European artificial intelligence "trusted throughout the world": Risk-based regulation and the fashioning of a competitive common AI market

Regine Paul

Department of Government, University of Bergen, Bergen, Norway

Abstract

The European Commission has pioneered the coercive regulation of artificial intelligence (AI), including a proposal of banning some applications altogether on moral grounds. Core to its regulatory strategy is a nominally "risk-based" approach with interventions that are proportionate to risk levels. Yet, neither standard accounts of risk-based regulation as rational problem-solving endeavor nor theories of organizational legitimacy-seeking, both prominently discussed in *Regulation & Governance*, fully explain the Commission's attraction to the risk heuristic. This article responds to this impasse with three contributions. First, it enrichens risk-based regulation scholarship—beyond AI—with a firm foundation in constructivist and critical political economy accounts of emerging tech regulation to capture the performative politics of defining and enacting risk vis-à-vis global economic competitiveness. Second, it conceptualizes the role of risk analysis within a *Cultural Political Economy* framework: as a powerful epistemic tool for the discursive and regulatory differentiation of an uncertain regulatory terrain (semiosis and structuration) which the Commission wields in its pursuit of a future common European AI market. Thirdly, the paper offers an in-depth empirical reconstruction of the Commission's risk-based semiosis and structuration in AI regulation through qualitative analysis of a substantive sample of documents and expert interviews. This finds that the Commission's use of risk analysis, outlawing some AI uses as matters of deep value conflicts and tightly controlling (at least discursively) so-called high-risk AI systems, enables Brussels to fashion its desired trademark of European "cutting-edge AI … trusted throughout the world" in the first place.

Keywords: artificial intelligence, competition state, critical policy analysis, cultural political economy, EU, risk-based regulation.

1. Introduction

> *We need to understand the possibilities and limitations of AI systems as a means to build the future we want.*
> AI High Level Expert Group (2019, p. 15)

> *By striving towards human-centric AI based on trust, we safeguard the respect for our core societal values and carve out a distinctive trademark for Europe and its industry as a leader in cutting-edge AI that can be trusted throughout the world.*
> European Commission (2019, p. 9)

After a decade of disruptive innovation and voluntary ethical guidelines (Hagendorff, 2020), state regulators have recently discovered artificial intelligence (AI)[1] technologies as a matter of more coercive regulation. In 2019, the United States proposed an Algorithmic Accountability Act (followed by a blueprint for an "AI Bill of Rights" in October 2022) and China published the Beijing Artificial Intelligence Principles. Canada established its Algorithmic Impact Assessment Tool one year later (Etziani et al., 2021). The most far-reaching so far, the European Commission published a comprehensive AI regulation proposal in April 2021, and Parliament submitted its position for tripartite negotiations with member states and the Commission in June 2023. The proposal is likely to be adopted in amended form in early 2024. With the so-called "AI Act", Brussels proposed the first ever *coercive* and *sanctionable* regulation of AI. It suggests outright bans, mandatory risk mitigation, and certification duties of some AI systems and uses, as

Correspondence: Regine Paul, Department of Government, University of Bergen, Christies Gate 17, Bergen, Norway. Email: regine.paul@uib.no

well as "dissuasive" penalties for non-compliance. Understanding what drives the EU's regulatory approach to AI is potentially crucial not just for Europeans but also for anticipating global regulatory dynamics (e.g., Hoofnagle et al., 2019). The EU's proposed AI regulation will also apply to providers and users of AI systems *in third countries* where their products and uses affect the rights of EU citizens. And experiences from the General Data Protection Regulation (GDPR)—and more recent discussions of the EU's Digital Markets Act and Digital Services Act—suggest Europe might externalize its strict(er) tech rules to other jurisdictions (Bradford, 2020).

Core component of the EU Commission's proposal is a "risk-based" approach to regulation (RBR). Based on an assessment of the probability and impact of potentially adverse effects, RBR proposes to develop proportionate forms of interventions—from standards to enforcement—as differentiated by risk level. Along those lines, the draft regulation suggests a nominally risk-based differentiation of regulatory measures to address "the risks associated with specific AI applications in a proportionate manner and of promoting the uptake of AI" (European Commission, 2021c, p. 6). The regulation distinguishes between unacceptable, high, low, and no risk AI systems and targets these risk spheres with a de-escalating set on interventions, from complete ban to laissez-faire (more details in Section 4). It puts a special regulatory focus on "high-risk" AI systems, mentioned 286 times and treated in half of all articles (Title III; art. 6–51). When developing its regulation proposal, the Commission had also discussed alternative policy options, including a voluntary labeling scheme, a sectoral ad-hoc approach, and a horizontal approach with mandatory requirements on all AI systems (European Commission, 2021c, p. 10). In a familiar-sounding narrative, the Commission justifies its eventual choice of a "proportionate risk-based approach" focusing on high-risk AI systems as "most effective" in enabling technology uptake in Europe all while combating some major harmful effects (European Commission, 2021c, p. 11).

This regulatory choice might seem neither new nor particularly worthy of scrutiny at first sight. The EU Commission adopted RBR already in 1997 in response to crises in the consumer health and food safety domain. Numerous other domains have introduced RBR since, including air quality, counter-terrorism, flood prevention, macro-prudential regulation, migration and border control, medical devices, or pharmaceuticals (cf. review in Paul, 2021). And yet, the regulatory fixation on the risk heuristic in the AI regulatory domain requires exploration *beyond* existing accounts of RBR in Europe. Only superficially does the Commission's regulatory proposal meet expectations of rational problem-solving in parts of the RBR literature (c.f. critical debates in this journal: Borraz et al., 2022; Rothstein et al., 2013). Those involved in drafting the Act acknowledge that AI-related risk analysis "is not systematic in any scientific way" (this foreshadows my original interview data, cf. methods in Section 3) and has instead involved high-level political choices about unacceptable use cases for AI as well as "rule-of-thumb" categorizations of high-risk AI systems. Why bother framing this approach as so centrally "risk-based" nonetheless? Prominent alternative explanations point to the symbolic appeal of risk heuristics as a *rhetoric* in organizational legitimacy-seeking (e.g., Black, 2010; Rothstein et al., 2006). But since the newly emergent regulatory domain lacks the organizational underbelly to render institutional risk management strategies plausible or even necessary at this stage, such more constructivist RBR accounts also require conceptual expansion (details in section 2).

Instead, and as the exemplary epigraphs highlight, EU Commission narratives are primarily concerned with how AI regulation can help bring into being "the world we want" by "carving out a distinctive trademark for Europe and its industry" in the AI domain. The Commission's articulation of the regulatory agenda seems more akin then to constructivist accounts of AI policies as "performative politics" which seek to "talk[ing] into being" envisaged techno-economic and politico-economic futures (Bareis & Katzenbach, 2022; also see: Krarup & Horst, 2023). They also resonate with a larger body of work on how visions of economic competitiveness and national sovereignty have shaped the EU's emerging technologies regulation, from cybersecurity to nanotechnology (Farrand & Carrapico, 2022; Mügge, 2023; Ulnicane, 2022; Wullweber, 2015). This paper elaborates critical political economy accounts of tech regulation further conceptually and empirically to grasp the politico-economic work done by the risk heuristic. *How, it asks in an explorative design, does the Commission's vision of a globally competitive common European AI market interact with its "risk-based" approach?*

In response to this query, I re-conceptualize and empirically explore risk-based AI regulation from a Cultural Political Economy (CPE) lens, based on individual and joint work of Ngai-Ling Sum and Bob Jessop (most importantly: Sum & Jessop, 2013). This sheds light on how powerful agents' (here: the EU Commission's) *strategically selective* pursuit of economic projects shapes the discursive as well as more hard-wired structuration of a novel regulatory domain. In particular, CPE helps expose the role of the risk heuristic in European AI regulation

in inscribing selective representations of the envisioned future AI world into regulation in ways that foreground Brussels' politico-economic agenda while coated in the guise of scientific objectivity (details in Section 3). More specifically, a CPE take on European AI regulation unravels how the semiotic framing of AI as *a matter of risk* folds into a seemingly clear-cut and apolitically conceived regulatory space which constructs a trademark of European "cutting-edge AI that can be trusted throughout the world" both discursively and through hard-wired regulatory interventions, including outright bans. Eventually, the risk heuristic serves as an epistemic tool—a meta-cognitive device which mediates how regulators think about the phenomenon to be regulated and how they structure their responses (cf. Tang, 2020)—that enables the Commission to fashion a rights-based *and* globally competitive European AI market beyond mere vision.

