of legal scholars is less focused on how awareness leads to politicisation, but is better able to answer how technology challenges the regulatory status and thus requires politicisation.

Moses (2007), for instance, identifies four reasons for a necessary regulatory response to technical change:[15] technology enables new forms of conduct; it introduces uncertainty as to how/whether the law applies; the law includes or excludes forms of conducts when it should not; and the rationale for existing rules is no longer valid. Crootof and Ard (2021) break down Moses' legal uncertainty point into three further categories:[16] '(1) application uncertainties, which raise the question of whether and how extant law applies; (2) normative uncertainties, which arise when the law is arguably unable to accomplish its aims; and (3) institutional uncertainties, which exist when there are questions about different regulatory entities' relative authority, competence, and legitimacy to apply and update the law'.[17]

To understand the relationship between technology and regulation better, it is important to understand the political science discourse on politicisation. 'Politicisation' refers to a certain level of interest, engagement and rhetorical strategies by various social actors (not necessarily only elected officials) to open up 'the appearance of an issue as being political'.[18] In light of the many and widely different definitions of the term, it is important to keep in mind how we use it here: as entering the political realm, being framed as such, and gaining prominence through this, echoing the assumption of Hilgartner and Bosk (1988) that public attention (or any individual's, for that matter) be a scarce resource, and therefore how it is being allocated should tell us something of its importance.[19] For elected officials, for instance, this would translate into creating room for the issue on the political agenda, discussing it within political institutions, and proposing bills or other forms of regulatory amendment. Politicisation might even turn into securitisation, tentatively depending on the level of distrust that the technology elicits, the perceived complexity of the technology, and the expected magnitude of (societal and economical) consequences. While the concept of securitisation emerged in

[14] Black and Murray (n 12).
[15] Moses (n 1).
[16] Ibid.
[17] Crootof and Ard (n 7) p.352.
[18] Matthew Wood, 'Politicisation, Depoliticisation and Anti-Politics: Towards a Multilevel Research Agenda' 14 *Political Studies Review* 521, p.4.
[19] Stephen Hilgartner and Charles L. Bosk, 'The Rise and Fall of Social Problems: A Public Arenas Model' 94 *American Journal of Sociology*.

international relations,[20] it has since found a wide range of applications, including on genetically modified organisms (GMOs)[21] and online disinformation.[22] The main point of departure for securitisation has been to understand security not in military terms, but in terms of survival. In other words, it is the two-pronged requirements of presenting the issue as 'posing an existential threat' and of pressing for special measures.

2.2 Methodology: Discourse Analysis of Case Studies

In this article we conduct a discourse analysis to dissect the reasons put forward behind politicisation, especially by those who 'claim to speak or act on behalf of the nation'.[23] We hence focus on 'speech acts' for which we can attribute a certain hypothetical impact in driving a change within the politicisation process.[24] As a consequence of this, and in line with the methodology,[25] our interest lies with how the speech act is carried out and with its consequences, rather than with its object (e.g., GMOs, or AI).

We combine this discourse analysis with a case study approach: we draw on four cases of regulation of technology – early aviation, GMOs, disinformation and AI – to identify correlations by teasing out common and differing factors within these cases, and draw on a fifth case (self-checkouts) in an attempt to show causation. Drawing on other cases for analogies and seeking comparisons with the past has a long tradition when new technology emerges.[26] We make analogies in order to understand the change that is being introduced, and its complexities, but it has also strong limitations, e.g., it introduces conceptual imprecision, emphasises certain aspects over others, and frames the debate in a particular light.[27] More maliciously, certain politicians may leverage the use of old metaphors they know resonate with the public in a particular way.[28]

---

[20] Barry Buzan, Ole Waever and Jaap de Wilde, *Security: A New Framework for Analysis* (Lynne Rienner Publishers 1998).
[21] Shane Markowitz, 'World of "Our" Making: A Socio-Material Constructivist Accounting of the Debate over Genetically Modified Organisms in the European Union' (Central European University 2017).
[22] Nayef Al-Rodhan, 'Post-Truth Politics, the Fifth Estate and the Securitization of Fake News' (June 7) *Global Policy Journal*.
[23] Buzan, Waever and Wilde (n 20) p.41.
[24] Ibid.
[25] Margaret Coulthard, *An introduction to discourse analysis* (Pearson Education Limited 1985).
[26] David J. Betz and Tim Stevens, 'Analogical reasoning and cyber security' 44 *Security Dialogue*.
[27] Crootof and Ard (n 7).
[28] Betz and Stevens (n 26).

3. Drivers of Politicisation in New Technologies

When the first automobiles started to emerge two centuries ago, fears and distrust over these new forms of self-mobility arrived quickly:[29] the British Locomotive Act (or Red Flag Act) came into force in 1865, and mandated that each 'automobile' used on public roads needed to be operated by at least three persons, with one person walking 'not less than sixty yards' in front of the moving car, waving a red flag. The job of the red flag holder was thus to 'warn the riders and drivers of horses of the approach of such locomotives', and 'signal the driver thereof when it shall be necessary to stop, and shall assist horses, and carriages drawn by horses, passing the same'. This is all the more extraordinary when considering that automobiles only started to become common around the 1870s. The use of the term 'locomotives' in the name of the act hints at worries that regulators had, not only about accidents (the Act introduced the first speed limits, but also that heavy, steam-powered engines could damage roads for which they were not built.

The example of the British Locomotive Act illustrates the transition of a new technology in a state of 'laissez-faire' with little attention paid by politicians – here the precursor to cars – to a state where it entered the political arena with a new piece of legislation as an outcome. A similar case is the banning of cars in the Swiss canton of Graubünden from 1900 to 1925 because cars were considered to threaten the traditional horse-drawn carriage.[30] We hypothesise that distrust must have been an (implicit) driver in this transition from new technology to politicisation; hence, distrust as a driving factor within our regulatory model is not entirely new. In order to test out this hypothesis further, case studies are considered below.

3.1 Understanding Distrust

It can be argued that regulators instinctively try to ensure trust in the 'object of regulation' through regulation itself. Put differently, if regulators are faced with new advancements that impact society (what we call 'objects of regulation'; e.g. GMOs, nuclear energy, self-driving cars) their reaction is to regulate these objects in a way that society feels comfortable with. In a sense, they react to the risk that the new technologies pose at changing expectations and relationships. But gauging levels of risk at such an early stage of a technology roll-out, without any framework or in-depth studies, is a rather subjective undertaking. Their reaction of seeking regulation hence betrays a perception of risk skewed towards 'riskiness'. Or, in other words, regulators' reaction betrays their distrust. And examples from past and current regulations illustrate this tendency of being reactive to distrust (see below).

