Boboc Roxana-Varvara, 2024

Digital Regulation in the European Union: Public Interests Through Digital Sovereignty or Commercial Expansion through Technological Innovation?

PhD Thesis Summary

Artificial intelligence (AI) is transforming societies and economies, raising critical questions about governance, ethics, and societal impact. Within the European Union (EU), AI regulation is a pivotal focus, reflecting the bloc's ambition to balance innovation with ethical responsibility. This thesis investigates the EU's approach to AI regulation, examining how public and private discourses shape policy frameworks, particularly in general AI and healthcare. By analyzing EU policy documents, private sector outputs, and civil society papers, this research highlights the tensions and synergies within the EU's digital public infrastructure strategy. It offers insights into the EU's efforts to lead in AI while safeguarding democratic values and public trust, emphasizing the need for balanced, inclusive regulation. This study contributes to the broader discourse on AI governance, emphasizing the need for a balanced and inclusive regulatory approach.

Reflecting on this thesis' project, it is best to start from the theoretical framing proposed. In this thesis, I have explored the concept of sociotechnical imaginaries and their crucial role in shaping technological innovation and governance. The investigation revealed how these imaginaries, which are collectively held and institutionally stabilized visions of desirable futures, significantly influence both the development and public perception of emerging technologies. Through this lens, I examined how various organized groups, such as supranational bodies, corporations, and non-profits, construct and perpetuate these imaginaries, thereby impacting global technological policies and societal expectations. My analysis focused particularly on the development of artificial intelligence (AI), driven by collective visions that encompass diverse aspects of society, with particular attention to healthcare. The influence of governments, research institutions, and tech companies was underscored, as these entities establish the regulatory frameworks that guide AI research and deployment. This stabilization process, in turn, affects public perception and discourse, highlighting the importance of those who shape these imaginaries. The thesis also delved into the notion of the risk, in an attempt to

propose a reevaluation of responsibilities among science, politics, and the economy. The concept of 'risk manufacturing' was explored to illustrate how societal perceptions of risk are shaped by various stakeholders, including governments and media. Understanding these perceptions is vital for grasping the complex dynamics of sociotechnical imaginaries, particularly in the context of AI.

The discussion focused on the rise of digital platforms and their monopolistic tendencies, which serve as critical infrastructure for social and economic interactions. These platforms exhibit monopolistic characteristics reinforced by network effects, complicating regulatory efforts. The dual role of platforms as mediators of digital services and as foundational infrastructures for the digital economy was emphasized. This duality challenges governance, as platforms shape user experiences and power structures while becoming essential to daily life. The influence of tech companies on economic, social, and political activities was highlighted, along with the impact of sociotechnical imaginaries on AI development and societal integration. The thesis emphasized examining technology within its socio-political and cultural context, advocating a critical approach that avoids technological determinism..

As I transition to the methodological implications of this research, the theoretical underpinnings established here serve as a crucial foundation. The discussions around the co-production of technology and society, the intricate relations of power and interests, and the shaping of sociotechnical imaginaries pave the way for methodologies capable of unpacking the dense interplay between technological advancement and societal impact. This approach anticipates a reflexive, interdisciplinary methodology, equipped to navigate the complexities of power, governance, and the ethical implications of AI integration into public and private spheres. Thus, this thesis provides a comprehensive theoretical framework for understanding sociotechnical imaginaries and their impact on technological development and societal expectations. By highlighting the power dynamics and risks associated with technological advancements, I set the stage for further investigation into these complex phenomena, contributing to a nuanced understanding of how emerging technologies are shaped by – and shape – sociotechnical imaginaries.

Upon this contextual and theoretical backdrop, the thesis embarks on responding to the following research questions:

RQ: How do the European Union and key private companies construct and employ sociotechnical imaginaries to shape regulatory frameworks and knowledge frames of AI at large and in high-risk sectors such as healthcare digitalisation?

RQ1: What are the main discursive techniques and narratives used in EU policy documents, private companies outputs and civil society position papers to reflect certain states of knowledge and anticipated futures for AI technology?

RQ2: Who are the key stakeholders in the EU's AI regulation landscape, and how do their interests, power dynamics, and contributions to public discourse shape the development of AI regulatory policies?

RQ3: What are the main percieved matters of concern, risks and solutions associated with AI as identified in public, private and civil society narratives?