The paper proceeds as follows. In Section 2, I problematize existing rational choice conceptualizations of risk-based regulation (both in general and for the specific case of AI technology) and discuss how a European political economy perspective and constructivist analyses of competitiveness discourses in tech regulation can broaden our analytical gaze. To better capture articulations of RBR vis-à-vis strong concerns over global economic competitiveness in regulatory domains, Section 3 conceptualizes RBR within a CPE framework (Sum & Jessop, 2013) and details related methodical implications and choices. I thus provide a novel conceptualization of RBR as an epistemic tool for *fashioning*, that is: *narrating and regulating into being*,[2] a vision of a rights-based and competitive European common AI market. Section 4 applies this analytical framework to empirical material, providing an extensive (mainly) qualitative interpretive analysis of 11 policy documents, one transcript of a formal presentation of the regulatory proposal by a Commission official, and 10 semi-structured interviews with Commission officials and regulatory advisors conducted between May and December 2022 (details in Section 3). Section 5 summarizes findings and discusses the wider relevance of exploring the Cultural Political Economy of risk-based AI regulation.

2. AI technology regulation and the European common market: What is *risk* got to do with it?

To understand why and how risk-based AI regulation appeals to the EU Commission, this section reviews standard rational choice accounts of RBR in emerging technology regulation. I problematize such accounts based on more constructivist and critical political economy work, suggesting that explorations of risk-based AI regulation require firmer grounding in the distinct politico-economic environment in which the EU Commission seeks to manage global competitiveness pressures through narrating and enacting its vision of the common market. Rather than treating risk analysis as a neutral epistemic device to rationalize regulatory interventions when confronted with the uncertainties of emerging technologies, we must analyze the political role it plays in structuring novel regulatory domains alongside key actors' visions of future markets.

2.1. Problematizing rational choice accounts of risk analysis in technology regulation

Scholars of emerging technology regulation duly consider the politics of regulation—for example, by pointing to the winners and losers of tech regulation, the ways in which regulators navigate uncertainty, or asymmetric processes of regulatory competition (Justo-Hanani, 2022; Taeihagh et al., 2021). However, such political interpretations do not usually extend to conceptualizations of how and why risk analysis matters so centrally in emerging tech regulation. So far, risk analysis counts as a tool for rational problem-solving when regulators face, as they usually do, trade-offs between tech innovation and risk mitigation (Jones, 2015; Krafft et al., 2022; Taeihagh, 2021; Taeihagh et al., 2021). With a special focus on automation technology, for instance, a recent study proposes that automated decision systems—also those drawing on AI and machine learning—"can be embedded in very different settings and vary widely in terms of their purposes as well as the decision consequences and the risks involved" so that "a higher risk … warrants greater regulatory efforts for ensuring algorithmic accountability" (Krafft et al., 2022, pp. 2, 12; Jones, 2015 develops a similarly "proportionate" framework based on "fair automation principles"). This assumes regulators have an overriding *rational* desire to forge an *optimal balance* between benefits and risks of deploying new technologies to avoid over- and under-regulation (Taeihagh et al., 2021, pp. 2–3; highlight added) and that risk analysis offers a neutral scientific epistemology to enable rational choice.

These readings conceptualize risk regulation as evidence-based responses to market failure (cf. the seminal summary of a large debate in Hood et al., 2001). As striving for absolute safety would require a disproportionate use of resources, RBR helps address the gravest risks and keep compliance costs low otherwise (cf. Majone, 2010; Sunstein, 2009). The notion of proportionality renders RBR as a seeming conciliator between ethical and safety-related considerations in technology development and hopes for innovation and growth. This line of reasoning also features in the OECD's generic risk regulation framework (2010), and it has already left an imprint on tech regulation. A policy review by Ernest & Young, written jointly with the OECD, specifies the benefits of RBR as follows: "By adopting a proportionate approach where the complexity of regulatory compliance depends on the risk that the AI system poses, policy makers can fulfil their duty to safeguard without unduly impeding the benefits that AI can bring to society" (Etziani et al., 2021, p. 4).

At first glance, this is exactly what the EU Commission has proposed for AI regulation: the differentiation of risk levels and proportionate interventions seems to meet the Commission's "twin considerations of value capture and ethics" (Larsson et al., 2020, p. 15). Through banning ethically unacceptable AI use cases, mitigating the risks of applications which cause considerable concern, and adopting a laissez-faire approach to areas of little concern, the Commission can seemingly wed its innovation and competitiveness agenda with safety and ethical goals. And yet, an abductive iteration between document and interview data and theoretical propositions (see methods) highlights the need to question this explanation (at least) for the case of risk-based AI regulation in the EU:

- Firstly, at a more empirical level, Brussels' rather superficial AI risk assessment methodology challenges the view that regulators are driven by rational problem-solving orientations[3] (cf. wider discussion of such orientations and how to measure them in Paul, 2021). The AI regulation introduces neither comprehensive risk analysis methodology nor does it create robust independent risk assessment units comparable to, for example, the European Food Safety Authority or the European Medicines Agency. Commission officials state that the delineation of risk levels "has not been systematic in any scientific way" (IP-02).
- Secondly, and more epistemologically speaking, scholars of tech regulation see us at an early stage of an AI "regulatory life cycle" and observe a "newly born" regulatory regime with neither established regulators, nor rules—some loose ethical guidelines apart—nor stable relationships with industry (Taeihagh et al., 2021, p. 5). This stage of the cycle usually involves unconsolidated and ad-hoc forms of providing and using expertise, expectations, procedures, and rules. Insiders propose that the epistemological limits to understanding AI and its adverse impacts on individuals and society are such that the EU's definition of risks focuses rather loosely on "where we already see problems" (IP-04). The recent controversy around risks associated with generative general-purpose AI like Chat-GPT underlines that precarious epistemology.
- Thirdly, and moving to ontological criticism of positivist RBR accounts, a growing body of sociological-institutionalist scholarship has argued that risk analysis should not be conceptualized as an objective scientific tool independent from context in the first place. Instead, it is an epistemic tool which is articulated and wielded in context-specific ways in regulators' efforts to structure the complexity and uncertainty of their environment and render it manageable in line with their own identity, norms, and goals (Black, 2010; Borraz et al., 2022; Rothstein et al., 2006).

This latter body of work counters rational choice interpretations of risk analysis from a constructivist position. Scholars in this tradition conceptualize risk heuristics as epistemological constructs which become part of a regulatory *rhetoric* meant to carve out and defend regulators' identity, mandate, activities, and decisionmaking processes against others' expectations and blame for adverse outcomes (Black, 2010). Such *institutional risk management* (Rothstein et al., 2006) explains, for example, why risk-based food safety inspections emerged in the EU in the early 2000s after the highly scandalized outbreaks of BSE (bovine spongiform encephalopathy or commonly "mad cow disease") to counter low consumer trust in the regulatory system; why European countries introduction of private risk management duties for homeowners to lower public expectation of state compensation for flood-related damages (Paul, 2021); or how the Commission has tried to increase authority for supranational decisionmaking in member-state dominated border control (Fjørtoft, 2020; Paul, 2017). Importantly, such risk-based ammunition works precisely because risk analysis preserves its "evidence-based" aura and credibility even where scientific rigor is lacking (Stone, 2012; for a similar discussion on the effectiveness of flawed objectivity claims in cost–benefit analysis see: Porter, 1995).

These debates provide useful stepping-stones for rethinking RBR from a more constructivist and context-sensitive research program. Alas, in our case of a "newly born" regulatory regime they do not reach far enough: in the absence of (yet to be) established mandates and institutions for regulating AI and justifying the appropriateness of related decisionmaking, risk-based institutional risk management lacks its usual target. At the same time, a meso-level analytical focus on individual regulatory agencies' risk-based ammunition strategies cannot capture how wider socio-economic formations, including global capitalism, shape articulations of RBR. In the next section, I discuss how a closer reading of critical political economy accounts, both in scholarship on emerging technology regulation and European integration more generally speaking, can enrich existing understandings of risk-based AI regulation.

2.2. AI regulation vis-à-vis global competitiveness and the European common market

I cannot comprehensively discuss the global politico-economic context of AI regulation here; but point to "moligopolic" structures, where a few big tech players in the United States and China dominate the global AI market (Petit, 2020; Zuboff, 2019). In an intensifying global competition (Smuha, 2021, describes this "race to AI" in more depth), Europe has been lagging behind. The Commission (2020, p. 5) remains concerned that "investment levels for AI in the Union are low and fragmented, relative to … the US and China," and worried about how brain drain to Silicon Valley hampers the creation of a "competitive environment" in Europe (Commission, 2020, p. 12). An earlier document detailed: "Europe is behind in private investments in AI which totaled around EUR 2.4-3.2 billion in 2016, compared with EUR 6.5-9.7 billion in Asia and EUR 12.1-18.6 billion in North America" (European Commission, 2018a, p. 5). In the same passage, the Commission situates its regulatory ambition amidst a clear economic imperative ("European industry cannot miss the train") and makes distinct causal assumptions about its future AI competitiveness: "for the EU to be competitive" it needs "to ensure the take-up of AI technology across its economy." To navigate this scenario of global competitiveness where Europe crawls behind, the Commission envisions the European common market as the provider of specifically "ethical" and "trustworthy" AI products which freely circulate in Europe and are sought after globally (see the paper's epigraphs). As the *Coordinated Plan on AI* suggests: "Spearheading the [AI] ethics agenda, while fostering innovation, has the potential to become a *competitive advantage* for European businesses on the *global marketplace*" (European Commission, 2018b, p. 17; highlight added).