Here, a brief exploration of the term 'distrust' is first required. Trust and distrust share the tripartite relationship of A trusting B in given matter X, as well as their cognitive

[29] John Agnew, 'Steam engines on UK roads, 1862–1865: Banning orders, agricultural locomotives and the 'red flag' Act' 90 *The International Journal for the History of Engineering & Technology* 53.
[30] Robin Schwarzenbach, 'Der Kampf ums Automobil' (4 July 2016) *NZZ*.

account, meaning that both trust or distrust must be understood as knowledge or a belief based on the observation of the surrounding circumstances at a given point in time.[31] In addition, both trust and distrust seem like sensible and rational beliefs depending on the circumstances (e.g., trusting trustworthy parties vs distrusting untrustworthy parties).[32] But trust and distrust are not simply opposites: the absence or lack of trust does not mean that there is distrust; one can both not trust and not distrust a party.[33] While trust is defined as featuring a confident positive expectation that the risk of an interaction will not materialise, distrust is the belief and confidence of a negative expectation materialising.[34] Therefore, they tend to have 'substantially asymmetric implications for behavior and for society'.[35] Contrary to trust, distrust tends to lead to non-reliance on other parties and avoidance of interactions with distrusted parties[36] and, in the context of automation, disuse of automation.[37]

While trust and distrust are not mutually exhaustive, they are mutually exclusive.[38] If for a specific domain in given circumstances one distrusts another party, one cannot simultaneously trust that party for that same domain given the same circumstances. However, with changing circumstances and time, calibration of (initial) beliefs and knowledge can change. While some authors have argued that trust and distrust can co-exist in relationships, such reasoning is only sensible if not reduced to one specific interaction. This is also where understandings of trust and distrust as continuums rather than on–off switches can become useful.[39] When trust and distrust are understood as continuums ranging from high to low trust and distrust, it seems reasonable to assume that different circumstances will impact the range of trust/distrust-levels. In some circumstances, just one change of context might be enough to alter the trust/distrust-level. In others, changing circumstances might need to be accumulated to alter the balance of (high/medium/low) distrust to (low/medium/high) trust.

---

[31] Russell Hardin, *Trust & Trustworthiness* (Russell Sage Foundation 2002).
[32] Balázs Bodó, 'Mediated trust: A theoretical framework to address the trustworthiness of technological trust mediators' 23 *New Media & Society* 2668; Hardin (n 31).
[33] Bodó Jason D'Cruz, 'Trust and Distrust' in Judith Simon (ed), *The Routledge Handbook of Trust and Philosophy* (Routledge 2020).
[34] Trudy Govier, 'Is it a jungle out there? Trust, distrust and the construction of social reality' (1994) 33 *Canadian Philosophical Review/Revue canadienne de philosophie* 237; R. J. Lewicki and B. B. Bunker, 'Trust in relationships' 5 *Administrative Science Quarterly* 583; Roy J. Lewicki, Edward C. Tomlinson and Nicole Gillespie, 'Models of Interpersonal Trust Development: Theoretical Approaches, Empirical Evidence, and Future Directions' (2006) 32 *Journal of Management*.
[35] Hardin (n 31).
[36] D'Cruz (n 33).
[37] Maha Salem and others, 'To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability' (2013) 5 *International Journal of Social Robotics* 313; Jakub Złotowski and others, 'Appearance of a Robot Affects the Impact of its Behaviour on Perceived Trustworthiness and Empathy' (2016) 7 *Paladyn, Journal of Behavioral Robotics* 55.
[38] D'Cruz (n 33).
[39] Lewicki and Bunker (n 34); Lewicki, Tomlinson and Gillespie (n 34).

3.2 Case Studies: Aviation, GMOs and Disinformation

We pursue our investigation of how new technologies have become politicised with three different case studies showing slightly different characteristics: aviation, GMOs and disinformation. Despite their different characteristics, they all exemplify how distrust correlates with the transition towards politicising the topic.

Figure 1 illustrates where the different case studies and their associated phases map out on our model.

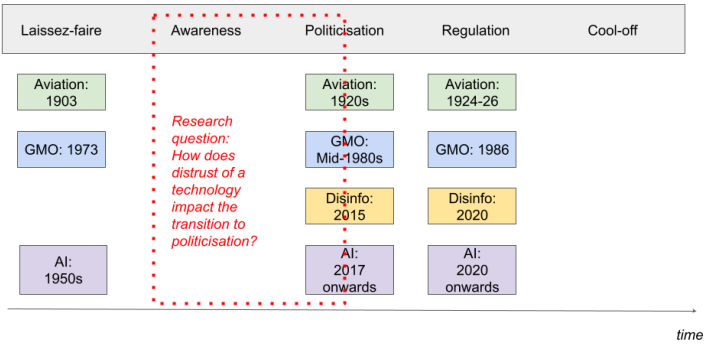

Figure 1: A model of how regulation transitions through different phases, with four cases as illustration and, for each case, the dates when the transitions occurred. The focus of this paper is on the cases placed onto the regulation model with the date.

3.2.1 Aviation

The case of regulating planes stands out for the twist that it brings to the usual regulation narrative. Until the mid-1910s, flying, as Elish and Hwang elegantly put it, was either 'a hobby for daredevils or the rich and insane'.[40] Because of the high number of accidents, it was understandable that there was a distinct distrust of planes.

Further to this, in the US, until at least 1924, commercial flying was regarded as having limited commercial applications,[41] apart from the niche market of aerial surveying, and the US Air Mail service. Slowly it became apparent that 'the absence of Federal regulation, besides creating a reluctance among the public to fly, had also created a reluctance among the financial community to invest in aviation'.[42] The industry

[40] Madeleine Elish and Tim Hwang, *Praise the Machine! Punish the Human! The Contradictory History of Accountability in Automated Aviation* (Intelligence and Autonomy Initiative 2015).
[41] Nick A. Komons, *Bonfires to Beacons: Federal Civil Aviation Policy Under the Air Commerce Act, 1926–1938* (U.S. Department of Transportation, Federal Administration 1978).
[42] Ibid.

therefore lobbied for federal regulations to come into force so that businesses would flourish (including insurance). Regulations therefore helped to change the perception from aviation being only a risky hobby for people with means to a useful technology with plenty of applications. It therefore had potential to displace other industries, namely the postal service and other forms of transport. In 1926, when the first commercial air regulation came into force in the US, the country was already behind some European countries, notably France, which had passed regulations two years earlier. The act of bringing it to the attention of legislators still arose from distrust: from people for their safety, and from banks and insurers for commercial viability.

