tise transition, innovation and sustainability environments

Co-funded by the Erasmus+ Programme of the European Union

OPEN-SOURCE AI AND THE ARTIFICIAL INTELLIGENCE ACT: EVALUATING EXPECTATIONS AND THEIR POTENTIAL EFFECTS ON EU POLICY-MAKING

Master's thesis for obtaining the academic degree

Master of Science

in the study programme Transition, Innovation and Sustainability Environments

submitted by

Franziska Schranz

Knowledge and Communication Management

at University for Continuing Education Krems

Advisor: Professor Dr. Marguerite Barry

Advisor: Professor Dr. Daniel Cardoso

Innsbruck, 09.07.2023

STATUTORY DECLARATION

I, Franziska Schranz, hereby declare,

1. that I have written my Master's thesis myself, have not used other sources than the ones stated and moreover have not used any illegal tools or unfair means,

2. that I have not used my Master's thesis or parts thereof as an exam paper in my domestic or any foreign country in any form to this date,

3. that, in case my Master's thesis concerns my employer or any other external cooperation partner, I have fully informed them about title, form and content of the Master's thesis and have his/her permission to include the data and information in my written work.

RESTRICTION OF ACCESS

ATTENTION:

The exclusion of utilization of the Master's thesis may be applied for to the Director of Studies (the officer responsible for the administration and enforcement of study law). A maximum of 5 years can be applied for, based on substantiated grounds that the student's legal or economic interests are endangered if it would be published immediately.

The approval document of the accepted exclusion of utilization must be included in the bound and electronic versions of the final Master's thesis instead of this page.

The application should be made with enough time to allow processing and return of the official document from the Director of Studies (minimum 2 months).

ABSTRACT

The EU draft law on artificial intelligence (AI) is the first piece of legislation to propose comprehensive regulations to protect human rights and ethical-democratic values in developing and using AI. This study analyses how open source (OS) developments have been incorporated into the legislative text since the European Commission presented the first draft in April 2021 and what expectations have been associated with it and disseminated in the media. In a first step, all adopted opinion statements of the legislative bodies of the EU were analysed in a document analysis to trace how OS was considered over time. Then expectations on what implications the inclusion of open-source AI (OSAI) would have on the open-source community and the wider public are examined in the course of thematic analysis. Contradicting views are compared and discussed how they resonate with the proposed changes of the AI Act. Findings show that supporters and opponents of OSAI regulation both base their arguments on the themes of security and transparency, as well as progress and dependence on large non-European companies. The stakeholders opposing OSAI regulation used a narrative that portrays the open-source community and BigTech as two distinct camps, a narrative which does not reflect the reality and we should be aware of.

Keywords:

AI Act, Open-Source, Open-Source AI, sociology of expectations,

CONTENTS

1 Introduction

In recent years, the European Union expressed concern about losing its strategic sovereignty as China and US rapidly advance in data collection, Artificial Intelligence (AI) talent acquisition, research & development, and uptake of AI applications by their domestic companies. This dependence on foreign AI providers hampers EU's ability to innovate, impacts on the data privacy of European citizens and limits the ability to shape a reliable and safe AI ecosystem for Europe. To restore and preserve "Europe's ability to act independently in the digital world" (*Digital Sovereignty for Europe*, n.d., p. 1) EU's AI strategy aims to promote the joint development of expertise among member states, invest in developing and disseminating AI and the corresponding infrastructure, and promote digital skills. Supported by a number of European policies and laws such as the "GDPR", the "Digital Service Act", the "Digital Market Act", the "EU Cyber Security Strategy", the "Data Governance Act" and the "AI Act", the EU is simultaneously striving to ensure a European trustworthy AI system of equity, inclusivity and reliability(European Commission, 2021, 2018b). First presented by the Commission in April 2021, the AI Act aims to create a harmonized AI policy among all EU members and avoid legal fragmentation and deployment barriers among EU states, while upholding EU values and human rights. Both European and non-European actors using, deploying, and developing AI in Europe must comply with the law's obligations (AI ACT, 2021) and it is seen as a lighthouse project for AI regulations worldwide. Consequently, the revision and amendment process by the Parliament and the Council is closely followed and commented by national and foreign stakeholders(Oxford Analytica, 2023).

1.1 Stakeholder Efforts to Influence the AI Act

Arguments like the endangering of Europe's digital sovereignty, negatively affecting Europe's Innovation ecosystem as well as the disadvantaging of start-

ups and SMEs were already brought forward by Microsoft and other industry representatives in 2021 with the aim to alter the AI Act to their benefits (Schys, 2023). The Corporate European Observatory reported that strong open and hidden lobbying efforts were made throughout the whole amendment process, especially by US tech corporations. The lobbying started already in the drafting phase, pushing for similar solutions like self-assessment regulations as the US installed. Particularly increased disputes are held since General Purpose AI is discussed to be included in the final draft. Big tech seems eager to exclude GPAIs from the Act and reduce requirements for high-risk AI systems (Schys, 2023). At the AI & Tech Summit organized by Politico, Marian Croak, representing Google, spoke out against the inclusion of General Purpose AI (GPAI). Specifically, she suggested that the inclusion of GPAI, even if used only for research purposes, and the primary focus of regulatory requirements on providers should be reconsidered. (POLITICO Europe, 2022). A similar request was sent by Microsoft in an open letter to Czech Presidency of the Council (*Open Letter on the Proposed Regulation of Artificial Intelligence*, 2022).

In other cases, it happened more surreptitiously. An example of covered lobbying by big tech addressing GPAIS regulation is the open letter of the international alliance of the Software Industry, in September 2022. BSA which was founded by Microsoft, strongly advocated in an open letter in autumn 2022 to exclude GPAIs in the AI Act. European start-ups and SMEs in the digital sector were requested by BSA to sign the letter since including GPAIs would obstruct innovation and development of AI in the EU and amongst large and small providers (Schys, 2023). However, SMEs were suggested by the European Digital SME Alliance to refuse, since this would only serve the tech giants to avoid the burden of complying with European regulations, while the obligations for SMEs would remain the same. Currently, all companies developing GPAI come from outside Europe and are funded by large technology companies. Not regulating GPAI or treating it as low-risk would reinforce the power imbalance between large technology companies and SMEs and just disadvantage the latter (Verdi,

2022). Therefore, even more concerning is the cart that the Allied for Startups, which focuses on a healthy policy environment for start-ups signed the letter (Schys, 2023).

1.2 EUs strategic foresight and the sociology of expectations

Strategic foresight plays a key role for the European Union to inform anticipatory governance and future oriented leadership. Methodologies like horizon scanning, megatrends analysis, scenario panning and visioning are utilized to understand current trends, extrapolate them and evaluating desired future directions and joined future visions among the European Union. Through the collaborative stakeholder engagement from all member states and multidisciplinary expertise the EU seeks for resilient evidence-based policy making and planning When the EU adapted strategic foresight in 1990 it was first mainly used for technological future planning but became now one of the core strategic tools of the European Union also in wider socio-economic contexts (*Strategic Foresight*, n.d.).