Several works have conceptualized the role of global economic competition as a key context for how EU (as well as other) regulators' frame AI technology regulation (without illuminating the specific part played by the dominant risk heuristic though). Ulnicane (2022), for example, analyzed 49 policy documents from across the world, finding that an "economic growth and competitiveness frame," inherited from previous "races" in emergent technology development, strongly shapes policies (with societal concerns and Sustainable Development Goals added only more recently and more marginally to the agenda). Inspired by work on socio-technological imaginaries, Bareis and Katzenbach (2022) conceptualize AI policies as "performative politics" which seek to "talk[ing] into being" envisaged techno-economic and politico-economic futures. They do not analyze the EU case as their focus is on China, Germany, France, and the United States instead. Krarup and Horst (2023) provide a closer analysis of EU AI policies (excluding the regulatory proposal from their corpus though) and propose that visions of the EU's single market are both "engine and structuring principle" (p. 1). The authors' mixed-methods analysis highlights how "deep-seated problems of market integration in the EU," including the tensions and contradictions that stem from casting the European integration project as both market-making and rights-based, have shaped policy framings (p. 3). More generally, these works also confirm findings on the competitiveness orientations of EU regulation in other emerging technologies, for example in cybersecurity and nanotechnology (e.g., Farrand & Carrapico, 2022; Wullweber, 2015).

At a conceptual level, diagnoses of the strong competitiveness undercurrents in European AI (and other technology) regulation also resonate with wider (and much older) debates in critical political economy. The first is inspired by Post-Marxist and Neo-Gramscian political economy and concerns the incorporation of economic competitiveness logics in regulation with the concept of "competition states" (Cerny, 1997; Cerny, 2010; Jessop, 2002). Rather than being one-sidedly captured by business interests, states try to manage economic globalization by pro-actively transforming into "competition states" (Cerny, 1997): these adopt policies and regulations to create business- and investment-friendly

conditions which can boost the national or regional economy's competitiveness on global markets alongside desired visions of its role on global markets.

The second source of inspiration concerns work which conceptualizes EU regulatory projects as deeply intertwined with the EU's identity as a tension-load "market-integrating undertaking" (Damro, 2012, p. 685; cf. debate in Krarup & Horst, 2023 on AI regulation as a case of "single market-making"). This suggests that the Commission and member states seek to become more competitive on global (product) markets *all while* also promoting a vision of a rights-based Europe which protects its firms and citizens against the detrimental effects of economic globalization. Political economist and state theorist Bob Jessop (2002) claimed already two decades ago that the increase of post-national regulatory powers has emerged from a distinct *European* project of advancing the global competitiveness of the common market all while mitigating the socially adverse effects of global capitalism. Scholars of European economic integration argue that this dual focus has fostered a Janus-faced regulatory approach (Finger, 2011; Rosamond, 2018). Internally, EU regulation seeks to further the common market in ways that protect its citizens against the most adverse effects of globalization and its firms against unfair competition based on social, environmental, or other forms of dumping through an extensive set of fundamental and individual rights and a dense regulatory portfolio. Externally, it promotes its regulatory standards, seeks to create a (high-standard) level playing field for national firms, and advances the EU's global competitiveness especially in high-quality product market segments (cf. on that external dimension: Bradford, 2015).

Importantly, more recent additions to Damro's (2012) "market power Europe" approach propose that the EU's focus on collective norms and rights foundations does not necessarily contradict its economic strategizing. The discursive articulation of its moral ambitions has rather become a chief element in how the EU conceives of its global regulatory power, in how it frames the distinct quality of products "made in Europe" on global markets, but also in how it seeks to enact such visions through its own regulatory interventions in global markets. In other words: the EU's regulatory projects draw both on a discursive framing of its identity as a rights-based common market and on the more material establishment of specific regulatory standards in ways that can help advance the common market's global competitiveness. A prominent—and already well-researched—example of this dual strategy of *narrating and regulating into being* a rights-based and *thereby* globally competitive common market is the *European Green Deal* where the EU "seeks to be a first mover in global regulatory competition around the green transition, but also wants to lead by example and diffuse its environmental norms" (Eckert, 2021, p. 82).

3. Analytical approach and methods: A Cultural Political Economy of risk-based (AI) regulation

The discussions in the previous section culminate in the paper's wider conceptual–analytical proposition: to account for the Commission's, otherwise empirically implausible, fixation on risk-based AI regulation, we must explore the risk heuristic's specific articulation in interaction with visions of the EU as a rights-based and globally competitive common market, as well as Brussels' more material attempts to structure global AI competition in its favor. In this section, I introduce Cultural Political Economy (CPE) as a conceptual–analytical framework for systematizing such inquiry and detail-related methodical choices, including the paper's abductive research logics.

CPE was developed by Ngai-Ling Sum and Bob Jessop (2013) to integrate poststructuralist discourse and historical materialist analyses of political economy for the exploration of how *both* ideational and more materially hard-wired strategies interact in socio-economic regulation. As it allows for an integrated analysis of regulatory discourses and more material decisionmaking, the CPE lens can contribute to regulation scholarship more widely (see discussion). In the case at hand, the CPE analysis focuses on how the Commission, as chief initiator of emerging technology regulation, employs the risk heuristic to *enact* (or "fashion" as I call it here to capture the constructivist act of narrating and regulating into being an envisaged common AI market) the EU's vision of a seamless, rights-oriented, and globally competitive common market that can protect its standards and norms internally and promote them externally.[4]

CPE conceives of two basic modes of complexity reduction for actors facing an incomprehensibly complex and uncertain world[5]: semiosis and structuration (Sum & Jessop, 2013, pp. 155–164; cf. Jessop, 2009). *Semiosis*, or meaning- and sense-making, concerns discursive strategies of accentuating some selected aspects of the social world, while ignoring or silencing others, or combining and recombining sets of meanings. This resonates with discussions of "frames" and "narratives" in interpretive accounts of (AI) policy (cf. Bareis & Katzenbach, 2022; Ulnicane, 2022).

Structuration concerns the "extra-semiotic" or "material" aspects of complexity reduction through institutionalizations of specific forms of social relations, such as relatively stable organizations, or transaction models. These establish a

> "mix of constrained opportunities, recursivity, redundancy and flexibility" and thus forge "a complex assemblage of asymmetrical opportunities for social action, privileging some actors over others, some identities over others, some ideal and material interests over others, some spatio-temporal horizons of action over others, some coalition possibilities over others, some strategies over others."   (Sum & Jessop, 2013, p. 150)

In regulation, such structuration includes the hard-wiring of opportunities and constraints within the socioeconomic formation through rules and compliance requirements (between the extreme poles of prohibition and laissez-faire), the ascription of regulatory mandates and responsibilities, and the device of sanctions and penalties for non-compliant behavior.

Semiosis and structuration do neither operate in a vacuum nor are they structurally determined: CPE applies an evolutionary understanding of socio-economic governance as the "variation, selection, and retention" of semiotic and extra-semiotic (i.e., more structural and ontologically material) elements from an irreducibly complex array of potential discourses and structural forms. Actors navigate the complexity of the social world in "strategically selective" ways; they thus produce "sedimented meaning" and "structured complexity"; and they hence also shape the future construal and construction of the social world (cf. Sum & Jessop, 2013, p. 192, fig. 4.3). But not even the most powerful actors can solve the inherent paradoxes and contradictions of the capitalist socio-economic formation for good. Nor can they fully control how complexity and uncertainty are being processed, leaving plentiful scope for the continuous re-politicization of regulatory projects (for a more recent comprehensive expression of the dialectic dynamics of capitalist reproduction, see Peck, 2023).

While I follow the authors' cue and distinguish between semiosis and structuration as interrelated basic modes of complexity reduction from the first part of their book (Sum & Jessop, 2013), a later chapter involves a typology of four "selectivities" or "strategic modes of selectivity." Based on a more detailed discussion of the works of Foucault and Gramsci, the authors separate structural, discursive, technological, and agential selectivity (cf. pp. 218f, tab. 5.1). Agential selectivity highlights that some actors have a privileged role in meaning-making and structuration, as would be the case of the EU Commission in the European regulatory setting (but see discussion for future research across a wider set of actors). Technological selectivity concerns "assemblages of knowledge," "rationalities," and the ways in which specific epistemologies support discursive and structural choices. I consider the dominance of the risk heuristic in European regulation a crucial part of such technological selectivity: its appearance and justification as a scientific epistemic tool helps rationalize highly political regulatory decisions as objective and trustworthy. Eventually, my specific CPE take on European AI regulation inquires:

- how the Commission—as a powerful regulatory agent—uses the risk heuristic as an epistemic tool in its semiotic and regulatory differentiation of the AI domain;
- how such risk-based differentiation draws on earlier discursive and regulatory articulations of the common market; and
- how it ultimately fashions the common AI market.