The first law in France, loi relative à la navigation aérienne, covered wide ground. It defined what an airplane ('aéronef') is (Arts. 1–3), where it needed to be registered (those in France can only be registered by French nationals; Art. 5), how the mortgage and forfeiture had to occur (Arts. 15–18), where they are allowed to fly (for French registered airplanes, anywhere where it did not encroach on property's right or on delineated security zones; Arts. 19–20), how high they have to fly ('at a height such that landing be always possible'; Art. 21), whether they could do aerobatic flying (not above cities; Art. 22), the mandatory use of airports for take-off and landing (Arts. 24–30), the pilot licences required (Arts. 31–38), how goods and people could be carried (Arts. 39–48), how liabilities are to be handled between pilot and owner (Arts. 51–60), and how various criminal laws ought to apply (e.g., for piloting without a licence; Arts. 61–82).

Overall, the law would today be largely considered a 'common sense' view of what needed to be regulated. But at slightly over five pages long in the official publication, the law is also non-technical and quite straightforward to understand, in stark contrast to today's flying regulations. Likewise, the Air Congress Act of 1926 in the US was very similar to the French act, not only in length, but also on topics covered and on how it allowed or restricted the use of aircrafts.

3.2.2 GMOs

The politicisation of GMOs in Europe also illustrates the case of distrust reposing both on hypothetical and yet uncertain health risk (as opposed to known health risk evidenced in concrete cases); the regulatory reactions both focused on the technology itself and occurred at a time of flux for institutional trust (against American corporations in the mid-1990s). More specifically, politicisation of GMOs in Europe occurred very early and NGOs led protests against them, whereas in the US farmers embraced their arrival.[43] At that point, there had been no cases of someone's health or life being put at risk because of GMO consumption, so it remained based, again, more on a risk perception than on tangible evidence.

At the government official level, even if there were many important disparities within the EU which led to opposition (or lack of alignment) between the European Commission and the Council, GMOs had already entered the political realm. An

[43] Jan M. Lucht, 'Public Acceptance of Plant Biotechnology and GM Crops' (2015) 7 *Viruses* 4254.

example of such a lack of alignment was that France, Germany and Spain did not want to lag behind a 'core technology that is transforming agriculture'.[44] We see a similar 'fear of missing out' in the case of AI regulation put forward by the European Commission (more below on that). In France, support for GMOs quickly waned as forces mobilised, but more importantly, as Markowitz explains, as this grassroot pushback displaced the issue of GMOs onto American companies, and their ethics.[45] More specifically, one American company, Monsanto, dominated the market, attracting considerable attention from the media and activists for its approach, which was labelled 'arrogant and aggressive'. This led to, for instance, the influential French farming group *Confédération Paysanne* (farmer confederation) to '[broaden] the issue of GMOs from environmental and health concerns to issues connected to neoliberal globalisation, laying out the stakes as a struggle over the capture of agriculture to industrial lobbies and the accompanying biodiversity loss associated with this takeover'.[46] It did not help that there was widespread distrust in research published on GMOs due to alleged conflict of interests and transparency, with NGOs regularly accusing companies of covertly financing so-called 'independent research' in order to strengthen their (lobbying) case.

With AI, the criticism has not been on covert funding, but has still led to directing distrust onto the scientific process. As big tech companies have conducted their own research, then presented their results in scientific venues, they have often been accused of bias, with criticism also extending to the broader scientific process.[47] In other words, other issues, e.g. bias, crop up around the core issue, i.e., GMOs, helping make the case that the political class should address the issue, which is allegedly a broad one affecting a large part of the population. Here again a parallel with regulating AI is striking, the issues being less about automation than about privacy, or about biases, for example, which have little to do with AI per se, but which AI worsens.

The first regulation on GMO at the EU level was the Council Directive 90/220/EEC of 23 April 1990 on the deliberate release into the environment of genetically modified organisms, which followed a 1986 resolution from Member States, as stated in its preamble. Running at a mere 14 pages for such a technical topic (its 2001 update totalled 38 pages), the directive was reflective at this point in time of three aspects: current scientific uncertainty;[48] the will not to jeopardise the possibilities that the technology could offer; and the difficulty in reaching agreement due to the complexity of the topic.[49] As one scholar put it, 'the search for the delineation of what is allowed with [GMOs] involves trade-offs between protecting human dignity and

[44] Yves Tiberghien, 'Competitive Governance and the Quest for Legitimacy in the EU: the Battle over the Regulation of GMOs since the mid-1990s' (2009) 31 *Journal of European Integration* 389.
[45] Markowitz (n 21).
[46] Ibid.
[47] Nathaniel Persily, 'Facebook hides data showing it harms users. Outside scholars need access' *Washington Post* (5 October 2021).
[48] Ellen Vos and Michelle Everson, *Uncertain Risks Regulated* (Routledge 2009).
[49] Isabelle Wildhaber, *Haftung für gentechnische Produkte* (Lit Verlag 2009).

freedom to research, individual and societal interests, and between points of view from natural sciences, medicine, ethics, and law'.[50]

The 1990 directive was more specific when it came to notifications (Arts. 5–18), including around consent from authorities to release a product. That the bulk of the directive is directed towards notification reflects the desire to assess the risk and to keep in place a tight control in an early phase of deployment of the technology. On other issues, the directive lacked precision, the consequences of which were acknowledged by the European Commission in a 2001 report: 'Many people are losing confidence in a poorly understood and complex system'.[51]

A 1996 report identified two areas of the directive where improvement was desirable: the clear labelling to buyers; and the conduct of risk assessment and classification.[52] This led eventually to an updated version in 2001, and the politicisation of GMOs remained high for the next decade – the update process continued and a new version of the directive was published in 2015. This continual process of updating was also required because technology around the production of GMOs and sequencing DNA also continued to evolve,[53] leading to an increasing number of products being commercialised. Likewise, with AI, even before the first regulation came into force, developments and rapid uptake by users had forced legislators to reconsider many parts of the draft legislation (see below).

3.2.3 Disinformation

Another case study concerns attempts at regulating disinformation, especially disinformation circulating on social media platforms, and most notably the change that occurred in the willingness to employ regulatory tools to tackle the issue. While in the early stages of social media, acceptance in companies grew rapidly,[54] the consensus in the US and the EU was initially that disinformation on the internet could not be effectively regulated without limiting individual rights. In the 2010s, it was thought that regulation meant giving too much power to a single gatekeeper to decide, subjectively, on whether content should be classified as harmful. This could be easily misused for political gains, and would also have lead to authoritarian states dictating what their citizens were allowed to hear and say, and thereby losing the moral high ground. Social media companies, on the other hand, argued that monitoring and removing content would be impossible due to the sheer amount of content on their platforms. In light of such challenges, policy answers remained timid. For instance, the EU's diplomatic arm decided to launch in 2015 a counter-propaganda service, the EUvsDisinfo.

---

[50] Ibid.
[51] European Commission, *European Governance: A White Paper* (2001).
[52] European Commission, 'Commission presents report on directive 90/220/EEC on genetically modified organisms' (10 December 1990).
[53] Wildhaber (n 49).
[54] Hannah Kuchler, 'How Facebook grew too big to handle' *Financial Times* (28 March 2019).