In order to answer these, a multidisciplinary, quali-quanti approach (Rogers, 2019; Venturini & Latour, 2010) is adopted. Each research chapter is developed in the basis of a mixed corpus, covering public, private and civil society public documents and outputs, resulting in 175 documents analyzed in the first research chapter, and 150 in the second one. Each chapter will answer the three sub-research questions, resulting in a discussion section that responds to the overarching question. To answer the first one, a mix of distant and close reading methods (Jänicke et al., 2015) unfolds. The corpus is parsed per stakeholder type and is analyzed through various digital tools to establish key words and phrases, trends in discourse, and particularities for each dataset. Thus, this approach allows for uncovering trends and directs attention towards particular narratives which are further investigated, resulting in a set of discursive characteristics that answer the first research question. The output of the first research question is visualized through alluvial diagrams, word clouds and circular dendrograms.

The second research question is approached through critical stakeholder analyses (Brugha & Varvasovszky, 2000; Varvasovszky & Brugha, 2000). This qualitative method provides insights into e stakeholder's engagement with the specific issue, their degree of interest, their influence or authority in the context, their stated position on the matter, and the impact of the issue on the stakeholder. The output of this approach is visualized as an ample alluvial diagram. The third research question is also tackled though a qualitative close reading approach, namely by developing a risk cartography and issue mapping, or risk manufacturing (G. Beck & Kropp,

2011; U. Beck, 2008; Rogers et al., 2015) approach. This method allows for depicting conflicting knowledge claims coming from key players though institutional contexts. To be able to visualize this process, the key insights are traced so that they reflect institutional structures, narratives, and knowledge claims. The result is a map summarizing the key information regarding the way in which the matter of concern is framed and what solutions are suggested by each type of stakeholder.

The first research chapter highlights how the narratives and strategies of key stakeholders not only reflect their knowledge and visions for AI but also actively shape regulatory landscapes and societal perceptions. The European Union's AI regulation aims to balance economic competitiveness with ethical considerations, positioning itself between the market-driven U.S. and state-controlled China. By promoting "Trustworthy AI," the EU seeks to create a competitive, ethically grounded AI ecosystem. However, this focus on ethics is often criticized as "ethics washing," where ethical discussions are used to delay effective regulation.

The EU's regulatory framework emphasizes building trust in AI through public-private partnerships, which are intended to align public and private interests and leverage European values for a competitive advantage. However, the tension between robust ethical guidelines and economic imperatives is evident. Critics argue that the EU's emphasis on economic value overshadows the need for accountability and strict boundaries for AI deployment. The concept of 'European added value' is used to position the EU as a guardian of principles threatened by non-European entities, fostering collaboration between public and private sectors to enhance innovation while adhering to European values. However, the ambiguity of this concept leads to debates about its effectiveness and alignment with public interests.

A key concern is that public-private partnerships may reinforce existing monopolies instead of dismantling them. The analysis shows how intellectual monopolies are maintained through controlled diffusion processes, where innovation is regulated to maximize economic returns. Civil society stakeholders argue that the EU's risk-based regulatory approach favors private sector interests over citizens' rights.

The research concludes that sociotechnical imaginaries are strategic tools for advancing specific perspectives on emerging technologies, shaping regulatory frameworks and public perceptions. The EU's approach to AI regulation integrates economic and ethical considerations, but the balance between these interests remains contentious, highlighting the

challenges of developing a regulatory framework that genuinely serves the public good while fostering innovation.

The second research chapter explores how sociotechnical imaginaries adapt to the healthcare sector, introducing new power dynamics. The EU's healthcare AI landscape is marked by tensions between empowering citizens and commercializing AI innovations. While the European Commission aims to provide secure access to health data, promote personalized care, and advance research, it also emphasizes creating and commercializing AI applications, positioning the EU as a global leader. This dual focus creates a conflict between individual empowerment and technology commercialization. A key debate centers on data collection and governance. Although the EU claims to support citizen empowerment and the public good, the focus often leans towards supporting innovators and private sector interests, prioritizing economic competitiveness over comprehensive ethical concerns. The EU's discourse frames AI as essential for improving health services and creating market-ready solutions to enhance global competitiveness. This prioritization of data access and innovation lays the foundation for cross-border public-private collaborations aimed at aggregating high-quality data and developing premium products. However, this technological solutionist approach can sometimes overlook potential threats and hastily address concerns.