Implementing this conceptual–analytical focus, this paper explores the EU's risk-based AI regulation with a stepwise empirical analysis (see summary of analysis in Table 1):

1 To explore *semiosis*, I reconstruct clusters of meaning behind the risk heuristic, and in particular the three risk levels devised by the EU Commission: these include articulations of the risk lens in relation to the wider regulatory context, regulatory goals, and (partial) visions of the common AI market.
2 To explore *structuration*, I reconstruct how the semiotic differentiation of risk levels, and related visions of the common AI market, interacts with Brussels' division of concrete regulatory instruments and forms of interaction between state, industry, and consumers.
3 To firmly situate the analysis in the macro-level *politico-economic context* of AI regulation, I explore the individual and joint work done by the specific semiotic and structural articulation of three AI risk spheres in the Commission's pursuit of a common European AI market.

Table 1 Overview of three risk spheres in EU artificial intelligence (AI) regulation

Ascribed risk level	Unacceptable risk	High risk	Low/no risk
Example of use cases	Social scoring systems; live remote biometric identification in public places	Predictive policing; law enforcement; migration and border control; education; recruitment	Chatbots; emotion recognition tools; smart spam filters
Regulatory context	Increase of undemocratic uses of, as well as severe rights violations with, AI tech globally	EU laggard in tech development; EU AI market fragmented; lack of risk knowledge; severe rights violations	EU laggard in tech development; EU AI market fragmented; strong SME segment
Regulatory goal	Protect EU core values (fundamental rights and democracy)	Protect fundamental and individual rights; enable innovation; create single market	Create a seamless single market for AI; enable innovation (protect consumers)
Vision of the common European AI market	European AI uses and development committed to fundamental rights and democracy; clearly different from (morally superior to) many competing AI system providers on global market; promoting "human-centric" AI globally	Common European AI market provides legal certainty, reduces liability for potential harms, counters market fragmentation; high-risk regulation enables "safe" innovation including in EU's strong public sector	Common European AI market as seamless and frictionless; EU as investor, R&D funder and aid for competitiveness of SMEs; firms in Europe voluntarily develop AI in line with EU principles and prosper in global competition
Role of state regulators	Strict watchdog outlawing abusive uses	Risk mitigator; testing ground for compliant AI uses in public sector	Enabler of voluntarism and innovation; key investor
Regulatory instruments	Prohibitions, bans, dissuasive fees	Conformity assessment based on various risk mitigation requirements; CE certification; oversight; regulatory sandboxes	Voluntary certification; codes of conducts; information duties; public investment
Role of risk sphere(s) in enacting common AI market	Specifying and banning "grossly unacceptable" use cases shows commitment to core values > legitimizes most other AI uses	Codification of balance between risks and benefits creates trust among consumers/users > enables "trustworthy" innovation	Light regulation of most use cases normalizes economic freedom in AI innovation as consumer market issue > free circulation of AI systems

The paper adopts an interpretivist research design to explore a case of regulation in its proper context (Schwartz-Shea & Yanow, 2012), empirically building mainly on qualitative[6] analysis of eleven policy documents, a transcript of an official presentation of the draft regulation, and ten interviews with experts involved in preparing the EU's AI regulation (see Appendix I). The choice of these various data sources ensures high levels of intertextuality and exposure to the empirical phenomenon, keeping with the quality criteria of interpretive research (Schwartz-Shea & Yanow, 2012). To reconstruct the Commission's regulatory strategy as comprehensively as possible, documents represent a full sample of Commission policy documents, strategy papers, but also policy recommendations by the consulting High Level Expert Group (AIHLEG). I targeted interviewees with a prominent role in drafting the EU Commission's AI regulation either in relevant Directorates General or as rapporteur in the AIHLEG. Document analysis enabled me to identify the three risk spheres, their specific discursive articulation, and more structural regulatory choices linked to risk-based differentiation. The analysis of expert interviews then allowed me to reconstruct actors' more subtle accounts of how the risk differentiation came about, how actors made sense of but also struggled over it, how they sought to structurally act upon these risk spheres, and how such risk-based differentiation of the AI regulatory space interacts with the EU's wider ambition to create a

globally competitive common European AI market. As the analysis will also highlight, the European Parliament has already contested some of these differentiations.

I coded primary data sources with qualitative data analysis software (MaxQDA). I first coded segments which articulate "low/minimal risk," "high risk," and "unacceptable risk" scenarios. I then coded for descriptions of "context conditions" of AI regulation, "visions of the common AI market" and specific "regulatory choices" and generated several subcodes inductively (see Appendix). The cross-coding of risk levels with both the Commission's visions of the common AI market *and* concrete regulatory choices helped reconstruct the semiotic and structural dimensions of risk-based AI regulation and its foundations in European political economy. The analysis builds on both quantitative analyses of the proximity of codes in and across coded segments (Table A3 in Appendix II) and a more limited qualitative analysis of argumentative threads within coded segments.

4. How risk-based regulation fashions a future common European AI market

This section explores how the EU Commission differentiates the AI domain by the help of the risk heuristic and how this work of carving out a threefold regulatory space interacts with its vision of the future common European AI market. The regulatory proposal distinguishes *three* risk spheres and justifies this with the wish to introduce "proportionate" rules:

> In order to introduce a proportionate and effective set of binding rules for AI systems, a clearly defined risk-based approach should be followed. That approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate. It is therefore necessary to prohibit certain artificial intelligence practices [unacceptable AI systems], to lay down requirements for high-risk AI systems and obligations for the operators, and to lay down transparency obligations for certain [low risk] AI systems. (European Commission, 2021c, p. 22; information in brackets added)

In the subsequent analysis, I inductively reconstruct these three risk spheres beyond the nominally clear-cut "proportionate" problem-solving rhetoric. This exposes how they vary in their articulations of several aspects: the relevant regulatory context, regulatory goals, the assumed epistemology of risk in the respective risk sphere, visions about the role of this risk sphere in the common European AI market, but also more *structural decisions* about the role of state regulators, regulatory instruments, and specific use cases. Table 1 summarizes the analysis.

4.1. "That's not our vision of Europe"[7]: The sphere of unacceptable AI systems

The Commission's AI regulation proposal, in Title II, outcasts a few AI systems as "contravening Union values" (European Commission, 2021c, p. 13). These values include "respect for human dignity, freedom, equality, democracy and the rule of law and Union fundamental rights, including the right to non-discrimination, data protection and privacy and the rights of the child" (European Commission, 2021c, p. 22). The protection of core values and fundamental rights features in more than half of all coded segments on "unacceptable risks" (Table A3), and sidelines otherwise pronounced discussions regulatory goals such as "innovation," "legal certainty," or the "single AI market."

Applications in mind involve the "AI driven 'optimisation' of social processes based on social scoring systems" which, according to the Commission, "violate(s) the basic idea of equality and freedom in the same way caste systems do" (European Commission and European Group on Ethics in Science and New Technologies, 2018, p. 11). These would represent an "attack on democratic systems" (European Commission and European Group on Ethics in Science and New Technologies, 2018, p. 11) and "go against our fundamental values" (Vestager, 2021, p. 2). In addition to social scoring, and following the same values-based justification, "subliminal techniques with a potential to cause physical or psychological harm," including among vulnerable groups such as children, are also part of the unacceptable risk sphere (Vestager, 2021, p. 2). A similar line of argument arises for the use of remote live biometric identification in public places. While the Commission supports the limited use of biometric identification in concrete law enforcement practices, for example in border control, it argues that "there is no room for mass surveillance in our society" (Vestager, 2021, p. 2).

9

The Commission connects its discursive articulation of values-based red lines in any marketization of AI products with a concrete regulatory choice: unacceptable AI systems are prohibited with only very limited exceptions for biometric identification and there will be "dissuasive" fines for violations of these rules (30 billion Euros or 6% of their annual yearly global turnover) (European Commission, 2021c, p. 82).