We should however acknowledge that formal foresight is entrenched with pre-existing beliefs and assumptions and cannot be easily separated from non-formal expectations. Van Lente (2012) argues that foresight is always flawed by the circulating future expectations and potentially perpetuate existing trajectories rather than encouraging novel solutions. Circulating statements about the future based on profound research and extrapolation or only intuitive estimates are not mere descriptions of an expected future but are "performative", meaning they do something, they induce how we interpret reality, whether they are true or false descriptions. *"Once they are voiced and circulated, they will legitimise, steer and coordinate efforts, also for unintended purposes."* (Van Lente, 2012, p. 778). Ideally by incorporating different stakeholder inputs, policy makers, like the EU try to break out of established thought patterns and prejudices and aim to inform their policy decisions by the most realistic outlook. However, often these stakeholders draw their stances from pre-existing assumptions and thus simply reproducing existing paths (Van Lente, 2012) or voicing assumptions that are

serving their cause and direct efforts towards their preferable future (Berkhout, 2006; Budde & Konrad, 2017). Kerr et al (2020) have examined this mechanism in more detail in relation to AI and communication governance in the EU, particularly in the UK and Ireland. It was noted that economic analysts, government agencies, and research centres presented a positive outlook on the future of AI in order to legitimize the allocation of public and private funds for AI research. Arguments relate mainly to the high expected economic benefits of AI, the need to accelerate AI research and development in order not to fall behind in the race with China and the US. This communication strategy seems to be effective, as there is a positive correlation between the increased emergence of such expectations and investment efforts. Likewise, technical trend predictions from consulting firms could be found in research priorities and legislative texts. On the other hand, critical voices of dissent seemed to focus on questionable aspects of AI handling and development. Ethical issues related to AI development and use have also been strongly reflected in the EU's public communication since 2016. However, Kerr et al (2020) could only demonstrate a "generic" performativity of these ethical expectations, but not an "effective" one. This means that the wordings are already widely used by private and public actors but are not yet visible in the way AI is developed and operated by companies and the public (MacKenzie, 2006). The AI Act could now change this and encourage AI providers, distributors, and users to practically realize the ethical guidelines in the architecture, distribution, and application. However, it is still under debate to what extend these values will ultimately be enforced under the AI Act.

In the last Press conference on the 14th of June 2023, shortly after the Parliament adopted its position on the AI Act, Brando Benifei and Dragos Tudorache, the two rapporteurs on the AI Act, highlighted that in the upcoming discussion on the final draft, they would carefully weigh the text of the law in order to accommodate all stakeholders.(*Press Conference by Roberta METSOLA, EP President, Brando BENIFEI and Dragoş TUDORACHE, Rapporteurs on the AI Act*

Plenary Vote, 2023). Hence, the policy makers will be influenced by the competing visions, assumptions, and fears of stakeholders (Borup et al., 2006) concerning AI and the proposed requirements.

In 2022, concerns were also increasingly voiced that the AI Act could affect OS developers. As a result, debates intensified among experts about whether this would hinder OS projects and translate into barriers for innovation, decentralization, and democratization of AI, especially in Europe (Engler, 2022; Sawers, 2023). As several stakeholders are currently publicly expressing their expectations about how the AI Act should take production into account and what the consequences might be, it is important to make them visible. Van Lente (2012) emphasizes that foresight can never be free from the influence of circulating expectations. Constructed future visions, voiced and circulated can shape the public view but also effect policy efforts. Analysis focusing on assertions concerning specific technologies can reveal underlying presuppositions that might guide the policymaking process. However, such an evaluation raises awareness that strategic foresight based on scientific research and methods is always influenced by commonly shared assumptions, beliefs and future images. Once those mechanisms are recognised and detected, it may help policymakers to apply foresight more efficiently. Since the actual veracity of a vision of the future can only be determined posteriori, it is more meaningful to understand how robust they are than whether they are true. The *robustness* of visions in a constructivist perspective is determined by the number of actors involved in reflecting on and forming these anticipations. Further, Van Lente concludesthat formal assertions are vulnerable to two mechanisms: If they deviate too far from common informal expectations, they risk not being effective because they lack legitimacy; if they overlap too much, they fail to attract public attention.

1.3 Research Aim and Research Questions

In light of the significant effect of voiced expectations on foresight and the decisions of policymakers, this work aims to investigate the circulating expectations surrounding the AI Act and its impacts on OS productions and its connected benefits.

RQ 1: What are the circulating expectations regarding the AI Act and its impact on open-source efforts and the characteristics attributed to open-source development?

To put the emerging expectations into context, the secondary goal of this work is to analyse how OS systems were taken into account in the first policy drafts of the AI Act and which proposed changes related to OS were put forward by the EU institutions.

RQ 2: How are open-source productions considered in the original proposal of the AI Act in 2021, and how did the consideration of open-source develop over time up to the first reading in June 2023?

And finally, to inform if the emerging mechanisms might have impacted the policy outcome, the results of RQ1 and RQ2 will be compared.

RQ 3: Are these expectations reflected in the position and opinion papers of the Parliament and the Council?

The overview of the circulating expectations might be even more important as with the first reading being adopted by the majority in the parliament, the legislative process commences now to the trialogue. This tripartite negotiation happens disclosed from the public in camera between few representatives of the Parliament, Commission and Council to agree on the final draft of the AI Act (Think Tank European Parliament, 2021). The Corporate Europe Observatory claims that particularly this stage is vulnerable to lobbying as it is the last chance

to influence the outcome of the final text and examples from the past showed that industry and other stakeholders intensified their influential efforts in this crucial phase. Additionally, as this discussion takes place behind closed doors there is no oversight from the public, while well connected lobbyists still were able to get information about the tripartite discussion and intervene (Corporate Europe Observatory, 2022).

This study will first provide an introduction to AI and OS, by elucidating the disputed definitions, their historical background, and associated characteristics. It seeks to offer an understanding of what encompasses AI and OS and how their significance is specifically depicted within the European Union. Followed by an overview of the current discourse on OS development and its relationship to AI law before moving on to the analysis.

2 AI

2.1 Artificial Intelligence (AI) definition

The term "artificial intelligence" (AI) was coined in 1956 by John McCarthy and introduced as a study field. Back then no common definition, theory or methodology concerning AI was agreed on the only consensus of the scientific community was the vision of computers being able to learn and perform intelligent tasks (James Moor, 2006). Noteworthy, is that even before becoming a scientific discipline, myths, fictions and theories on artificial intelligence or thinking machines existed (Sheikh et al., 2023). In philosophical and academic circles AI theories, without referring to it specifically as AI, emerged with the efforts to describe the process of human thinking. Rosenblueth et al. (1943) attempted to outline the behavioural processes of organisms, including humans, and described purposeful behaviour as a process initiated by a goal and adjusted by negative feedback (Arturo Rosenblueth et al., 1943). Building on this theory, Norbert Wiener introduced the term cybernetics, which describes the concept of circular control systems that realign themselves based on feedback and, among

other things, describes the learning process that can be applied to computers and is comparable to the human learning process. This idea can be seen as an inspiration for modern techniques such as reinforced learning for training algorithms (Li et al., 2019). Around the same time, in 1950, Allan Turing, in his paper "Computing Machinery and Intelligence," introduced the imitation game now commonly known as the Turing Test. To play the game successfully, a machine must convince an interviewer that it is human. Turing points out at the beginning of his paper that the question is not whether machines can think, but rather whether computers can perform tasks-in the case of the game, answering questions-as well as humans. He believes that the minds of computers can be compared to those of children, who can learn how to act through reward and punishment. Moreover, he assumed that the ability of machines to pass the test and perform human tasks is only a matter of memory and computing capacity, but refers mainly to tasks based on logic. (Allan M. Turing, 1950). However, logic and pure mathematical calculations were already disputed as the only prerequisite for human-like intelligence. Descartes argued in the 17th century that logic can be imitated by machines, but only humans can reason, i.e. justify why they do something, reflect on their behaviour and adapt it. While we know that at least the latter already refutes the concept of cybernetics, it is evident that in the 20th century, the notion of intelligence is still strongly characterized by logic and mathematical problem-solving. (Bruce Malzlish, 1994). Academics argue that we are still failing to define AI as we still do not fully understand what constitutes human intelligence. And maybe the human-like intelligence is not the one we should aspire to replicate as the output might seem the same, but the process is very different. On this note, we must keep in mind that our perception of AI is also changing in accordance with our technological abilities. Currently advances in pattern recognition and self-learning algorithms techniques that are referred to as Machine Learning (ML) or for more advanced algorithms Deep Learning (DL) are shaping our perception. Nowadays, many people perceive DL

and ML algorithms as AI. This perception will most likely change again with innovations and developments in this field (Sheikh et al., 2023).