However, since 2017, several incidents involving high ranking politicians have changed the tide. In 2017, while campaigning for the French presidency, Emmanuel Macron was the repeated target of Russian 'bots'; in 2018, the Cambridge Analytica scandal brought to light how political actors were using targeted ads to increase the impact of their disinformation campaigns; in 2019, Nancy Pelosi, the Speaker of the House of Representatives in the US, battled to have a deep fake video (a doctored video easy to mistake for genuine) of her taken down; in 2020, in the midst of the Covid-19 pandemic, Twitter started marking posts from the then President Trump as containing inaccurate information about the disease. Further incidents also played a role in initiating regulatory change – most notably in the context of Covid-19 and Russia's war with Ukraine.[55]

These events led to EU Member States passing their own legislation to tackle disinformation. Germany started discussion on the *Netzwerkdurchsetzungsgesetz* in 2015 (and voted on it in 2017), a law originally designed to combat hateful content online, but extended to include 'fake news' following the US 2016 Presidential elections. In France, upon taking office in early 2018, Macron's government introduced in parliament the 'Loi contre la manipulation de l'information', the law being voted on in November of the same year. The law gives a judge 48 hours to rule on issues brought forward by any individual. At the EU level, similarly, the introduction in 2018 of the EU Code of Practice on Disinformation prior to the 2019 EU Parliament elections preceded the introduction in 2020 of the Digital Service Act (DSA), which sought to tackle the same problem, namely illegal content (and not the broader harmful content, which is more difficult to pin down). The DSA seeks to make social media companies with more than 45 million users in the EU (more) liable for what occurs on their platforms. The new obligations include a mandatory 'notice-and-action' requirements, mandatory redress for content removal decisions, and a comprehensive risk management and audit framework. The European Parliament passed the DSA in April 2022, and it is in force as of February 2024.

The EU's DSA has attempted to approach disinformation in its broadest sense possible to include misinformation, the influencing of operations by foreign intelligence services, and information manipulation.[56] But in opposition to France's approach which gives a tight deadline to rule on cases, the DSA puts the onus on platforms, especially for what it calls Very Large Online Platforms, to adopt safeguards and to be transparent about them (see e.g., Art. 14 on publishing restrictions imposed, as well as Art. 17 on reporting accounts suspended, or Art. 34 on conducting systemic risk assessment). However, the DSA still preserves exceptions for service providers and other hosts in terms of their liability for content (Arts. 4–6), in continuation of Arts. 14–15 of the 2000 E-Commerce Directive (Directive 2000/31/EC), which had already

[55] Antonio Manganelli and Antonio Nicita, *Regulating Digital Markets: The European Approach* (Palgrave McMillan 2022).

[56] Sharon Galantino, 'How will the eu digital services act affect the regulation of disinformation?' (2023) 20 *SCRIPTed: Journal of Law, Technology and Society* 89.

removed this liability by treating hosts as mere conduits for information (with caveats) and requested 'no general obligation to monitor' content.

There are therefore asymmetric requirements, with Very Large Online Platforms defined as platforms with a 'number of average monthly active recipients of the service in the Union equal to or higher than 45 million' (Art. 33). The DSA singles these out to curb systemic risk of disinformation by mandating them to: 'diligently' review their algorithms (Art. 34); adjust their recommender systems (Art. 27); have mechanisms in place to give notice of possible disinformation (Art. 16); and publish yearly audits (Art. 37) – all in essence to 'focus on promoting reliable and information over-sensationalised, and misleading content'.[57]

The Act has not been without criticism, for example that the platforms do not have to be transparent about how they take decisions regarding content moderation,[58] and that the split between different authorities to enforce the Act – notably between the Digital Service Coordinator at the Member State level and the European Commission – could cause frictions.[59] Yet another criticism has been the 'excessive' reliance on transparency mechanisms with little evidence that users behave differently when presented with more information.[60]

A testimony to the effect of the law, and even to the climate of distrust from the EU politicians prior to legislating, has been how Facebook's growth of its user base slowed during this time, with the user base starting to decline in 2022 – although it is not possible to make a causal link between users' distrust and them leaving the platform.[61] A key driver to regulate disinformation has been the fear that disinformation could pose to functioning democracies, a point which also comes up regularly with AI.

3.3 Politicisation of AI in the EU

With the publication of the AI Act in April 2021, the EU had clearly entered the fourth stage of our model: regulating. Of interest to this study is what, prior to the regulating stage, led to the transition from the 'AI is left to academic and industry researchers' stage to the one of politicians making room in the agenda for the topic. EU politicians are not detached from the rest of the world, and developments to both the west and east of the EU influenced this politicisation. There is little indication that politicisation of AI per se – taken narrowly to understand the explicit mention of AI and not of any

[57] Mark Leiser, 'Reimagining Digital Governance: The EU's Digital Service Act and the Fight Against Disinformation', https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4427493.
[58] Galantino (n 56).
[59] Alain Strowel and Jean De Meyere, 'The Digital Services Act: transparency as an efficient tool to curb the spread of disinformation on online platforms?' (2023) 14 *Journal of Intellectual Property, Information Technology and E-Commerce Law.*
[60] Leiser (n 57).
[61] Rob Thubron, 'Meta shares plunge 20% after Facebook user numbers fall for the first time ever' *Techspot* (3 February 2022).

related terms – began prior to 2016: no countries had an exclusive national strategy for it, and, if the topic was picked up by politicians, it was usually only marginally.

Still, the topic was present, at least in the press. For most of the 2010s, certain successes led to changes, with AI increasingly finding its way into news pieces and, consequently, into conversation.[62] But the tone was futuristic, hopeful, and at times science fiction. Consider this quote from a 2014 *Wired* article on the topic:

> 'Amid all this activity, a picture of our AI future is coming into view, and it is not the HAL 9000—a discrete machine animated by a charismatic (yet potentially homicidal) humanlike consciousness—or a Singularitan rapture of superintelligence. The AI on the horizon looks more like Amazon Web Services—cheap, reliable, industrial-grade digital smartness running behind everything, and almost invisible except when it blinks off. [...] Like all utilities, AI will be supremely boring, even as it transforms the Internet, the global economy, and civilization'.[63]

The general press has echoed these views, with the BBC prominently relaying Stephen Hawking's comment that AI 'could end mankind' (while asking at the same time whether AI threatens humanity).[64] This threat to mankind (and to human autonomy) is not purely anecdotal. A study of news media revealed that the top topics on AI, regardless of date, were movies, novels and AI research, much in line with futuristic forecasts and science fiction.[65]