Interviewees report that the discursive and regulatory treatment of some AI uses as unacceptably risky has been "a purely political decision" (IP-02). There was no risk assessment, they argue, but an arbitrary value statement.[8] As a Commission official reflects:

> *The prohibitions - if you take remote biometric identification by law enforcement, this is of course a purely political decision. How do we deal with mass surveillance by law enforcement? … there's no evidence that suggests this is always good or always bad, you know, this is sort of a political judgment: how do we want to deal with it? Or … social scoring the Chinese way … where we say, 'that's not our vision of Europe, we don't want social scoring system à la China, we don't want mass surveillance.' These are pure political decisions.*   (IP-02)

Some thus interpret the Commission's structural intervention of banning some AI uses as an attempt to frame the EU as a "project based on human rights" (IP-10), as a union building on fundamental rights and democratic values which has the "courage to ban certain uses" of AI (IP-09). Others highlight that this is "partly a symbolic gesture to say, you know, there are things that we are ready to forbid" (IP-04). This almost sacrificial language of foregoing innovation for more noble civilian aims must be set against an otherwise dominant rhetoric of wanting to enable AI deployment by default (and, for that matter, also against the decision not to touch autonomous lethal weapon systems in its regulation). Indeed, the Commission emphasizes the exceptional character and limited empirical scope of this sphere as "the top of the iceberg," "very narrow," and as "something which we find *outrageously* abusive" (Gross, 2021, transcript).

Interview data suggest that the targets of this discursive and regulatory out-lawing of some AI applications are mainly external. Even if the proposal does not mention any jurisdiction explicitly, spelling out "things the EU would never do" (IP-02) seeks to discredit existing AI applications elsewhere. It sets a normative counterpoint to developments such as the "rushed deployment of algorithms in government but also by platforms" in the United States (IP-09) or "undemocratic uses" of AI technology in the Chinese social credit system and other forms of mass surveillance (IP-10, similar statement also by IP-02, IP-07). Some see prohibitions as a way of signaling to the world that "China can do what it wants, but the law of the land in Europe is we don't do social scoring. … It is clear that we are different." (IP-05); or as inspired by "the need to make a clear distinction … we're not going to become China. We won't do social scoring … for Europe to show how different it is from the Chinese approach" (IP-07).

Being strict on AI systems which grossly violate existing EU values may also serve as a prerequisite for the EU's continued legitimacy as a rights-based regulator, both internally and externally. Some argue that the discursive articulation and regulatory structuration of unacceptable AI risks also reflects Europe's lasting commitment to the GDPR with its individual rights focus both toward its own citizens and as a signal to global markets (IP-09, IP-10). By narrating its own AI regulation approach as committed to democratic values and fundamental rights—to the point of sacrificing innovation and economic growth—Brussels seeks to render the competition of value systems as part and parcel of the global competition on AI in more material ways, too. This thread is illustrated in this longer statement by a Commission adviser:

> *Here we're talking about … the EU's project based on human rights. And this is typical of Europe, because, at the end of the day, it's true that many countries have signed the relevant [human rights] declarations, but who is really fighting for it, and really applying and implementing it? Well, it is Europe. But of course … there is also a lot of work to be done in-house, we should do our own work first, and then go outside to tell others what to do.*   (IP-10)

This line of argument sees the future success of the EU's "tech diplomacy" in global fora such as the Global Partnership for AI—and particularly its efforts to gather a powerful group of "like-minded countries" to counter the increasing Chinese dominance of international standardization bodies—as reliant on its gaining moral credibility by hard-wiring some base norms on AI uses within Europe. A pointed version of this geopolitical argument runs like this: "if we lose the battle around AI standards, basically, we will have to be ready to accept that [country A] will remain a dictatorship forever, and that they will actually increase their capacity to be non-democratic because they use the best AI technologies that [country B] provides them" (IP-10). If the EU constrains the scope

for AI competition to those systems deemed morally acceptable by European standards, both discursively and by constraining some AI development, it may tilt both the geopolitical and the economic tech race in its favor.

4.2. "You know what you have to do … and you have a huge market at your footstep"[9]: The sphere of high-risk AI systems

The discursive articulation and regulatory treatment of "high-risk AI systems" dominates[10] the draft regulation and related documents. Contrary to the explicitly values-grounded articulation of unacceptable risks in AI technology, the Commission engages the language of actuarial risk here: "the extent of the adverse impact caused by the AI system on the fundamental rights protected by the Charter[11] is of particular relevance when classifying an AI system as high-risk" (European Commission, 2021c, p. 25). Pursuing this notion of risk further the regulation suggests any stand-alone AI systems shall be categorized as high-risk "in the light of their intended purpose, [where] they pose a high risk of harm to the health and safety or the fundamental rights of persons, taking into account both the *severity of the possible harm and its probability of occurrence*" (European Commission, 2021c, p. 26, highlight added). The annex to the proposal lists a limited number of "specifically pre-defined areas" where the use of AI systems counts as particularly likely to infringe health and safety or fundamental rights. These contexts include critical infrastructure, law enforcement, migration and border control, employment, or education (European Commission, 2021a).

This nominal expression of an actuarial risk concept contrasts with evidence from interviews. The Commission has not calculated the probability and adverse impact of AI systems on these fundamental rights with any numerical measures or clear indicators. Some AIHLEG members remember that this list "was literally pulled out of the air, based on a focus group" (IP-07), others share that they advised against using this terminology because "risk actually is something you have to evaluate based on concrete data, knowledge, … it must be grounded in real situations, real experiments, real results" (IP-03). Indeed, Commission officials agree that given the lack of data and precedents they "don't really have any … way of calculating probability" where "most of what you're regulating doesn't exist yet" (IP-04). This suggests that the Commission does not perceive its turn to RBR here as a case of rational probabilistic problem-solving. Instead, interviewees report on a "pragmatic" (IP-05) rule of thumb approach to defining high-risk AI systems as "the issues that stick out"—including through a White Paper consultation process.

Such lack of scientific underpinning does not come as a surprise neither to constructivist RBR scholars nor to critical political economists studying EU regulation, of course. We rather need to explore the political workings of the risk heuristic: as an epistemic device which helps create the *perception of balance* between fundamental rights protection and AI-based growth and competitiveness. The hierarchy between these two goals seems to tip in favor of innovation: through its *narrating* as an act of actuarial analysis of the decision to treat a few hand-selected AI uses as "high risk," the EU can discursively claim to tackle chief concerns about the emerging technology all while structurally maintaining AI-driven innovation as a default regulatory option. This works irrespectively from a lack of scientific method or transparent participatory processes underpinning risk analysis (several interviewees criticized the consultation process: IP-01, IP-06, and IP-07).

This default is omnipresent when interviewees describe high-risk AI systems as a "limited" space. An official recalls that the Commission "wanted to avoid at all costs being seen as an over-regulator. And so, I think, the compromise was struck that the AI Act will only cover a tiny subset of all AI in the single market" (IP-09). The high-risk terminology thus sends a message to competitors, industry, and potential users of AI systems that "it is not because we regulate AI that we don't like it" (IP-04). An EU official reassured industry representatives that these uses are "not forbidden but need to be carefully checked before they enter the market" (Gross, 2021, transcript). Industry reportedly pushed the Commission to carve out a small and clearly demarcated space for regulatory intervention so as not "to stifle innovation" (IP-01, IP-06, and IP-07). One member of the HLEG (IP-01) argues, for example, that "the whole thing was an industry initiative from the beginning. They saw we're losing; we're losing against the Silicon Valley and China. And that was the motive behind it [focusing on a nominally risk-based approach]: to have trustworthy AI from Europe." Indeed, the Commission has been adamant about the need to clearly define the regulatory space of "high-risk AI systems" as "limited to those that have a *significant harmful impact* on the health, safety and fundamental rights of persons in the Union" so that "such

limitation minimises any potential restriction to international trade" (European Commission, 2021c, p. 25, high-light added). In the words of two Commission officials: "we've tried to make sure that most of the AI development is not touched by what we're doing, we're only focusing on a very small number of high-risk use cases" (IP-04), rendering the regulatory approach "as innovation-compatible as we can" (IP-02).

The AI regulation acts upon this semiotically constructed "balance" between risk mitigation and innovation goals on the common European AI market through several forms of structuration. As a bottom line, it allows for the deployment and use of high-risk AI systems in the EU but makes them subject to several mandatory requirements. Providers and users of high-risk AI systems must comply with specific documentation duties, risk, and quality management obligations, they must ensure by design that human oversight and control of the system is possible, they must undergo a conformity assessment, and they must engage in post-market monitoring of the risks stemming from (uses of) AI systems. Fines for non-compliance are lower than for the unacceptable risk realm (20 billion Euros or 4% of the total global turnover). Three regulatory instruments are worth mentioning as structuration devices in this risk sphere:

- Firstly, those wanting to deploy a high-risk AI system have to conduct an *ex-ante conformity assessment*. The assessment shall be based on "a full, effective and properly documented ex-ante compliance with all requirements of the regulation and compliance with robust quality and risk management systems and post-market monitoring." The choice of self-assessment—rather than third party of statutory assessment—for most high-risk AI systems speaks to the Commission's desire not to impose overly high compliance costs and to enable innovation, but it also recognizes the lack of public auditing expertise in this "early phase of the regulatory intervention" (European Commission, 2021c, p. 15).
- Secondly, all high-risk AI systems which went through conformity assessment and have been registered on a new database can then use the *CE label* "to indicate their conformity with this Regulation so that they can move freely within the internal market" (European Commission, 2021c, p. 34). The Act portrays this declaration of conformity with EU standards as chief marker of "trustworthiness": high-risk AI systems underwent strict assessment in the EU and their adverse effects on core values and fundamental rights are properly mitigated. The CE label lends trust to an AI system in ways which should then enable its free circulation on the common market (European Commission, 2021c, p. 34), raise consumer acceptance, and potentially become a global trademark of "ethical" AI products "made in Europe." One official sees the CE label as a competitiveness booster specifically for European small and medium entreprises (SMEs) to "compete with the bigger ones": "if we now have a label for trustworthiness … every start-up can have this label and they will compete on equal footing" (Gross, 2021, transcript).
- Thirdly, public sector AI innovators can draw on a specific set of *regulatory sandboxes*, an experimental tool first introduced in the fintech sector and meant to create testbeds for innovative projects in emerging regulatory settings (Ranchordas, 2021). In the AI case, public sector organizations can navigate and carve out how compliance *could* be achieved when using high-risk AI systems. Norway and Spain are already experimenting with sandboxes to "connect innovators and regulators and to provide a controlled environment that facilitates the development, testing and validation of innovative AI systems with a view to ensuring compliance with the requirements of the AI Regulation" (European Commission, 2022, p. 3).