AI has generated much debate not just within scientific communities but also among science fiction writers. It is important to include the narratives that have emerged in SF in the discussion of AI, as they reflect and shape our expectations of AI, its capabilities, and risks, which in turn can influence policymakers and developers (Hermann, 2023). Therefore, a brief overview of how AI has been portrayed in science fiction books and movies will be provided here. Widely recognized as one of the first works of SF and depictions of artificially created intelligence is Mary Shelley's Frankenstein from 1818. In this story, the artificial human, Frankenstein's monster, demonstrates the ability to learn and assist his creator but ultimately becomes vengeful due to negative influences from society. (Byrd & Paquette, 2023). Some may argue that Shelley's creation of the monster in her novel "Frankenstein," which primarily consists of organic components, differs significantly from contemporary conceptions of AI. However, *HAL 9000*, a supercomputer featured in the 1968 science fiction work *2001: A Space Odyssey*, represents a much closer alignment with the modern understanding of AI. Unlike the monster in *Frankenstein*, *HAL 9000* does not possess an organic body but autonomy and self-awareness of its existence (Hermann, 2023). Since then, many SF novels and films have explored different aspects of AI, including empathy. For example, in SFs like *Ex Machina* 2014 ("*Ex Machina* (Film)," 2023)or *Her* 2013 ("*Her* (Film)," 2023), the AIs systems demonstrate the capacity to recognize and respond appropriately to human emotions. 2021, in the latest novel of Kazuo Ishiguro "Klara and the Sun," protagonist Klara, a humanoid robot, is even described as a sentient being with feelings (Ajeesh & Rukmini, 2023). The properties of agency, as described by Chopra and White (2011) as the capacity to choose for oneself how to accomplish a task and carry it out, self-awareness, and the ability to learn, appear to be present in all of the SF narratives that were discussed above. Interestingly to note, is that in the narratives of Frankenstein's Monster and HAL 9000, where AI turns against humans, AI is shown as great

support for the people who use it and is not inherently bad but turns maleficent due to the shortcomings of their developers and surroundings influencing its learning process (Byrd & Paquette, 2023; Hermann, 2023). Further, it can be argued that the increased emergence of empathic AIs in recent works might be attributed to the fact that the recognition of empathy, also called emotional intelligence and its importance grew at the end of the 20th century (F Indoannidou & V Konstantikaki, 2008). Certainly, there are already technologies on the market that display empathy (Zhou et al., 2020) or autonomy (Chopra & White, 2011) to some degree, but there is no technology yet that has the full range of the above-mentioned cognitive abilities as humans do (Sheikh et al., 2023). AI can currently be seen mostly as a marketing term rather than a description for the aofal technical capabilities of the product or service being offered (Michael Atleson, 2023).

Today experts often divide AI into artificial general intelligence (AGI) and artificial narrow intelligence (ANI). AGIs are described as reflective, flexible, and learning and possesses information processing capacity able to be applied to a vast variety of tasks and contexts. While this form of AI is closest to the first definition given by John McCarthy and the other participants of the Dartmouth Summer Research Project, there are no applications called AI today that would deserve this name including the large language models (LLMs) like Chat GPT. However, we have already numerous systems which would be categorised as ANIs which are able to fulfill tasks in specific domains (Shevlin et al., 2019). Even if AI, which could meet the criteria of the original definition, is estimated not to be achieved in the next 80 years, already "simple", narrow systems are already having the ability to greatly impact our lives (Sheikh et al., 2023). Let's take the emphatic chatbot Xiaolce as an example. Although this chatbot is not a sentiment entity, Xiaolce was still able to establish several emotional relationships with its users and influence them (Zhou et al., 2020). Therefore, a definition is required that allows policymakers to already govern these “simple” AI components and systems.

Even though there have been various attempts by governing institutions around the world to provide a definition for AI, none has become universally accepted. An extensive work by AI Watch compiled and compared numerous existing definitions and identified common taxonomies (European Commission et al., 2020). In the strategic and informative EU texts of the last five years, the predominant AI characteristics are: machine-based system; taking inputs from the environment; ability to intelligently process and perform a specific task; the ability to learn and improve (Annoni et al., 2018; European Commission, 2018a; High-Level Expert Group on Artificial Intelligence, 2019).

The Artificial Intelligence Act (AI Act), in addition to its legislative purpose, also aims to introduce a unified definition of AI. In the first proposal, the definition was as follows: *'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with* (AI ACT, 2021, art. 3 (1)). However, this very broad definition by the European Commission is fleshed out by reference to Annex I, which defines AI by the technical circumstances that produced it. For example, software is considered an AI system if it was developed through "machine learning", "including deep learning", "statistical approaches" or "logic-based" (ANNEXES 1 to 9, 2021, p. 1). The enumeration with the possibility of expansion was specifically included in the definition to reflect the rapidly evolving field of AI, allowing the law to be expanded to include new developments as needed. Despite efforts to narrow the definition by looking at the techniques used to develop the software, it still leaves too much room for interpretation or loopholes to move to other development measures (Nicolas Kayser-Bril, 2021). The text recently adopted by the European Union Parliament on June 14, 2023, however, seems to broaden the scope even further by changing the text from "designed using specific techniques" to "designed to operate with varying degrees of autonomy" and talking about "machine-based

systems" instead of software (Texts Adopted, 2023, art. 3 (1)). The definition for the AI Act was also chosen to be consistent with the official definition of the U.S. National Artificial Intelligence Initiative Act of 2020 (National Artificial Intelligence Initiative Act of 2020, 2021) to facilitate the process of creating common rules and standards (*Press Conference by Roberta METSOLA, EP President, Brando BENIFEI and Dragoş TUDORACHE, Rapporteurs on the AI Act Plenary Vote*, 2023).

2.2 The role of AI

Artificial intelligence is expected to play important economic role and to bring systemic change and advancements across several industries, sectors and facets of our life (Klaus Schwab, 2023). A simulation study by McKinsey on the impact of AI on the global economy has shown that the introduction of AI will have a positive impact on GDP growth and that countries and companies that are early adopters of AI in particular will benefit economically. The simulation estimates that AI will contribute 1.2 percent annually to global GDP growth from 2018 to 2030. Compared to other general-purpose technologies, which accounted for a maximum of 0.6 percent, this increase would be enormous. These improvements in GDP would accumulate and lead to a 16% increase in global economic output in 2030 compared to the 2018 baseline Broken down to individual countries, early adopters of AI could see growth of 20 to 25 percent in their overall economic output, and companies having fully adopted AI for their processes or business model until 2025 are expected to double their returns by 2030. Enterprises, on the other hand, who miss to do so or only absorb AI on a marginal level will face a 20 percent revenue loss (McKinsey&Company, 2018). Despite the economical net gain AI is expected to support reaching the environmental and societal focused Sustainable Development Goals (SDGs). For example, accelerating green energy integration through enabling smart grid or improving health care by AI supporting diagnoses and predictive health care (Vinuesa et al., 2020).

2.3 The distribution of research and development efforts

However, there are concerns about the equitable accessibility of AI opportunities, potentially leading to a divide between those who have the means and knowledge to benefit from AI and those who do not. This phenomenon, which is derived from the term digital divide that includes factors such as connectivity, access to digital devices, and the skills necessary to derive socioeconomic benefits from digital products, is now evolving into an "AI divide." This divide additionally relates to the availability of training data needed to develop and improve AI systems, which adds another layer to the issue. Such a divide can already be observed in the research and development activities concerning AI. When looking at the main representatives in academia we can see a clear geographic concentration of expertise as 10 countries account for 75% of academic literature on AI (Irene Kitsara, 2022). The biggest contributor are countries from Asia and the Middle East followed by the United States and select European nations. In Europe, the United Kingdom, Germany, France, Spain, and Italy are the top performers, but from an international perspective, China and the United States are clearly the top performers in academia. Their publication volume is more than three times that of the third-placed United Kingdom (Irene Kitsara, 2022).

A similar picture emerges when looking at innovation and development in the AI field. When considering the number of patent families as a measure of innovative activity, only two European companies, Bosch and Nokia, rank within the top thirty. Elsewhere the higher ranks are filled with entities from China, USA, Japan and Korea. Also 25 out of the top 30 are corporate entities, with IBM, Microsoft and TOSHIBA leading the way. Geography wise China and the US are clearly the main competitors in AI innovation, development, and research, mainly filed by big multinational companies in the Information and Communication Technology (ICT) sector (Irene Kitsara, 2022). We can therefore see a clear AI gap between countries, SMEs and multinationals, different sectors

and on the job market. Increased competition for expertise among companies for workers with AI utilization or development skills are expected. The share of occupations requiring digital skills is projected to increase by 10 percentage points within 12 years by 2030 while employment of workers with low digital skills would decline, as would their wages (McKinsey&Company, 2018).