And yet, in 2014 and 2015, politically, very little occurred in relation to AI – or if it did, it did not cross the level of visibility that either a national strategy, new department, allocation of resources, or bill creates. Despite the far-reaching consequences predicted across all aspects of life – aspects which could have elicited needs for regulation – no US, Chinese or EU agency moved towards politicising or regulating AI. Decision-makers may be excused for having been dismissive of such predictions, if not at least for the tone in which authors presented them. Only in 2016, with yet another seemingly benign and niche breakthrough, did the political sphere start not only taking notice but acting too – and legal research on the topic also changed gear.[66]

On 19 March 2016, Google's DeepMind programme beat the world's best player at the Chinese game Go, trouncing experts' expectations of advances within 10–15

---

[62] Lea Köstler and Ringo Ossewaarde, 'The making of AI society: AI futures frames in German political and media discourses' (2022) 37 *AI & Society* 249.
[63] Craig Karl, 'The Three Breakthroughs That Have Finally Unleashed AI on the World' *Wired* (27 October 2014).
[64] Rory Cellan-Jones, 'Stephen Hawking warns artificial intelligence could end mankind' *BBC* (2 December 2014); Mark Ward, 'Does rampant AI threaten humanity?' *BBC* (2 December 2014).
[65] Yujia Zhai and others, 'Tracing the evolution of AI: conceptualization of artificial intelligence in mass media discourse' (2020) 48 *Information Discovery and Delivery* 137.
[66] Constanta Rosca and others, *Return of the AI: An Analysis of Legal Research on Artificial Intelligence Using Topic Modeling* (CEUR 2020).

years.[67] Gaining less media attraction but exemplifying a more problematic side of AI, on 23 March 2016, Microsoft released on Twitter the Tay AI 'bot', only to withdraw it 16 hours later as it was making offensive and inflammatory tweets, presumably fed by internet 'trolls'. A month later, on 28 April 2016, Google's CEO advertised in a marketing coup still telling of the time that the company was now going to be 'AI First'. While this could initially appear almost irrelevant and far-removed from the political realm it was nothing but; it led to the securitisation of AI, pitting countries in a nationalistic race against each other in this field – or an 'AI cold war' as one author put it.[68] On 3 May of the same year, the White House under President Obama called for the 'preparing for the future of AI', citing 'a series of breakthroughs in the research community and industry'.[69] A few months later, in October 2016, as the country was about to elect President Trump, the Office of the President released its AI strategic plan.[70] China was not mentioned once in the document at the time, but framing the strategy as a competition between the two countries would come to shape their relation and the field.[71]

In the EU, a May 2017 Commission review of the Digital Single Market Review mentioned AI only passingly: 'the Commission will continue to monitor the opportunities and challenges brought by artificial intelligence solutions'.[72] Only in October 2017 can we find the first EU note, when the European Council invited the European Commission to act, out of 'a sense of urgency to address emerging trends'.[73] Interestingly, the call strongly shaped the outcome as it mentioned a 'risk-based' framework; and it strongly displayed a nationalistic competition frame of mind with the aim of the EU to 'reaffirm the leading role of its industry'. The European Commission subsequently obliged and then published a call to create a High-Level Expert Group (HLEG) in March 2018; in June 2018 the Group was ready; and in April 2019 it published its first report. More political negotiations ensued until April 2021, when the EU unveiled its proposed AI Act.

This timeline is important if we are to understand how, why and when AI entered the political realm. Did anything scare decision-makers, and if yes, what? AlphaGo represented AI surpassing humans at a game, which had not been expected anytime soon. Although some interpreted this is a matter of concern in and of itself, it also gave rise to fears of what could come next, and a 'fear of missing out' on regulation,

---

[67] Christopher Moyer, 'How Google's AlphaGo Beat a Go World Champion' *The Atlantic* (18 March 2016).
[68] Denise Garcia, 'Stop the emerging AI cold war' (2021) 593 *Nature*.
[69] Ed Felten, 'Preparing for the Future of Artificial Intelligence' *The White House* (Washington, DC, 3 May 2016).
[70] National Science and Technology Council, 'The National Artifical Intelligence Research and Development Strategic Plan' (2016).
[71] Jinghan Zeng, 'Securitization of Artificial Intelligence in China' (2021) *The Chinese Journal of International Politics* 417.
[72] European Commission, 'Communication on the Mid-Term Review on the implementation of the Digital Single Market Strategy' (2017).
[73] European Council, 'European Council meeting (19 October 2017) – Conclusions' (2017).

especially in relation to the US and China – China's first strategy was published in July 2017 – also played a role.

There is evidence for both of these points. First on the 'fear of missing out': the March 2018 call for experts clearly mentioned the upside of AI as a reason for setting up the strategy.

> 'To make the most of it for the citizens and the economy'
> 'AI offers major business opportunities for European industry, SMEs and start-ups and contributes to productivity growth'

The April 2018 AI strategy also has a lengthy, blunt section on competitiveness, which makes clear what drives it: 'the EU risks losing out on the opportunities offered by AI, facing a brain-drain and being a consumer of solutions developed elsewhere'.

On the second point of fearing what would come next, the March 2018 call had already hinted at the centrality of trust: 'the emergence of AI also raises legitimate concerns that should be addressed to build trust and raise awareness'. The first deliverable was to be a draft of AI Ethics Guidelines,[74] further reflecting the high priority that this represented for the EU, and unleashing much (academic) debate on what constitutes 'ethical AI'.[75] The 2018 AI strategy took this much further, stating AI to be 'one of the most strategic technologies of the 21st century': 'The stakes could not be higher', it wrote, before adding that 'European leaders have put AI at the top of their agendas'.[76] Further evidence includes: 'The first challenge is to *prepare the [sic] society as a whole*' [emphasis in bold in the original], and references to an 'AI revolution'. Within six months, a change of tone was noticeable, with this type of discourse bringing it close to framing the topic as an existential topic, a first indication of the needle moving from politicisation to securitisation (albeit not quite passing the threshold). The strategy seeks to focus on three areas, one of which is on being competitive internationally, and another one on the need to build trust as 'AI applications may raise new ethical and legal questions' (trust is mentioned seven times in the 19-page document). If trust must be 'built', it means that distrust is likely present.

A ten-page communication published on 7 December 2018, repeating the strategy's argument that AI is like electricity in that it transforms the world, takes only until the third paragraph to jump from opportunities to issues.[77] It makes sense that this led to further EU documents with focus on issues and trust in AI. It is hence no surprise that 'trust' appears 145 times in the very first 41-page document produced by the High-Level Expert Group (April 2019). Such communication and strategy documents are evidence of how thinking and approaches at the highest level of EU institutions

[74] European Commission, 'Concept note: The High-Level Expert Group on Artificial Intelligence' (2018).
[75] Anna Jobin, Marcello Ienca and Effy Vayena, 'The global landscape of AI ethics guidelines' (2019) 1 *Nature Machine Intelligence* 389.
[76] European Commission, 'Communication on Artificial Intelligence for Europe' (2018).
[77] European Commission, 'Communication on Coordinated Plan on Artificial Intelligence' (2018).

emerged, evolved and, in the end, formed part of the views that shaped the regulation of AI coming out of these very EU institutions.