The clear semiotic delineation and narrow scoping of high-risk AI systems, paired with a choice of specific regulatory interventions in this realm, engage with visions of the common European AI market in several interconnected ways. Firstly, it relates to a desire to create "legal certainty"[12] for developers and users of the technology as a booster of innovation (European Commission, 2020, p. 19). Several interviewees suggest that "a clear legal framework" was a key demand by an industry keen to avoid uncertainty around their future investments and worried about the "risk that someone might regulate" in the future in unexpected ways (IP-02, similar statements by IP-01, IP-05, and IP-06). Industry representatives in the AIHLEG would have been "keen to avoid too much regulation … or too many rules" and eventually welcomed a focus on high-risk mitigation, but for a clearly curtailed and small number of applications, as beneficial for their innovation and investment strategies (IP-03). One Commission official conveys the assumed benefits for businesses as follows: "you have 450 million consumers that you can serve. Rather than having 27 different regulatory frameworks … you know what you have to do, you do it, you can carry on with it, … you have no legal risk, and you have a huge market at your footstep" (IP-02). In other words: the

clear definition and regulation of the high-risk sphere does not only (or predominantly) reduce risk of rights violations for EU citizens; through addressing an older constitutional concern with fragmentation on the common market (cf. Krarup & Horst, 2023) it also lowers companies' legal (and financial) risks when innovating in this sector.

Secondly, the Commission assumes that its tight regulation of high-risk AI systems—and the simultaneous outlawing of a few extreme cases of unacceptable risks—can *generate higher trust levels* among citizens who "will be more ready to accept AI" so long the regulation offers "something to believe in" (IP-04). The 2020 White Paper already linked citizen-consumer trust and innovation goals in a utilitarian manner: "Building an ecosystem of trust … should give citizens the confidence to take up AI applications and give companies and public organizations the legal certainty to innovate using AI" (European Commission, 2020, p. 4). Another Commission official summarizes the "trust-creating" role of the regulation's high-risk focus as follows: "if people do not trust technology, that will slow down its uptake … it is important that people think: okay, the way in which AI is developed and used in Europe is properly regulated, so we can trust that this is okay" (IP-02). A more critical member of the AIHLEG sees a subordination of rights protection to innovation goals in the Commission's definition of "trustworthy" AI: "The industry is preparing a massive future market. And they cannot ever have it that the general public, who's already nervous about killer robots, loses trust in their products" (IP-01). Regulating high-risk systems in a "trustworthy" manner—including through conformity assessment and CE label—thus legitimizes and enables technological innovation on the European AI market.

Thirdly, the Commission considers citizen trust particularly relevant in public sector AI applications, an area which the AIHLEG has pointed to as a large market segment for the EU to harvest. In its policy recommendations (2019), the group suggested that Europe's "strong public sector" (p. 17) "has a competitive edge to scale Trustworthy AI" (p. 16) and can thus "act as a catalyst for innovation and growth" (p. 10). As violations of fundamental rights are more likely and impactful in public sector AI applications, framing some of these as high-risk and safeguarded experimentation with them in regulatory sandboxes can, hopes the Commission, convince public authorities to procure, co-develop, and deploy them without too much resistance by worried citizens.

More outward-looking, lastly, the Commission perceives the establishment of a single and legally certain European AI market, with clearly banned AIT uses as well as tightly regulated risks, as a prerequisite for boosting the EU's *global* competitiveness in this domain: "Guaranteeing the respect for fundamental values and rights is not only essential in itself, it also facilitates acceptance by the public and increases the *competitive advantage of European AI companies by establishing a brand of human-centric, trustworthy AI* known for ethical and secure products" (European Commission, 2019, p. 8, highlight added). Some thus describe the project of "trustworthy AI from Europe" as "an industry initiative" where "they [European businesses] saw we're losing against Silicon Valley and China" (IP-01), and were the EU would "not be competing in the same way" but would be "creating a *Europe-branded* AI" in which "ethical" and "trustworthy AI" would "by synonymous with *European* AI" (IP-07, highlight added). RBR hence enables the rhetorical and regulatory construction of a competitive space for *ethically checked* European AI products much in line with the EU's earlier strategies to capture global high-quality market segments by strict, and morally underpinned, unilateral regulation (cf. Damro, 2012; Eckert, 2021).

4.3. "We are convinced that most of the AI out there is not risky"[13]: The sphere of low/no-risk AI systems

Brussels' AI regulation articulates most applications as low/no-risk and installs light-touch information requirements plus some incentives for adopting codes of conduct. The Commission narrates otherwise imposed regulatory requirements on AI applications as *exceptional*: "we are convinced that most of the AI out there is not risky. … we believe that 88 to 85 percent of the AI applications or uses out there don't need any regulation, they should just be placed on the market, used, developed …; they don't need restrictions as they're not more dangerous than any other software" (Gross, 2021, transcript). This vision of a European common AI market with freely circulating tech applications and much scope for innovation *relies on* the earlier description—and regulatory inscription—of unacceptable AI system as "the top of the iceberg" and high-risk systems as "a very small number" of cases (IP-04, also IP-02).

While the Commission articulates no potential for harm in this risk sphere, it also re-iterates its overarching vision of a European common market based on consumer protection and rights. The focus on boosting the development and uptake of "low-risk" AI would not generate a rights-free vacuum for innovation (Gross, 2021, transcript). The regulatory sub-differentiation between applications posing *residual* risks (e.g., spam filters) and those

considered *low* risk (chatbots, emotion recognition devices, and deep fakes), for example, helps the Commission specify consumer rights for AI applications: while the former can circulate freely on the single market as long as they do not violate "the legal rules that already exist to protect consumers," the latter come with some information duties "to make crystal-clear to users that they are interacting with a machine" (Vestager, 2021, p. 1). Such light touch regulation should enable EU consumers to make "informed choices or withdraw" from AI technology (European Commission, 2021b). The proposed regulation also points to existing mechanisms for legal redress should rights violations occur in the low/no risk realm. Overall, then, the Commission normalizes low/no-risk AI systems discursively as matters of standard consumer protection which is sufficiently handled through existing regulation. Interestingly, in the original Commission proposal this normalization included chatbots. This changed with the wider application and critical discussion of generative AI technology based on large language models (such as Open AI's ChatGPT). As a response, the European Parliament—in its June 2023 amended regulatory proposal—re-opens the case of whether chatbots should be considered low-risk and requires generative AI to be submitted to review before commercial release.

Leaving aside the GPE debate, overall, the EU's regulatory proposal enacts the discursive construction of low/no-risk AI systems as a matter of free market circulation with a laissez-faire approach that can do without binding requirements (some information duties apart). Commission officials emphasize that the bulk of AI system will not be governed by regulatory requirements, but by the more investment-oriented parts of the EU's policy package containing "everything from skills to funding for research, to infrastructure deployments, to specific sectors" (IP-04). At the same time, the Commission supports self-regulation as part of European firms' successful marketing and competition strategies on the AI market. For example, member states and companies should introduce voluntary codes of conduct "if they want to go beyond legal obligations and … illustrate this" to their customers and users (Gross, 2021, transcript). The White Paper already discussed voluntary CE certification as an economically desirable branding device: firms who adopt codes of conduct in line with the AI regulation's requirements for high-risk systems, even if they do not fall into this risk category, "would then be awarded a quality label for their AI applications" and eventually be able to "signal that their AI-enabled products and services are trustworthy … in compliance with … standardized EU-wide benchmarks" (European Commission, 2020, p. 25). The Commission constructs a link here between the (tighter) requirements for high-risk AI systems, firms' voluntary adoption of these in the low/no-risk realm, the creation of "trust" among consumers and users, and the much-desired increases of AI development and deployment on the common market (European Commission, 2021c, p. 36).