2.4 Importance of decentralized AI development

As AI enables industries and companies to increase their efficiency and cut costs through automation and optimized processes the competition gap between those industries and companies applying AI and those who do not widens (Irene Kitsara, 2022), which aligns with McKinsey's simulation results. This means that competition and power might be distributed amongst only a few big ICT companies. Early adopters of AI and entities involved in AI development may retain their competitive advantage by acquiring specialized knowledge and skills and having access to large amounts of proprietary data while increasing the entry barrier for newcomers (Irene Kitsara, 2022). Regular data collection and retrieval processes enable continuous updating and enrichment of their datasets, which can be used as a basis for continuous improvement of their AI systems. Newcomers will find it difficult to catch up with data volume and acquired skills, which could undermine the mechanisms of healthy competition and concentrate power in the hands of a few providers. This will also mean that a few large AI models will dominate, and with them the biases that are built into these systems. The convergence of algorithmic methods is further driven by the release of these methods by tech giants like Alphabet via open source so that other developers can build on them. Due to the lack of alternatives and low diversification, the same inequalities and errors will always be reproduced (Ekkehard Ernst, 2022).

The risk of power asymmetries and reinforcement of inequalities was also highlighted by the UN High Commissioner for Human Rights and the EU Fundamental Rights Agency(Nicolas Kayser-Bril, 2021). In line with that warning Vinuesa et al. (2020) found in their study that a total of 134 targets out of the 169

SDG targets were positively impacted by AI. However, the benefits AI is going to create will be unequally distributed, gained mainly by those who have the computational power and skills to develop AI. Since most AI developers are currently working in advanced economies, AI solutions are specifically designed to address the challenges of the Global North, and so is their design. Digitally disadvantaged countries therefore not only have difficulty accessing AI applications, but the applications offered are also less relevant to their needs and cannot be easily adopted because they are not adapted to the infrastructure and context of the region in question. Thus, less prosperous regions also lag behind in achieving the SDGs (Vinuesa et al., 2020). At the same time workers in the global south are exploited to train the data intensive AI models (Mohamed et al., 2020). A Times investigation revealed that OpenAI was paying Kenyan workers a maximum of 2$ per hour to label maleficent texts to detect and remove toxic data from ChatGPTs training data set. Workers were not only underpaid, but also suffered the psychological consequences of being constantly confronted with harmful content. (Julia Zorthian, 2023) "Responsible development" (*Safety Standards*, n.d.) as advocated on OpenAIs safety-standard declaration was thus strived for the end user on the expense of ghost workers earlier in the value chain. The company was reported to have cancelled their contract with OpenAI by now (Julia Zorthian, 2023), however that does not change the fact that machine learning models still need huge quantities of labelled, classified and cleaned data usually done by so called "ghost workers" in ICT poor countries(Mohamed et al., 2020). This makes it all the more important to put in place robust policies and oversight mechanisms to ensure safe and equitable AI systems, but also to decentralize AI development to enable diversified solutions tailored to different cultural and socioeconomic conditions, regardless of location and economic power (Vinuesa et al., 2020).

3 Open-Source AI

A prime example of decentralized production and democratization of AI are OS projects. Ideally everyone can contribute, resources like software and sometimes even hardware are shared, and ideally no single entity has complete control over the direction of the project. Of course, there must be an understanding to commission and modify these systems, but in theory, anyone can modify and use the AI system according to their needs so that there are equal opportunities to benefit from AI advances (Yochai Benkler, 2006).

3.1 History and Definition

The development of open-source software (OSS) can be traced back to the software project GNU founded in 1983 by Richard Stallman. After UNIX, as a platform that was a collaborative effort of developers, went private, Stallman, was politically motivated to create a non-proprietary system as a freely available alternative to Unix. He wanted to enable free access to and use of information and software. To ensure that the software he developed in collaboration with others did not become closed software downstream, he formed the Free Software Foundation (FSF) (Yochai Benkler, 2006), whose main goal is to promote and protect "computer user freedom"(Free Software Foundation, n.d.). However, if any work is shared with the public, it is not automatically made freely available, but by default protected under copyright, implying that every project contributor also holds copyright by default. These principles of intellectual property law were used by the FSF to create a special copyright license called the General Public License (GPL). The copyright terms of the GPL authorize anyone to use, modify, and distribute the GNU software, but the derivative software is required to be redistributed under the same terms. This method was defined by the FSF and is now widely known as "copyleft". While the free software movement prioritizes the importance of software freedom and political ideology, the open-source community which later emerged chose the term "open-source" to distance itself from any political motivation. OS is merely a decentralized

production model which promotes effectiveness and efficiency and can run under different licenses which may also not prohibit proprietary software downstream. Therefore any free software may be open-source but not any OSS is free software(Yochai Benkler, 2006).

In general, two kinds of OS licenses can be distinguished: “permissive” and “strong copyleft”. Popular "permissive licenses" include Apache 2.0 or MIT, which exclude any conditions on the use or further processing, allowing downstream developers to turn the OS project into a closed project. “Strong copyleft” licenses, on the other hand, grant the same rights but with the obligation for any directive work, to be released under the same license. Meaning all downstream projects must also stay open source and attribute the original source. The GPLv3 license is even prohibiting any connected closed services (*The Legal Side of Open Source*, 2023).

A definition of open source is maintained by the Open Source Initiative (OSI), founded in 1998 to create a common understanding for the community of what open entails (*About*, 2006) including.:

1. “Free Redistribution”: The license should not limitate the user in distributing the software solitarily or integrated in an aggregated software.

2. “Source Code”: The underlying code shall be provided comprehensively and free of charge. The source code and any compiled form should be accessible within the software or at least over a corresponding platform without any constraints.

3. “Derived Works”: The source code must be modifiable and the derivatives of it unrestricted distributable.

(*About*, 2006)

OSS development demands open authorship. Developers do not provide any guarantee for whatever they produce. Instead, there is a record of who the author is and of what part, along with their contact details. The authors name and how to contact them is made public. Trust is established since every author takes public ownership of their work and this work determines their reputation (Perrin, 2021). Revenue is sometimes generated through different strategies (Knut Blind et al., 2021).

Accompanying services: Offering and charging for auxiliary services that complement the free OS product. These additional services, which include maintenance, customization, consulting, and training, serve to enhance the value proposition of the core OS offering. A notable example of this business model is Red Hat, which successfully provides training, support, and maintenance services to meet the needs of users of the Linux operating system, which is based on the OS kernel (Knut Blind et al., 2021).

Corporate development and distribution: When a company intentionally distributes software, created and adapted by a paid developer to the company's needs, as open source. The revenue the company gets out of distributing their software freely, is not a direct profit, but the free maintenance, improvement and contribution of the community using the open source-software. Raspberry Pi is a known example using this kind of business model. (Okoli et al, 2015)

Software as a Service (SaaS) with distribution of server software is a strategy that relies on an online usage of OSS. The software is provided free of charge. However, due to the software's complexity and requirement for specific hardware, companies can turn a profit by offering an easy but charged online access to servers running the software and by hosting cloud services. WordPress is an example of a cloud service provider using SaaS as a business model(Knut Blind et al., 2021).

Offering Open Core/ Dual-licensing/ Selling: For this business model a charge free core version of a software is released as open source. If the users or developers need additional features, a paid proprietary license can be purchased offering these additions. Such additional features are often the ones that are mainly of interest to businesses rather than individuals(Knut Blind et al., 2021).

Further business models are supporter membership, crowd funding or playing advertisements. For OSS, where frequent updates on fixes and bugs are necessary, a paid subscription to receive the latest update also is a profit strategy(Knut Blind et al., 2021).