Similar to the GMOs case study mentioned above, many elements also come into play within the politicisation of AI: while early on, discussions focused on discriminative models, which use machine learning to classify content (e.g., AI used for the creation of recommendations, spam filtering, image classification, etc.), generative AI models became more popular in 2017, especially with the introduction of the Transformer architecture.[78] This led to the establishment of different generative AI systems (e.g., GPT, Bert, LaMDA).

An important development for our study happened in November 2022 with the release of OpenAI's online demonstrator of the ChatGPT system, which was widely shared on social media and caught the attention of the broader public including policymakers (European Parliament, 2023).[79] Even though ChatGPT is based on GPT-3, which had been previously available on OpenAI Playground, the ChatGPT version spread more quickly through the general population; this led, for example, to school teachers fearing how they could ensure that their current curriculums and assignments were still adequate in light of the text generation abilities of ChatGPT,[80] as well as, separately, concerns in the journalistic and creative industries.[81]

Within the European political discourse, as mentioned above, regulatory shaping started after the publication of the HLEG report. That report set the foundational principles for ensuring a trustworthy AI environment within the EU and included requirements on human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being; and accountability. These high-level ideals were integrated into the proposed AI Act, which was made public in April 2021. The proposed AI Act states that AI is a 'fast evolving family of technologies' (Rec. 4 of the adopted text) and has a broad material scope which includes machine learning approaches as well as logic- and knowledge-based approaches (Annex I) – the large scope has also been criticised in the legal scholarship.[82] The AI Act sets obligations for providers of AI systems, i.e., entities that develop AI systems, as well as implementers, distributers, and importers of AI systems (Art. 3).

---

[78] Ashish Vaswani and others, 'Attention Is All You Need' 31st Conference on Neural Information Processing Systems (NIPS 2017) .

[79] European Parliament, 'EU AI Act: first regulation on artificial intelligence' *News European Parliament* (8 June 2023), www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

[80] Alex Hern, 'AI-assisted plagiarism? ChatGPT bot says it has an answer for that' *The Guardian* (31 Dec 2022).

[81] David De Cremer, Nicola Morini Bianzino and Ben Falk, 'How Generative AI Could Disrupt Creative Work' *Harvard Business Review* (13 April 2023).

[82] Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 22 *Computer Law Review International* 97.

The AI Act follows a clear risk-based approach: it classifies risk as unacceptable and bans AI systems that, for instance, use subliminal techniques to manipulate a person's behaviour in a manner that can cause harm; exploit vulnerabilities of people due to their age, physical, or mental disability and thereby cause harm; or enable governments to use general-purpose 'social credit scoring' (Art. 5). The AI Act classifies as high-risk AI systems that are used as a safety component of a product, or if it is covered by one of the pieces of EU single market harmonisation legislation listed in Annex II (Art. 6). In addition, AI systems deployed in the following sectors are deemed to be high-risk to safety or fundamental rights: critical infrastructure where the AI system could put people's life and health at risk; in education contexts (e.g., when AI systems are used to determine access to education); or in employment contexts (e.g., when AI is employed to recruit employees and manage their performance). To develop or use a high-risk AI system, an organisation must meet a range of technical and regulatory requirements before the system can be brought to market. This includes: establishing a risk management system (Art. 9); establishing safeguards against various types of biases in datasets and using prescribed data governance and management practices (Art. 10); having technical documentation in place (Art. 11) and fulfilling record-keeping duties (Art. 12); ensuring the ability to verify and trace back outputs throughout the system's life cycle and incorporating provisions for acceptable levels of transparency and understandability for users of the systems (Art. 13); ensuring appropriate human oversight over the system generally (Art. 14); and ensuring accuracy, robustness and cybersecurity (Art. 15).

While the risk-based approach and key concepts of the proposed AI Act are included within the adopted version of the EU Parliament from June 2023, the policymaking discourse has shifted from the focus on discriminative models to including foundation models that are at times referred as 'general purpose AI'.[83] The foundation models challenged the current political approach, as the Ada Lovelace Institute puts it:

> 'As policymakers begin to regulate AI, it will become increasingly necessary to distinguish clearly between types of models and their capabilities, and to recognise the unique features of foundation models that may require additional regulatory attention.'[84]

The term 'foundation model' thus entered the revised AI Act (Parliament version adopted on 14 June, P9_TA(2023)0236) and a new focus on generative AI emerged (e.g., Art. 28(b) of the Parliament version) with more rules on demonstrating the appropriate design of the models with respect to mitigating foreseeable risks to health, safety, fundamental rights, the environment and democracy (Art. 28(b)(2)(a) Parliament version); disclosing content that was generated by foundation models, especially ones that generate text, images, audio or video (i.e., generative AI) (Art. 28(b)(4) with reference to Art. 52 of Parliament version, which further makes explicit

[83] Elliot Jones, 'Explainer: What is a foundation model?' (17 July 2023); Luca Bertuzzi, 'AI Act: EU countries headed to tiered approach on foundation models amid broader compromise' *Euractiv* (17 October 2023).
[84] Jones (n 83).

that deep fake information must be disclosed in a clear and visible manner (Art. 52(3)); designing models that do not enable users to generate illegal content (Art. 28(b)(4)(b) of the Parliament version); and further transparency and documentation obligations with respect to the data used to train the models (Art. 28(b) and Art. 52 of the Parliament version).[85] Art. 52 further outlines the relevant information that must be provided, which includes 'if there is human oversight, and who is responsible for the decision-making process, as well as the existing rights and processes that, according to Union and national law, allow natural persons or their representatives to object against the application of such systems to them and to seek judicial redress against decisions taken by or harm caused by AI systems, including their right to seek an explanation' (Art. 52(1)). To enforce the regulation, discussions on establishing an EU AI Office have emerged rather than just leveraging national supervisory authorities.[86]

***

Each of the cases in section 3.2 raise different issues in relation to AI. In the case of disinformation, the fear was more for the state of democracy than one related to health and safety. In addition, to take two extremes, there is a notable difference between the uncertainties that were introduced by early aviation than by GMOs: on the one hand, it was well known that a plane crash could kill you, even if assessing the probability of a plane crashing was difficult, while on the other hand, it was difficult to discern early on what the long-term effects of GMO cultivation and ingestion could be. The disinformation case is somewhat similar to the GMO case: there are hypotheses, some with more early evidence than others, on how they further polarise society, stymie the political process, undermine institutions and authoritative experts, and more.[87] But a definitive account of such consequences is yet to be drawn.