5. Conclusion and discussion

This paper explored the political work done by the omnipresent risk heuristic in the EU's emergent AI technology regulation. It addressed the limited explanatory power and methodological shortcomings of standard conceptualizations of RBR as an objective rational problem-solving tool, connecting instead to constructivist and critical political economy accounts of the performative politics of technology regulation. More specifically, I conceptualized the risk heuristic as an epistemic tool used by the EU Commission to bring into being a globally competitive common European AI market. Within a CPE framework and through an abductive, interpretive, and qualitative research program, I explored how the Commission's vision of a globally competitive common European AI market interacts with articulations of the "risk-based" approach at the semiotic and structural level. The CPE analysis revealed how the discursive differentiation of AI systems into three risk spheres helps constitute a seemingly clear-cut and objective regulatory structure, ranging from prohibitions, via risk mitigation, to laissez-faire. It showed how Brussel's ambition of creating a trademark of European "cutting-edge AI … trusted throughout the world" drives the risk heuristic's ordering work in the emerging regulatory domain. I summarize key findings next, before discussing the paper's wider contributions and limitations.

At its most basic, the research highlighted how the Commission employed the risk heuristic to carve out three much diverse spheres for AI regulation: an outlawed sphere for unacceptably risky AI applications, a tightly regulated but innovation-friendly sphere for high-risk AI systems, and a free market sphere for low/no-risk AI systems. While this differentiation might speak to assumptions of proportionate regulation superficially, neither concern over the limited scientific rigor of risk analysis nor concrete problem-solving orientations were part of the Commission's sense-making of RBR in the AI domain at the time of this research. Instead, justifications for

adopting RBR, for drawing lines between the three risk spheres (which remain politically contested, as Parliament's amendments show), for the regulatory goals to pursue in each sphere and across them, and for the regulatory instruments to deploy, all heavily draw on visions of the common AI market and Brussels' wish to enact this very vision through its discourses and regulatory choices.

Each of the three risk spheres, as the Commission narrates and regulates them, performs a specific role in the fashioning of the European common AI market. The clear delineation and regulatory outlawing of *unacceptable AI risks* portrays the common market as deeply committed to fundamental rights and democratic values. Banning a handful of applications based on deep value conflicts legitimizes all AI applications outside this exceptional space, including the high-risk ones, as trustworthy and worthy of deployment on the common market. Moreover, by defining and prohibiting unacceptably risky AI systems, the Commission sets Europe apart from other large players on the global AI market and unilaterally constrains the AI competition space to those systems deemed morally acceptable by European standards. As for the *high-risk sphere*, its semiotic articulation and regulation serve to provide legal certainty and reduce potential liability risks for harms for those seeking to innovate and deploy AI systems which might violate individual rights. The Commission creates a space for "safe" and tightly checked innovation, including in the public sector, both through the narrow scoping of the high-risk sphere to a limited number of applications and through concrete instruments such as the conformity assessment and CE label. The discursive articulation and regulatory laissez-faire treatment of the *low/no-risk sphere*, lastly, normalizes the free circulation of AI applications on the European market as a default option and treats most risks—especially information asymmetries—as already covered by consumer protection regulation. This focus renders the European common AI market as a seamless and frictionless place for AI innovation which can flourish without intervention precisely because "trustworthiness" has been created in the other two risk spheres. Of course, the more recent controversies over general purpose generative AI, such as GPT technology, but also ongoing discussions about autonomous lethal weapons, highlight the continuous contestation of both the semiotic articulation and the regulatory structuration of AI applications (and these deserve our attention in future research).

Overall, RBR allows the Commission not only to *discursively cohere* competing policy goals of rights protection and innovation, by accentuating them differently in different risk spheres. The risk-based outlawing of some AI uses as matters of deep value conflicts, coupled with the tight risk mitigation of high-risk systems, also serves as a *branding device* in Brussels' fashioning of a distinct trademark of European "cutting-edge AI that can be trusted throughout the world" (see epigraph). As the Commission embarks on buildings "the future we want" in the AI domain (see epigraph), risk-based differentiation serves as its chief epistemic tool—a meta-cognitive device which mediates how itself and others think about AI as an object for regulation—for narrating and regulating into being its vision of a globally competitive common European AI market.

These findings nourish further reflections for scholars of AI regulation, risk-based regulation, and critical political economy. Firstly, the relative dominance of rational choice prescriptions and discourses in research as well as policy consultancy on AI regulation (e.g., Etziani et al., 2021; Krafft et al., 2022; Taeihagh et al., 2021) needs conceptual, methodological, and empirical re-consideration. The analysis illustrates not only the severe epistemic constraints on risk analysis in this domain, but how "performative politics" (Bareis & Katzenbach, 2022) underpin the discursive framing and regulatory treatment of AI systems as nominally *risk-based*. This warns against buying too quickly into "proportionate" AI regulation schemes and warrants critical reflection both on the epistemological requirements of risk analysis and on the potentially more political functions of presenting regulatory decisions as "risk-based" (Paul, 2022).

Secondly, and beyond the specific case of AI, I suggest that a critical political economy foundation is crucial for capturing the performative politics of RBR. As parts of the literature on emergent technology regulation highlight, fierce struggles over global economic competitiveness (but also concerns over national sovereignty and security, cf. Farrand & Carrapico, 2022; Mügge, 2023) strongly shape regulatory agendas and render them part of wider "AI competition state" projects (cf. Paul, 2024 in print). It has so far been unclear, however, what is risk got to do with regulatory competitiveness agendas. Insights from existing sociological institutionalist work on RBR highlight that the risk heuristic appeals because it allows for a well-trusted and allegedly straightforward distinction of regulatory goals, visions, roles, and instruments. Risk analysis tends to preserve its air as objective analytical tool for managing the regulatory space *regardless* of its scientific rigor (cf. Paul, 2021). Future work on RBR can expose in more detail then how regulators employ RBR's coating with an objective aura strategically to

give legitimacy, credibility, and regulatory power to their competitiveness projects, not least in regulatory competition with others. Whether RBR succeeds in convincing consumers, producers, and other jurisdictions of the trustworthiness of European products is another research-worthy question outside the scope of this paper.

Thirdly, European political economy accounts of regulation, while expanding the conceptual–analytical gaze of both RBR scholarship and AI technology regulation debates, would benefit from deeper understanding of the epistemic politics of risk analysis in the common market project. The research shared in this article provided first insights into the links between the epistemic credibility of risk analysis and regulators' pursuit of competitiveness agendas. The role of risk analysis in structuring actors' meta-cognition is well researched in constructivist accounts of RBR (e.g., Black, 2010; Borraz et al., 2022; Paul, 2021; Rothstein et al., 2006), but we can harvest these insights more in accounts of emerging tech regulation as well as European political economy. For example, while it was evident in this paper's research that the Commission engages the risk heuristic to brand European AI as a "trustworthy" high-quality product, we know little about how such articulations change throughout regulatory processes, for example, during the 2022/2023 negotiations of AI regulation between the Commission, the European Parliament, and member states. From a CPE perspective, we could explore further different *agents'* ability *to* contest the *technological selectivity* of the risk heuristic (indeed: interviews with member of the HLEG suggest fierce debates about whether RBR was appropriate for EU AI regulation and where to draw the lines between risk levels). If risk analysis is a critical epistemic tool to structure the regulatory domain in ways accepted by many as legitimate, as this paper argues, then struggles over the specific definitions of risk, methods of actuarial analysis, and related decisions of AI risk management require closer *political* analysis beyond rational choice utopias.

This research further confirms that moral aspirations for democratic values and consumer rights are a crucial part of how Europe seeks to set itself apart in a *globally* competitive space. While this dynamic deserves scrutiny beyond the domain studied here, in the AI context we already observe attempts to advance "market power Europe" (Damro, 2012) in international interactions. Consider, for example, Brussels' efforts to upload its high-risk definition to the Trade and Technology Council, a bilateral forum with US regulators. In this instance, the risk-based fashioning of the European common AI market could leave the realm of narratives and wishful thinking and help *regulate into being* the vision of a future AI world "we want" that guided EU AI regulation to begin with.

Acknowledgments

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Endnotes

[1] The abbreviation "AI" typically describes "computer systems built (at least partially) using machine learning" (Henman, 2020, p. 210) based on complex algorithms that can identify, process, and sort large amounts of data at speed and (partly) self-organize such computation processes based on internal feedback mechanisms ("learning").

[2] This is in extension of Bareis and Katzenbach (2022), who focus on the narrative part but do not consider how such narratives structure regulatory choices.

[3] Importantly, the notion of "orientations" highlights an analytical focus on policymakers' *perceptions* of the role of scientific analysis and rationality in policymaking, in line with my constructivist–interpretivist methodology, detailed in Section 3.