3.2 Degrees of openness

When discussing access to these OS models, it is important to define what types and degrees of openness actually exist in terms of software and hardware. Balka et al. (2010) categorizes and describes three forms of openness: transparency, accessibility, and replicability. Transparency refers to the quantity and quality of information revealed to the public. Accessibility refers to the ability of the public to participate in the product development. Replicability refers to the level of ease in which the product could be self-assembled. The researchers found that software tends to be more open than hardware, but also that participants generally agreed that openness of software was more important than openness of hardware. However, this survey was conducted in 2009 and would be interesting to revisit in today's technology landscape.

Specifically with Large Language Models (LLMs), however, a distinct issue arises in terms of hardware capabilities, regardless of the degree of openness provided to the public. To examine this issue in detail, we can compare and contrast two organizations with differing degrees of openness: OpenAI and Hugging Face. OpenAI, with it's massive popularity with ChatGPT and even including "open" in it's company name, is actually not the ideal open source platform. While the web interface to ChatGPT is free for the greater public, the GPT models, as well as

some of the refinement techniques used to produce the GPT models, remain proprietary. Indeed, OpenAI has publicly published technical papers for each of the GPT models, but these papers explicitly do not include specifying the size of the model, the architecture of the model, or details of how the models were trained (OpenAI, 2023). In a separate interview, however, CEO Sam Altman estimated the total training cost over $100 million. Additionally, while the web interface is free to the public, API access to the GPT models remains as a pay-per-use service, following a typical API SaaS business plan, where users can fill their accounts with a specified amount to use the APIs (Knight, 2023).

A stark openness contrast to OpenAI is Hugging Face, an completely OS organization currently consisting of 159 developers that seeks to democratize and transparently share the source code, model architecture, and training techniques of all models within their organization. They have also been known to implement OS versions of some of the newest models described in machine learning literature. The most popular LLM from Hugging Face is BLOOM, created by the BigScience (Le Scao et al., 2023) Research Workshop. It should be noted that one of the original motivations in forming BigScience Research Workshop was to combat exactly the closed source pattern of companies like OpenAI. The BigScience Research Workshop website states: "The resources for this endeavor are found mainly in the hands of big technology giants. The stranglehold on this transformative technology poses problems, from a research advancement, environmental, ethical and societal perspective." (*BigScience Research Workshop*, n.d.). In Le Scao et al (2023)., the original BLOOM publication, every step of data sourcing, training, and model architecture is explained in great detail. However, it should be stressed that despite every informational aspect of the model being made public, a typical individual would likely not have the financial resources to train the model. The training of BLOOM itself was only possible through a French public grant allowing the team to use the IDRIS Jean Zay supercomputer.

In conclusion, regardless of openness of these models, there is then a clear risk that training of the largest of the AI models continue to be restricted by organizations with large, which could potentially lead to a further centralization and restriction of which entities control these models. Exactly because these LLMs require extensive and thus financial restrictive amounts of training creates a unique challenge for OS groups to maintain a similar pace to that of technology leaders including OpenAI, Amazon, Microsoft, and Google. Furthermore, even after they are trained, these models require massive amounts of memory to run. In the case of BLOOM, likely not even as large as GPT-4, requiring about 350 GB of RAM just to run (Le Scao et al., 2023). Using a popular GPU provider like Vultr, we can see a private instance of BLOOM would cost an upwards of $7000 per month to run (*Cloud Compute*, n.d.). The ability for smaller players to afford both the training cost and operation costs of these models becomes questionable.

3.3 Benefits of open-source productions communicated in the EU

On behalf of the European Commission, a broad study was conducted to analyze the impact of open source on the European economy. A positive impact on the European economy ranging from €65 to €95 billion, was identified for the year 2018 as a direct result of globally available OSS. The study further suggests that OSS provides a cost-benefit ratio of above 1:4, it is also predicted that a 10% annual increase of OSS contributes to a 0.4% to 0.6% increase in GDP for EU Member States. (Knut Blind et al., 2021)

Compared to proprietary software, open source does not require high costs for maintenance efforts and licences. It is free and maintained publicly by several developers. In addition to the cost savings, OSS has the advantage of not being restricted to one provider, leaving a high motivation for the public sector to also take advantage of these benefits (European Commission. Directorate General for Communications Networks, Content and Technology., 2021). Additionally, a study specifically on OS and AI commissioned by the European Parliament

concluded that OS fosters the local uptake of AI due to the low entry barriers and easy interoperability through open standards (Theben et al., 2021). For example, Hugging Face, one of the largest nonprofit OS providers, makes its machine learning models available to small developers who would not be able to train such a large model themselves. It also lowers the barrier to entry for developing countries by providing access to their models via APIs, as the high computational power required to run large machine learning models is often a limiting factor for countries with low computing capacity (ALGORITHM WATCH, 2023).

To train big machine learning models enormous computing power is needed which subsequently produces carbon emissions. This is amongst other reasons a motivating factor for Hugging Face to make their LLMs "Bloom" openly available (ALGORITHM WATCH, 2023). Developers can fork the project, meaning they can copy the code of the LML and build upon and refine the model to their needs (Knut Blind, Mirko Böhm, et al., 2021) avoiding the energy-intensive work of training it from scratch. Not creating a project from scratch. Further, it produces well-structured code, improved by revision and competition which is fostered through open sources and consume less power (ALGORITHM WATCH, 2023).

Forking practice also means security for the end user. When a proprietary software is abandoned by its developer, it is almost impossible for the end user to get technical support from another source to update the software and fix bugs. In contrast, an OSS has no copyright ownership and can be easily forked and maintained by a new developer. (European Commission. Directorate General for Communications Networks, Content and Technology., 2021) Further, the resulting peer monitoring, of making AI open-source available increases the probability of bugs and risk detection and mitigation (Shen et al., 2021) as Linus's law describes "given enough eyeballs, all bugs are shallow" (Eric Raymond, 2002, p. 19). Even tough, unlike proprietary software provider, OSS developers do not warrant f the functionality and safety of their software trust is s established through transparency. The exact information about who contributed

what code at what time, with the author's contact information, is a common documentation process in most OS projects and is publicly visible (Perrin, 2021).

Direct beneficiaries of OSS are start-ups from the information and communications technology sector. An increase of OSS is therefore expected to be proportional to the amount of Start Ups founded in Member States of the European Union. However, the study revealed a high dominance of big US companies basing their business models on OSS, constantly benefiting from the significant OSS contributions of developer from small European companies (Knut Blind et al., 2021).

Direct access to the source code, access to constant exchange of knowledge, and the drive for innovation were identified by the study to be further motivating factors for using and contributing OSS.

4 Methods

This thesis employs a set of qualitative methods to identify and examine the circulating visions, assumptions, and statements regarding the impact of the AI Act on open source. Additionally, it offers an overview of how open source is explicitly or implicitly addressed within the AI Act and investigates whether the identified expectations align with the decisions made by policymakers.

Expectations are constituted through formal mechanisms like research activities, and informal mechanisms including opinion statements and prophesies once circulated and shared they are performative (Van Lente, 2012). Hence, to identify expert statements concerning OS and the AI Act shared and referenced in media, a Thematic Analysis (TA) is applied on 10 expert opinions circulated in news articles and blog post published online since the first proposal of the Commission. TA accommodates to detect recurring themes and interpret viewpoints and therefore appropriate for this study. By utilizing TA the circulating expectations regarding the AI Act and its impact on OS efforts and the characteristics attributed to OS development (2) will be detected.

To explore how OS productions are considered in the original proposal of the AI Act in 2021 and how the consideration of OS did develop over time (RQ2) a document analysis was applied. A document analysis allows a systematic exploration of textual content and to understand how specific themes were addressed and how they evolved over time.(Karppinen & Moe, 2012). In this study a total of 29 texts of primary sources (Jääsaari, 2007) that were officially adopted by governing bodies during the AI Act legislative process and analysed and compared over time the specific articles and paragraphs that could potentially impact open source developments. The collection of the relevant documents was informed by the ordinary legislative procedure (OLP) of the EU which will be outlined below.

Finally, to investigate if the formulated future visions correspond with the amendment process of the AI Act. The results of the first and second research question were compared.

Subsequent chapters outline the distinct methodologies and datasets employed, followed by a comprehensive overview of the obtained results.