4. An Attempt at Causality: Distrust → Politicisation

4.1 A Case Study: No Distrust, No Politicisation

Having illustrated the correlation between distrust and the politicisation phase of regulating a new technology, the question is now whether it would be possible to derive a stronger logical relation, namely one of causality. Causality is a fraught and complex topic. In order to tease out causality, one social science methodology has been to use counterfactuals.[88] By identifying (seemingly) similar causes and similar

[85] We note that references to generative AI and foundational model was removed (previously Art. 28), but that the Art. 52 in relation to deep fakes was retained in the final version under Art. 50.
[86] Bertuzzi (n 83).
[87] Thomas Rid, *Active Measures: The Secret History of Disinformation and Political Warfare* (Profile Books 2020).
[88] Henry E. Brady, 'Causation and Explanation in Social Science' in Robert E. Goodin (ed), *The Oxford Handbook of Political Science* (Oxford University Press 2011).

effects across situations in the real world, we might come as close to establishing causality as possible.

Here, as we are interested in showing that distrust of a technology leads to its politicisation ($D_{AI} \rightarrow P_{AI}$), we need to find a similar world (x) where no distrust of a new technology led to it not being politicised ($\neg D_x \rightarrow \neg P_x$). Such technologies exist but they are difficult to find organically: VCR raised issues of copyright, radio of airwaves, and cars of road damage, just to name a few. But one technology where we could not find any attempts to politicising it in the EU are self-checkout systems in retail environments ($\neg P_{self\text{-}checkouts}$). While many regulations (e.g., on specific sales) will automatically apply to the self-checkouts, there has not been any call in the EU for further regulations. How similar are the worlds of AI and self-checkouts? Both automate human tasks. Naturally, one is much more general than the other, and depending on the exact implementation, AI could (but does not have to) take away agency from users.[89] The parallels do extend to the few macro factors on which we based our comparison for the other case studies. For a start, AI and self-checkouts follow a generally similar timeline, indicating that the socio-politico-cultural environment in which these technologies emerged ought to be considered similar: both have their original roots in the 1960s,[90] but both only started picking up steam from 2010 onwards (the number of self-checkout users grew to hundreds of thousands globally by 2017).[91] Beyond the environment being comparable, that the user numbers keep growing shows the interest of retail vendors in reducing their staffing costs – and hence the obvious associated potential that self-checkouts have to replace cashiers – as well as the interests of shoppers who have found certain comfort in them.

Several fears due to their usage are also inherent, but remain for now open questions: do they lead to more theft and theft attempts, thereby increasing safety risk for retail staff – notwithstanding that many cashiers now have to operate as ersatz security guards, monitoring and intervening in potential theft cases? Do they lead to more sales of tobacco and alcohol (to youngsters or any other groups), thereby increasing health risks due to consumption of such products? Do they increase liability lawsuits against retail chains, either from any injuries occurring to buyers in the process, or from unauthorised sales of cigarettes or alcohol to minors, or yet again, for genuine mistakes from buyers now accused of attempted theft? Hypothetical fears are therefore present. In all of these cases, the EU would have had competency to introduce regulation; Art. 114(3) of the Treaty on the Functioning of the European Union (TFEU) states that the EU can do so when it comes to 'health, safety, environmental protection and consumer protection', so it could therefore have emphasised how self-checkouts are a potential jeopardy to cashiers' safety,

[89] Aurelia Tamò-Larrieux, Andrei Ciortea and Simon Mayer, 'Machine Capacity of Judgment: An interdisciplinary approach for making machine intelligence transparent to end-users' (2022) 71 *Technology in Society* 102088.
[90] Adriana Hamacher, 'The unpopular rise of self-checkouts (and how to fix them)' *BBC* (9 May 2017).
[91] Ibid.

consumers' health, and consumer protection against accusation of theft. Safety, health and consumer protection are also '[s]hared competence between the Union and the Member States' according respectively to Art. 4(2)(k) and Art. 4(2)(f) of the TFEU. In addition, the EU's competence to regulate could have stemmed from the widespread presence of self-checkouts which therefore arguably present potential for disturbance to the functioning of the internal market for which the EU has the mandate to act under Art. 4(2)(a) of the TFEU.

Can we establish that there is no distrust of self-checkouts? The trust relation can be somewhat complicated: tweaks in the user experience (with the interface, voice, etc.) and marketing campaigns to use them point at attempts to increase both their usage and related public trust.[92] The increase in the number of self-checkout systems, and studies noting increased user satisfaction as well as consumer acceptance, paint a picture in which there is no clear distrust.[93]

As such, we conclude that self-checkouts are a case study in which no distrust leads to no politicisation ($\neg D_{\text{self-checkouts}} \rightarrow \neg P_{\text{self-checkouts}}$).

We are cautious of establishing a strong causal link, as we are cognisant of the multitude of factors that play a role in seeking historical explanations for the phenomenon. For this reason, we make two caveats. First, this is not a deterministic causal link but a probabilistic one: the occurrence of the cause – distrust – increases the probability of the effect: politicisation. Second, we are not saying that distrust is necessary for politicisation to happen, as other causes can also lead to politicisation (see below). The question then becomes whether distrust is sufficient for politicisation, and if not, what other conditions must be present for it to occur. To strengthen our case for the logical relationship between distrust and politicisation, we therefore need to investigate how distrust is situated within the sufficient and necessary conditions for politicisation.

4.2 INUS Conditions for Politicisation

INUS stands for 'a set of conditions requiring that a cause be an insufficient [I] but necessary [N] part of a condition which is itself unnecessary [U] but exclusively sufficient [S] for the effect'.[94] (The N part can also be 'non-redundant'.) INUS can be used to break down complex, multi-causal relationships. The typical example used to explain the INUS condition is the analysis of the cause of a fire. Three elements are needed to start a fire: oxygen, flammable material and a spark. For instance, if a house is on fire there are multiple factors that can be the cause of it. The [U] means that the condition is not necessarily exclusive, i.e., that other conditions can lead to the same effect, rather than meaning that is not needed at all to generate the cause (Table 1

[92] Ibid.
[93] Denis Vuckovac and others, *From Shopping Aids to Fully Autonomous Mobile Self-checkouts – A Field Study in Retail* (AIS 2017); Kerem Katri and Saar Sirli, 'Consumer acceptance of self-service checkouts in Estonian retail market' 49th Proceedings of the European Marketing Academy (2020).
[94] Brady (n 88).

gives two examples that would cause a fire, thus the condition being unnecessary on its own as other conditions could lead to the same outcome).