Big Tech companies play a crucial role in this transformation, leveraging their technical and infrastructural capacities to drive healthcare digitalization. The EU and private companies target patients with different narratives—while the EU seeks to empower citizens through digital health tools, private companies emphasize individual empowerment and access to personal health data. The narrative positions Big Tech as a societal benefactor in healthcare, but the authority to shape AI development strategies often lies with those controlling the narrative. The research shows that private interests dominate public discourse, with commercial actors gaining access to vast amounts of data and regulatory flexibility.

The EU's public-private collaboration framework is not a true public option but an opportunity to address societal and economic needs through competitive, innovative private companies. This approach aims to establish a European tech hub capable of global competition while leveraging public institutions' support. However, the digital transformation in healthcare introduces complexities that exacerbate existing tensions, particularly around data flow management and responsibility. The benefits of AI as a public utility are often reserved for a limited set of participants, potentially widening the digital divide among Member States.

Despite prioritizing commercial interests, the EU remains well-suited to drive digital health innovations, with its digital regulation being more effective at safeguarding digital rights than non-European counterparts. A legally accountable and heavily regulated public-private partnership is seen as the optimal approach to create an alternative to Big Tech services. The research highlights the need for independent oversight, clear guidelines for stakeholder actions, and the coherent development of unified projects. This balanced approach is essential for creating a regulatory framework that genuinely serves the public good while fostering technological innovation.

The overarching insights resulting from both chapters are that while the EU's public discourse aims to integrate ethical considerations and public accountability with technological advancement, it often leans towards supporting commercial interests. The private sector, on the other hand, prioritizes economic competitiveness and innovation, frequently advocating for regulatory flexibility and emphasizing the transformative potential of AI. Both discourses acknowledge the importance of public-private partnerships but differ in their underlying motivations and priorities. The EU seeks to create a balanced framework that ensures ethical AI development while fostering economic growth. In contrast, private companies focus on maximizing market potential and innovation, sometimes at the expense of comprehensive ethical oversight. The tendency of corporate actors to collaborate among each other to either bypass regulation – as seen in research chapter 1, or to expand their technical and infrastructural power – as observed in both research chapters, is an overarching trend within the AI market. This dynamic interplay shapes the evolving landscape of AI regulation in the EU, reflecting the complex negotiation between societal benefits and economic interests.

To conclude, this thesis contributes to the critical STS approach to emerging tech regulation within the EU, highlighting AI as a pivotal governance interest. AI drives a digital revolution reshaping global power in an era of disruption. The thesis reveals how the EU balances public good with an economically-driven agenda, positioning itself as an entrepreneurial facilitator of AI cooperation and a significant global competitor. This ambition is often framed within digital sovereignty, which the EU leverages to justify its AI strategies. The EU's approach emphasizes commercial interests over fundamental rights, promoting responsible experimentation to foster innovation rather than enhancing democratic scrutiny and legal liability.

The thesis critiques the EU's commercial focus, highlighting the need for honest, independent oversight that prioritizes fundamental rights and public value. Despite its ambitious goals, the

EU's digital regulation has been more effective at safeguarding digital rights than non-European entities. A more responsible and heavily regulated public-private partnership is seen as the optimal way to produce an alternative to Big Tech services. The thesis emphasizes the importance of contesting and responsibly engaging with the values that should underpin the EU's proposed public alternative. As the findings indicate, civil societies are a crucial critical voice which hold other entities accountable and emphasize incongruencies and tensions. Overall, the research calls for clear instructions for actions targeted at specific actors and a coherent development of a single project, rather than a cycle of unfinished plans. The EU's role as a platform developer and coordinator of action is crucial in providing an alternative digital public infrastructure, but it must ensure that public interests are not subordinated to commercial ends and that it does not fall prey to technological solutionism or permissionless innovation. While in highly regulated contexts that prioritize fundamental right, technology can improve certain aspects of society, there is – nor it should be – an inevitability associated with AI deployment in all sectors of society, from low to high-risk ones. Technological is advancement is not the outcome of progress, but a byproduct usually of commercial needs. It should be encountered with caution and adapted with scrupulousness, with attention to each sector's risks in particular. This approach is essential for fostering a functional democracy that responsibly engages with the values guiding AI development.