4.1 Step 1: Document Analysis

4.1.1.1 Sampling

Since the resulting first draft of the AI Act was officially proposed in April 2021 the draft has been revised and amended. The council adopted its official position on the 6th of December, commonly known as the General Approach. Most legislative-associated committees of the Parliament handed in their opinion papers. In June 2023 the Committee on Internal Market and Consumer Protection (IMCO) and the Committee on Civil Liberties, Justice and Home Affairs (LIBE), leading the discussion adopted their proposed changes (European Parliament, n.d.). In June the Parliaments position was presented in the first reading and accepted by the three institutions. The final legislative text will now be discussed

in the trilogue to reach a mutual agreement on a definitive text between the three institutions (*Policy, Law – Decision-Making Process | European Union*, n.d.).

To point out how OS was considered along this amendment process, opinion and position papers by the EU institutions with legislative power involved in the ordinary legislative process (OLP) as outline above were collected for the document analysis (see Table 1)

Table 1: Documents used for the document analysis

Date	Document Title	Pages	Source
2021-04-22	Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	108	Eur_lex
2021-04-23	ANNEXES to the Proposal REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	17	Eur_lex
2021-04-23	From: Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNIONLEGISLATIVE ACTS	97	Eur_lex
2021-04-23	From: Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT ANNEXES Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	51	Eur_lex
2021-04-23	From: Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director COMMISSION STAFF WORKING DOCUMENT EXECUTIVE SUMMARY OF THE IMPACT ASSESSMENT REPORT Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	5	Eur_lex
2021-04-23	From: Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director REGULATORY SCRUTINY BOARD OPINION Proposal for a Regulation of the European Parliament and of the Council aying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts	8	Eur_lex

2021-04-23	From: Secretary-General of the European Commission, signed by Ms Martine DEPREZ, Director Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	121	Eur_lex
2021-06-14	From: The Chamber of Deputies of the Czech Republic Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts - Opinion on the application of the Principles of Subsidiarity and Proportionality	20	Eur_lex
2021-06-15	From: General Secretariat of the Council Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Optional consultation of the Committee of the Regions	2	Eur_lex
2021-08-13	From: The Senate of the Parliament of the Czech Republic Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS	5	Eur_lex
2021-08-13	From: The Portuguese Parliament Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) ANDAMENDING CERTAIN UNION LEGISLATIVE ACTS - Opinion on the application of the Principles of Subsidiarity and Proportionality	12	Eur_lex
2021-09-22	Official Journal of the European Union Opinion of the European Economic and Social Committee on Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts	6	Eur_lex
2021-10-12	From: The Senate of the Parliament of the Czech Republic COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Fostering a European approach to Artificial Intelligence - Opinion on the application of the Principles of Subsidiarity and Proportionality	6	Eur_lex
2021-10-18	From: The Polish Senate COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Fostering a European approach to Artificial Intelligence - Opinion on the application of the Principles of Subsidiarity and Proportionality	6	Eur_lex
2021-12-02	Official Journal of the European Union COMMITTEE OF THE REGIONS Opinion of the European Committee of the Regions — European approach to artificial intelligence — Artificial Intelligence Act	26	Eur_lex

2021-12-29	Official Journal of the European Union OPINION OF THE EUROPEAN CENTRAL BANK on a proposal for a regulation laying down harmonised rules on artificial intelligence	7	Eur_lex
2022-04-22	From: The Senate of the Italian Chamber of deputies Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Opinion on the application of the Principles of Subsidiarity and Proportionality	8	Eur_lex
2022-12-06	From: General Secretariat of the Council Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach	208	Eur_lex
2022-04-22	From: Committee on the Internal Market and Consumer Protection, Committee on Civil Liberties, Justice and Home Affairs DRAFT REPORT on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts	161	www.europarl.europa.eu
2022-04-22	From: Committee on the Environment, Public Health and Food Safety OPINION on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts	66	www.europarl.europa.eu
2022-06-14	From: Committee on Industry, Research and Energy OPINION on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts	49	www.europarl.europa.eu
2022-06-16	From: Committee on Culture and Education OPINION on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts	41	www.europarl.europa.eu
2022-07-12	From: Committee on Transport and Tourism OPINION on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts	70	www.europarl.europa.eu
2022-09-12	From: of the Committee on Legal Affairs OPINION on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts	74	www.europarl.europa.eu

4.1.2 Method and Analysis

The extracted documents are all adopted position and opinion papers are European Commission, the Council of the European Union, and the European

Parliament, as well as the opinions of the Economic and Social Committee, the European Central Bank and the European Committee the latter are advisory bodies without legislative power but involved in the OLP. This collection of documents has been informed by the EUR-Lex website, the official website of the European Union for publication and information on Union law, which provides an overview of the legislative process of the AI Act (*EUR-Lex - Ordinary_legislative_procedure - EN - EUR-Lex*, n.d.). All documents except the ones from the Parliament were retrieved from EUR-Lex. The adopted Opinions of all parliamentary committees dealing with the AI Act were taken from the website artificialintelligenceact.eu. A website reporting on the latest developments related to the AI Act and operated by the Future of Life Institute (FLI) (*Artificial Intelligence Act | Think Tank | European Parliament*, n.d.). As a result, the final corpus is comprised of 29 policy papers. For performing the document analysis, the READ process proposed by Dalglish et al.,(2021) was applied. This process comprises four distinct steps: "(1) ready your materials, (2) extract data, (3) analyse data and (4) distil your findings." (Dalglish et al., 2021, p. 1426). In the initial step, relevant documents are identified guided by the topic, time frame, and the database from which the documents will be retrieved. Subsequently, a predetermined set of categories is established to guide the extraction of data from each document. This includes essential information such as title, author, and date, as well as additional categories that are relevant to the research question. Careful reading of each document allows for the extraction of data to populate these categories. The next step involves analysing the extracted data by identifying patterns, similarities, differences, and other relationships within and across the documents. Finally, the findings are synthesized, often achieved by grouping the analytical categories, utilizing graphs, or, in the case of tracking a policy process, constructing a timeline (Dalglish et al., 2021).

According to the READ process papers were categorized by time and origin before being examined for texts that explicitly addressed open source and texts that might implicitly refer to open source. During this process, papers that did

not address OSAI were screened out and paragraphs that did address OSAI were coded, following an inductive scheme informed by the knowledge gained through the previous literature review regarding OS. The codes were used to compare how the inclusion or exclusion of open source has evolved over time.

4.2 Step 2: Thematic analysis

4.2.1 Sampling

Expectations are constituted through formal mechanisms like research activities, and informal mechanisms including opinion statements and prophesies, no matter their origin once circulated and shared they are performative (Van Lente, 2012) hence, to answer the first research question formal and informal sources were considered. Online published documents were selected based on their ranking and publishing date within the period from April 19, 2021, the start of the AI Act proposal, to May 15, 2023. This selection included informal publications such as news articles and blog posts that specifically related to the AI Act and addressed the inclusion or omission of open source under the Act, and more formal documents, that supported their conclusions by research. The search engine DuckDuckGo was selected for this task due to its impartiality, relying on popularity rankings and query matching rather than personalized search history. Key terms like "AI Act" and "open-source" were employed as search queries, and all relevant findings from the initial two result pages, as they represent the most popular and therefore most circulated findings, were examined, and selected if they presented expectations pertaining to the AI Act and the potential impact of OS implementations. Expectations could be assumptions and statements regarding the impact of the AI Act on the OS community (primary impacts), the resulting impact on business and society (secondary impacts), and expectations for policymakers (see table 2).