Table 1: Breakdown of IN/US conditions for a fire

IN (insufficient but necessary)	US (unnecessary but sufficient)
Faulty wiring in the house, which in case of a short circuit cannot cause a fire on its own	A short circuit producing a spark near flammable material
Any flammable material on its own cannot cause a fire	Striking a match near flammable material in an open-air environment

Within this example, the open flame is sufficient to cause a fire and the short circuit in the wiring contributes to the fire when combined with the open flame. This helps to break down the causes and the relationship between the causes. In this case, the short circuit contributes to the fire, even if insufficient on its own. INUS thus does not rule out situations with common causes; and the causes do not have to be exclusive to one set of conditions only (this is of lesser importance to us, as we only attempt to determine that distrust is a causal factor).

With respect to our analysis of distrust and AI regulation, we analyse the different conditions to determine whether distrust can be seen as an IN condition, i.e., that it is insufficient on its own but will, when combined with other factors, lead to politicisation. Table 2 illustrates this case, and we elaborate on it below.

Table 2: Breakdown of IN/US conditions for politicisation of new technologies

IN (insufficient but necessary)	US (unnecessary but sufficient)
Help a politician be re-elected	Fear of worker displacement
Symbolism of EU institutions over domestic ones	Distrust

4.2.1 IN

The question of what makes it onto the political agenda is one of political science's key questions,[95] often addressed under the heading of 'agenda-setting'. The question of which major factors contribute to it can be answered by the snappy quote: 'successful politicians instinctively understand which issues benefit them and their party and which do not. The trick is to politicise the former and deemphasise the latter'.[96] For a topic to benefit a party may mean a few different things: it can be a strategy to own a topic (e.g., labour topics for a labour party) and be associated with it for the next election campaign; or it could mean the party can push for certain preferred policies.[97] The consequence of either of these is that politicians, as

[95] Simon Otjes, '"No politics in the agenda-setting meeting": plenary agenda setting in the Netherlands' (2019) 42 *West European Politics* 728.
[96] Edward G. Carmines, 'The Logic of Party Alignments' (1991) 3 *Journal of Theoretical Politics* 65.
[97] Sebastiaan Princen, 'Agenda-setting in the European Union: a theoretical exploration and agenda for research' (2007) 14 *Journal of European Public Policy* 21.

gatekeepers, determine which issues to deal with and hence give the issues legitimacy. Princen (2007) makes the point that the goal is often not merely to move a topic into the political arena, but rather to move its importance higher up (and to move other topics further down).[98] This happens – although *how*, rather than *why*, politicisation occurs is out of scope here – either via 'outside lobbying' (e.g. public mobilisation) or 'inside lobbying' (e.g. by trying to directly influence decision-makers). Lobbying offers, however, an interesting insight as to why firms often have recourse to it and why politicians play along: because firms seek to raise costs for competitors (from abroad or even domestic ones).

With politicisation through EU institutions, there are a few further particularities. Media within Member States can certainly play a role in triggering politicisation at the Member State level,[99] and to possibly bring this up to the EU level: for instance, an EU Member State may choose to bring an issue to the EU level in order to overcome domestic opposition, or to associate it with specific values and symbols,[100] a theory known as 'venue shopping'. A reason for choosing the EU over a domestic audience may also be, depending on the Member State, because the EU offers so many more access points to try to influence policy, from the European Commission to working groups in Parliament.[101]

Our goal here is not to draw up a comprehensive list of all possible factors explaining why politicisation occurs, but merely to show that there are other combinations of factors possible, making the case that distrust is an insufficient but necessary factor to result in politicisation. Interestingly, though, some of the possible drivers considered in this article receive little attention in the literature. For instance, possible safety issues can become emotional for a population, bringing awareness of the frailty of life. These fears are low-hanging fruits for political parties seeking to capitalise on them, as the cases for GMOs or aviation exemplified. And the literature does make mention of it as a driver, albeit scarcely.[102]

Also interesting is that certain of these factors were present in the case for self-checkouts: fears of safety (e.g., for cashiers turned security observers against theft) could have emerged, or the companies implementing the technology could have sought to create an 'economic moat' to guarantee a competitive advantage. But so far, this is not noticeable, which leads to us cautiously conclude that these factors on their own are insufficient and need to occur in combination with others.

---

[98] Ibid.

[99] Peter Van Aelst and Stefaan Walgrave, 'Political agenda setting by the mass media: ten years of research, 2005–2015' in Nikolaos Zahariadis (ed), *Handbook of Public Policy Agenda Setting* (Edward Elgar Publishing Limited 2016).

[100] Frank R. Baumgartner and Bryan D. Jones, *Agendas and instability in American politics* (The University of Chicago Press 1993).

[101] Guy Peters, 'Agenda-setting in the European Union' in Jeremy J.Richardson (ed), *European Union: Power and policy-making* (Routledge 1996).

[102] Nikolaos Zahariadis, *Handbook of Public Policy Agenda Setting* (Edward Elgar Publishing Limited 2016).

Table 3: Summary table of a few other factors leading to politicisation

Why did an issue become politicised?	True for the AI case study?
Benefit to a politician and their party	No
Strategy for elections	No
Legitimate an issue	Yes
Safety / emotional	Yes
Resulting from a push from abroad	Yes
To change the importance of the issue (up or down)	No
To create an economic moat (for bespoke businesses)	Yes
Venue shopping with more access points	Possible

4.2.2 US

In order to narrow down which exact combination of factors on top of distrust lead to politicisation, we need to compare once more our case studies and find the smallest set of factors where the effect (i.e., politicisation) still occurred. It is clear that our case studies on aviation, GMOs, disinformation and AI share the common element of a fear of potentially displacing workers in a specific industry. Formally, we would need to show that politicisation did not happen if this factor was not there. However, this is exactly a limitation of INUS conditions where it is not possible to rule out common causes in a situation, and we are therefore in a dilemma.[103] But we can also note that some factors, such as the fear of a functioning democracy, sometimes played a role and at other times, did not; for instance, when in the presence of uncertainty about the technology and hypothetical safety concerns. The same can be said about the lack of agency, where it plays a role for passengers on planes and for shoppers without GMO labels, but where it does not for readers of news in the context of disinformation, thus it is probably only a driver in combination with other factors. We consider these examples to be sufficient for us to conclude that distrust also needs to be in combination with other factors to lead to politicisation – an almost foregone conclusion when considering that social developments rarely or never consist of a single causal element. This hence allows us to conclude that distrust is an INUS condition for politicisation.

5. Conclusion

The politicisation of AI in the EU following distrust of the technology clearly has two aspects: although there is a willingness to harness the upside of the technology, especially within this framing of a race to the top, the EU has also sought to achieve 'trustworthy AI', both as a result of distrust as a driver and as the consequence of seeking to leverage the technology's potential. With any new technology, there are

[103] Brady (n 88).

certain positive and negative aspects. Distrust acts as a bias towards one of these sides: it gives prevalence to seeking regulatory actions to control the new technology. Some of it might, in the end, be justified. But there is a danger in prioritising prejudices over a more in-depth risk assessment, especially so early in the regulatory process.