[4]   I am not concerned here with whether such strategizing "succeeds" and results, for instance, in higher market shares for European companies or better protection for EU citizens. The analysis focuses on how the EU's identity as common market shapes what it "says and does" (Damro, 2012) with risk-based AI regulation.

[5]   As Taeihagh and colleagues (Taeihagh et al., 2021) state, these issues are "profound and pervasive" in emerging technology regulation, they perpetuate the usual information and power asymmetries in novel regulatory domains, and they imply constant iteration in regulatory development until a regulatory regime is more mature (p. 3).

[6]   To plausibilize interpretivist reconstruction of the Commission's semiotic distinction of the three risk levels, I ran several smaller quantitative analyses in MaxQDA, including percentual shares of codes across coded segments and the code-relations-browser (see Appendix II, Tables A1–A3).

[7]   IP-02; cf. Vestager (2021).

[8]   Several interviewees report heated normative debates around "red lines" already in the AIHLEG's work on policy recommendations and ethics guidelines (IP-01, IP-04, IP-05, IP-06, and IP-07). These particularly concerned the potential ban of autonomous lethal weapons and research on machine consciousness, which has also been taken up by Parliament.

[9]   IP-02.

[10] I coded almost 100 segments in only three policy documents under this heading (2020 White Paper, 2021 proposed AI regulation, 2021 Communication on Fostering a European Approach to AI), high-risk systems reach a two-third share of all coded segments on the three risk spheres (Table A1 in Appendix II), and the proposal dedicates more than half of all articles (Title III; art. 6–51) to such systems. Arguably, such dominance is in line with standard conceptions of RBR where the high-risk sphere of the risk pyramid requires most "balancing" between the more extreme prohibition and laissez-faire approaches.

[11] The regulation proposal mentions a myriad of specific rights such as "the right to human dignity, respect for private and family life, protection of personal data, freedom of expression and information, freedom of assembly and of association, and non-discrimination, consumer protection, workers' rights, rights of persons with disabilities, right to an effective remedy and to a fair trial, right of defense and the presumption of innocence, right to good administration".

[12] This regulatory goal is mentioned in 24 overlapping segments jointly with "high-risk uses" in interviews and documents, see Table A3, Appendix II. See same table for mentions of "innovation," "trustworthiness," and the idea of a "single AI market" in the high-risk sphere.

[13] Gross (2021) (transcript).

References

Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values*, 47(5), 855–881.

Black, J. (2010). The role of risk in regulatory processes. In R. Baldwin, M. Cave, & M. Lodge (Eds.), *The Oxford handbook of regulation*. Oxford University Press.

Borraz, O., Beaussier, A. L., Wesseling, M., Demeritt, D., Rothstein, H., Hermans, M., Huber, M., & Paul, R. (2022). Why regulators assess risk differently: Regulatory style, business organization, and the varied practice of risk-based food safety inspections across the EU. *Regulation & Governance*, 16(1), 274–292.

Bradford, A. (2015). Exporting standards: The externalization of the EU's regulatory power via markets. *International Review of Law and Economics*, 42, 158–173.

Bradford, A. (2020, December 17). The Brussels effect comes for Big Tech. Project Syndicate.

Cerny, P. G. (1997). Paradoxes of the competition state: The dynamics of political globalization. *Government and Opposition*, 32, 251–274.

Cerny, P. G. (2010). The competition state today: From "raison d'État" to "raison du Monde". *Policy Studies*, 31, 5–21.

Damro, C. (2012). Market power Europe. *Journal of European Public Policy*, 19(5), 682–699.

Eckert, S. (2021). The European Green Deal and the EU's regulatory power in times of crisis. *Journal of Common Market Studies*, 59(S1), 81–91.

Etziani, G., Koene, A., Kumar, R., Santiago, N., & Wright, D. (2021). *A survey of artificial intelligence risk assessment methodologies. The global state of play and leading practices identified*.

Farrand, B., & Carrapico, H. (2022). Digital sovereignty and taking back control: From regulatory capitalism to regulatory mercantilism in EU cybersecurity. *European Security*, 31(3), 435–453.

Finger, M. (2011). Towards a European model of regulatory governance? In D. Levi-Faur (Ed.), *Handbook on the politics of regulation* (pp. 525–535). Edward Elgar Publishing.

Fjørtoft, T. N. (2020). More power, more control: The legitimizing role of expertise in Frontex after the refugee crisis. *Regulation & Governance*, 16, 557–571.

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120.

Henman, P. (2020). Improving public services using artificial intelligence: Possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, 42(4), 209–221.

Hood, C., Rothstein, H., & Baldwin, R. (2001). *The government of risk: Understanding risk regulation regimes*. Oxford University Press.

Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: What it is and what it means. *Information & Communications Technology Law*, 28(1), 65–98.

Jessop, B. (2002). *The future of the capitalist state*. Polity.

Jessop, B. (2009). Cultural political economy and critical policy studies. *Critical Policy Studies*, 3, 336–356.

Jones, M. L. (2015). *The ironies of automation law: Tying policy knots with fair automation practices principles*. Social Science Research Network.

Justo-Hanani, R. (2022). The politics of artificial intelligence regulation and governance reform in the European Union. *Policy Sciences*, 55(1), 137–159.

Krafft, T. D., Zweig, K. A., & König, P. D. (2022). How to regulate algorithmic decision-making: A framework of regulatory requirements for different applications. *Regulation & Governance*, 16, 119–136.

Krarup, T., & Horst, M. (2023). European artificial intelligence policy as digital single market making. *Big Data & Society*, 10(1), 1–14.

Larsson, S., Bogusz, C. I., & Schwarz, J. A. (2020). *Human-centred AI in the EU. Trustworthiness as a strategic priority in the European Member States*. European Liberal Forum.

Majone, G. (2010). Strategic issues in risk regulation and risk management. In *OECD reviews of regulatory reform* (pp. 93–128). Organisation for Economic Co-operation and Development.

Mügge, D. (2023). The securitization of the EU's digital tech regulation. *Journal of European Public Policy*, 30(7), 1431–1446.

Paul, R. (2017). Harmonization by risk analysis? Frontex and the risk-based governance of European border control. *Journal of European Integration*, 39(6), 689–706.

Paul, R. (2021). *Varieties of risk analysis in public administrations: Problem-solving and polity policies in Europe*. Routledge.

Paul, R. (2022). Can critical policy studies outsmart AI? Research agenda on artificial intelligence technologies and public policy. *Critical Policy Studies*, 16(4), 497–509.

Paul, R. (2024). The Politics of Regulating Artificial Intelligence Technologies: Towards AI Competition States. In R. Paul, E. Carmel, & J. Cobbe (Eds.), *Handbook on public policy and artificial intelligence*. Edward Elgar Publishing.

Peck, J. (2023). *Variegated economies*. Oxford University Press.

Petit, N. (2020). *Big tech and the digital economy: The Moligopoly scenario*. Oxford University Press.

Porter, T. M. (1995). *Trust in numbers. The pursuit of objectivity in science and public life*. Princeton University Press.

Ranchordas, S. (2021). *Experimental regulations for AI: Sandboxes for morals and mores*. Social Science Research Network.

Rosamond, B. (2018). The European Union. In T. G. Weiss & R. Wilkinson (Eds.), *International organizations and global governance* (pp. 268–282). Routledge.

Rothstein, H., Borraz, O., & Huber, M. (2013). Risk and the limits of governance: Exploring varied patterns of risk-based governance across Europe. *Regulation & Governance*, 7(2), 215–235. http://onlinelibrary.wiley.com.ezp1.bath.ac.uk/doi/10.1111/j.1748-5991.2012.01153.x/pdf

Rothstein, H., Huber, M., & Gaskell, G. (2006). A theory of risk colonization: The spiraling regulatory logics of societal and institutional risk. *Economy and Society*, 35(1), 91–112.

Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': Regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84.

Schwartz-Shea, P., & Yanow, D. (2012). *Interpretive research design: Concepts and processes*. Routledge.

Stone, D. (2012). *Policy paradox: The art of political decision making*. W.W. Norton & Co.

Sum, N.-L., & Jessop, B. (2013). *Towards a cultural political economy. Putting culture in its place in political economy*. Edward Elgar.

Sunstein, C. (2009). *Laws of fear: Beyond the precautionary principle*. Cambridge University Press.

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40(2), 137–157.

Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*, 15, 1009–1019.

Tang, K.-S. (2020). The use of epistemic tools to facilitate epistemic cognition & metacognition in developing scientific explanation. *Cognition and Instruction*, 38(4), 474–502.

Ulnicane, I. (2022). Emerging technology for economic competitiveness or societal challenges? Framing purpose in artificial intelligence policy. *Global Public Policy and Governance*, 2, 326–345.

Wullweber, J. (2015). Global politics and empty signifiers: The political construction of high technology. *Critical Policy Studies*, 9(1), 78–96.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supporting Information.