Table 2: Content Analysis: List of codes, descriptions, and examples

Code	Description	Example
negative	The overall content argues for excluding or reducing the pressure on developers and vendors of open-source systems to comply.	"Open-source AI models deliver tremendous societal value by challenging the domination of GPAI by large technology companies and enabling public knowledge about the function of AI. The European Council's former approach—exempting open-source AI until it is used for a high-risk application— would lead to far better outcomes for the future of AI."(Engler, 2022)
positive	The content contains mainly arguments for the inclusion of open-source within the scope of the AI Act.	"Research in the field of AI is often open access and open-source as codes are shared and uploaded to platforms such as GitHub, which is accessible and free to use. This gives anyone the opportunity to use new AI techniques and codes for any purpose or practice, also high-risk ones, the latter possibly without any adherence to the requirements for high-risk AI." (Catelijne Muller, LL.M et al., 2022, p. 11).
neutral	The content includes assumptions and statements that argue for and against the liability of open-source developers and providers under the AI Act.	"The road to regulation hell is paved with the EU's good intentions," Etzioni said. "Open source developers should not be subject to the same burden as those developing commercial software. It should always be the case that free software can be provided 'as is …Not every practitioner believes the AI Act is in need of further amending. Mike Cook, an AI researcher who's a part of the Knives and Paintbrushes collective, thinks it's "perfectly fine" to regulate open source AI "a little more heavily" than needed."(Wiggers, 2022)

From there following the theoretical sampling strategy progressively adding new documents by the principle of maximum and minimum contrast (Oliver Dimbath et al., 2018), using sources that advocate for and against the inclusion of open-source. Here it should be highlighted that only few documents arguing explicitly for an inclusion of open-source could be found. The sample size of 20 documents was not decided ex-ante but in course of the data collection by the adequacy "(richness, complexity)" (Braun & Clarke, 2019, p. 28) of the data in line with the concept of information power (Malterud et al., 2016). This process served mainly to identify expert opinions disseminated through news articles and blog posts, which were subsequently utilized for the Thematic Analysis. In this regard, the

examination focused on identifying the individuals or sources referenced within the news articles and blog posts. The referenced documents were then extracted and utilized for the Thematic Analysis (see Table 3).

Table 3: List of TA Samples

Publication Date	Stance	Authorship	Author Classification	Text Type
25.08.2021	Positive	IEEE	technical professionals association	blog post
13.02.2022	Positive	ALLAI	AI research institute	Research report
24.08.2022	negative	Brookings (Alex Engler)	Policy research institute	Blog post
27.09.2022	negative	BSA	Software business association	Open letter
20.10.2022	negative	GitHub	OS provider	Opinion letter
30.11.2022	negative	Mozilla	OS provider	Opinion letter
31.12.2022	negative	LAION	OS provider	Opinion letter
03.02.2023	negative	GitHub (Thomas Dohmke)	OS provider	Opinion letter
11.04.2023	positive	AI Now	Policy research institute	Research report
11.05.2023	negative	Open Future	OS advocacy group	Blog post

4.2.2 Method and Analysis

To analyse reflective thematic analysis was conducted according to the phases described by Braun and Clarke (2021). The familiarization with the data set already started in the content analysis was deepened and initial codes were formed. Once the codes represented a sufficient summary of the content, the individual codes were sorted and clustered to generate initial themes (candidate topics) based on their shared meaning and patterns. An inductive, semantic approach was used. The codes were based on the text's literal, explicitly

expressed message rather than the hidden, latent meaning, and the themes were identified from the content of the data during the process. Subsequently, the candidate themes were reviewed, refined in an iterative process, and synthesized into overarching themes. Finally, to answer the research questions, the themes of positively and negatively classified content were compared and related to the results of the preceding document and content analysis. This was done by matching the timeline of the policy documents with the themes that emerged and the stakeholders who circulated those themes.

5 Findings: Document Analysis

The document analysis has shown that open source was never mentioned explicitly in the first draft of the AI Act proposed by the European Commission in April 2021, nor in the delivered opinion statements of the Parliament or the Council until the inclusion of General-Purpose AI systems. OS was first mentioned by the Committee on Legal Affairs in September 2022 and finally adopted in course of the common position of the Council in December 2022. However, related terms were used that may imply OS, namely "free of charge", "private use", "sole purpose of scientific research and development" and "open data" and were therefore as well included in the examination (see table 4)

Table 4: List of codes, descriptions and examples of the document analysis

Code	Description	Example
research	Articles excluding or including research	This Regulation shall not apply to any research and development activity regarding AI systems.
Open Source	articles mentioning explicitly open-source	(1b) 'open source AI systems' means AI systems, including test and training data, or trained models, distributed under open licenses.
free of charge	Excluding or including AI systems that provided for and monetary return	'making available on the market' means any supply of an AI system for distribution or use on the Union market in the course of a commercial activity, whether in return for payment or free of charge;
open	Text concerning efforts of	experimentation facilities and researchers,

data	establishing or providing open data repertoires	should be able to access and use high quality datasets within their respective fields of activities which are related to this Regulation. European common data spaces established by the Commission and the facilitation of data sharing between businesses and with government in the public interest will be instrumental to provide trustful, accountable and non-discriminatory access to high quality data for the training, validation and testing of AI systems. For example, in health, the European health data space will facilitate non-discriminatory access to health data and the training of artificial intelligence algorithms on those datasets, in a privacy-preserving, secure, timely, transparent and trustworthy manner,
private use	Exclusion or inclusion of private use	6a. This Article only applies to users acting in their professional capacity and not to those using AI in the course of a personal non-professional activity.

In the following section I will first introduce the first proposal of the Commission to provide an overview how the legislation intends to approach AI, followed by a map of the current process and what amendments affecting OS productions were presented by the Council and the Parliament.

5.1 Risk Approach

For AI regulation, the EU has implemented the AI Law, which is based on a risk model with four levels. Those levels include "low-risk systems", "limited- or minimal-risk systems", "high-risk systems", and "unacceptable-risk systems" (Samuel Brown, 2023). The rule of proportionality applies, in the sense that duties and requirements are implemented according to the level of risk. For low or limited risk systems, minimal information must be provided to users or other stakeholders. However, for high-risk systems, providers and distributors must comply with an extensive list of obligations. Systems with unacceptable risks are completely banned in the EU. When artificial intelligence interacts with humans, it must be clearly labelled. (AI ACT, 2021, p. 8).

5.1.1 High-risk AI Systems

High-risk AI systems have the potential to endanger fundamental rights, health, or safety of individuals. This determination is not solely based on the system's design but also its intended application. The act includes an initial list of use cases in Annex III, which will be expanded further. For instance, these use cases encompass systems employed to determine eligibility for public services or loans, establish priority, and make personnel decisions such as promotions and new hires. Additionally, in education, high-risk scenarios include using AI systems to assess students or authenticate identity documents at border controls, make legal or profiling judgments, perform real-time identification using biometric data, and manage infrastructure such as water or electricity supply (ANNEXES 1 to 9, 2021, apps. 1–9).

High-risk AI systems require:

- A risk management system of iterating testing and improvement cycles throughout the lifespan of the system (Article 9);
- Training datasets, adequate in quality and quantity for the intended use and free of any bias (Article 10);
- extensive technical documentation needed to prove compliance with the here mentioned requirements of chapter two of the AI Act. It must be provided for the national authorities for assessment if put into service or placed on the EU market and constantly updated (Article 11);
- Result tracing mechanism through constant creation of logs to monitor the operations of the system (Article 12);
- Comprehensive instructions for use and transparent information on the systems operations and risk assessment (Article 13);

- Integrated mechanisms for human oversights and interventions (Article 14);

- Operational consistency and accuracy as well as strong safety measures against cybercrime or maleficent use (Article 15);

Ensuring the safety and reliability of high-risk AI systems necessitates the provision of appropriate training datasets, meticulous technical documentation, effective result-tracing mechanisms, comprehensive instructions for use, and the integration of human oversight. Furthermore, maintaining operational consistency, utmost accuracy, and robust safety measures against cybercrime are imperative in safeguarding these systems. (AI ACT, 2021, arts. 9–15). Failure to adhere to the prescribed obligations specified in the AI Act for high-risk classified systems may lead to penalties, with fines reaching up to 30 million EUR or, in the case of corporate entities, equivalent to 6% of their total annual revenue. (AI ACT, 2021, art. 71(3)).

5.1.2 Actors addressed along the value chain of AI system dissemination.

Albeit the word open source has not been raised the 21 of April 2021 presented AI Act, OS actors are not excluded by the definitions in Article 3. Therefore I want to outline the definition and the obligation of the most important actors with respect to OSAI systems.

5.1.2.1 Impacts on OSAI providers

The primary onus of compliance lies with the AI system provider, encompassing developers, whether individuals or institutions, who introduce their AI system or its outcomes into the Union market or operate it within the scope of commercial activities, irrespective of whether the AI system is distributed freely or for financial gains (AI ACT, 2021, retical 10). Including free provision within the AI Act extends its applicability to OSAI as well. As providers, OSAI developers are obligated to establish the required safety mechanisms and quality management

systems discussed earlier in this chapter, compile comprehensive technical documentation, provide complete information and instructions to third parties, and bear the expenses associated with conformity assessment and, if necessary, implementing corrective actions in cases of substantial system modifications (AI ACT, 2021, arts. 9–15).

5.1.2.2 Impacts on OSAI distributors

The AI Act defines a "distributor" as any entity that makes an AI system available on the Union market without affecting its properties. (AI ACT, 2021, art. 3(7)). Before placing the AI system on the market, distributors must ensure compliance with AI regulations by verifying its conformity with the instructions for use and technical documentation provided and taking necessary steps to align it with the prescribed requirements. (AI ACT, 2021, art. x) regardless of if such a transaction involves remuneration or not (AI ACT, 2021, art. 3 (10)), which subsequently includes OS platforms like GitHub where development and sharing take place (GitHub, 2022).

5.1.2.3 Impacts on OSAI users

The term "user" in the AI Act means any person or entity that interacts with or uses an AI system, other than private, nonprofessional use (AI ACT, 2021, art. 3(4)). In accordance with the AI Act, users are obligated to employ and oversee the AI system in adherence to the provided "instructions for use" and ensure the input of relevant data. However, users are not obliged to conduct additional risk assessments or undertake risk mitigation measures (AI ACT, 2021, art. 29). In that sense, provider fear being held accountable for unauthorized actions performed by the user, which they could not have foreseen (Paul Keller, 2022). I As per the AI Act 2021 provisions, providers are exempt from their obligations only if the user significantly modifies the AI system or if it is made available within the European Union under the user's name. In such instances, the responsible actor

assuming the role of the provider inherits all associated obligations (AI ACT, 2021, art. 28).

5.1.3 General Purpose AI Systems

The draft of the AI Act incorporates General Purpose Artificial Intelligence (GPAI) systems that "may be used in a plurality of contexts and be integrated in a plurality of other AI systems" with the capacity of "image and speech recognition, audio and video generation, pattern detection, question answering, translation and others" (General Approach, 2022, art. 3(1b)). GPAI systems can also function as high-risk systems, so the provider must adhere to the abovementioned requirements unless any high-risk applications are explicitly excluded in the accompanying instructions and information. The adoption of the General approach in November 2022 was a response to the widespread implementation of foundation AI models. The term "foundation" denotes AI models that have undergone extensive training on large datasets and can be utilized as a fundamental framework for various applications, offering adaptability for various purposes. (*Developing and Understanding Responsible Foundation Models.*, n.d.). The lion's share of the foundation models was developed and distributed as open source. Prime examples of foundational models are Chat GPT-3, BLOOM and BERT.

5.2 AIA Amendment Process

Figure 1 shows all texts submitted by the Commission, the Parliament, the Council, and the official advisory committees that directly and indirectly relate to OS productions. They range from April 21, 2021, the first proposal by the EU Commission, to June 14, 2023, the first reading and the adopted text by the Parliament. The timeline provides an overview of the proposed and adopted legislative paragraphs and definitions, which are described in detail below to contextualize the analysed circulating expectations.

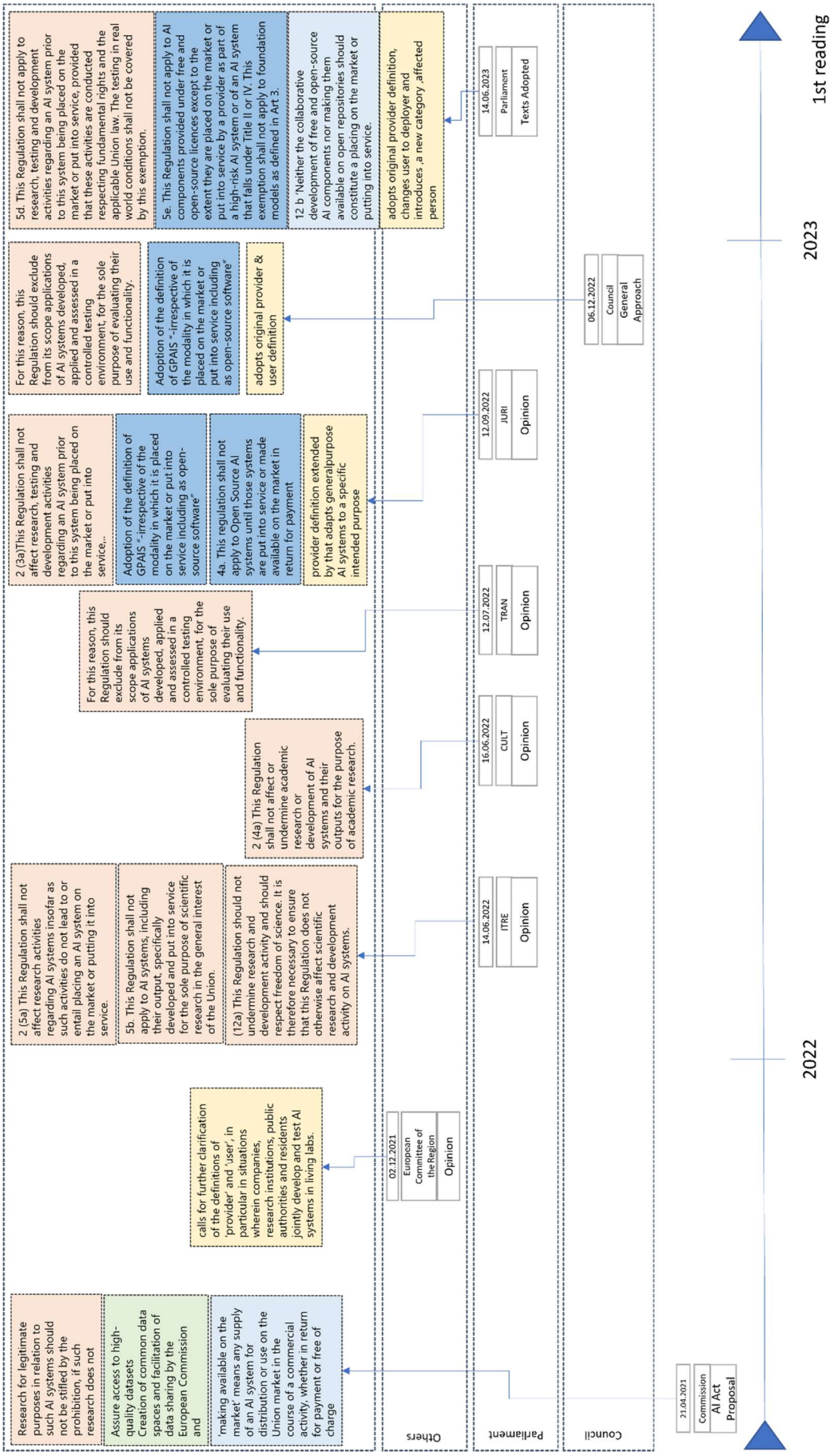

Figure 1: AIA Amendment Process: Texts concerning open-source AI requirements and obligations of the actors

5.2.1 Open-Source

In the adopted opinion of the Committee on Legal Affairs (JURI) opinion of September 12, the European Parliament proposed amendments stating that GPAIs, if open source, should be excluded until they "are put into service or put on the market in return for payment" (Amendment 35 on Article 2, 4a), which includes payment for ancillary services such as technical support. The JURI further extended the definition of provider by including any natural and legal person "that adapts general purpose AI systems to a specific intended purpose." (*Opinion of the Committee on Legal Affairs*, 2022, p. 22) Since OSS licenses offer free use, dissemination, and adaptation the compliance obligation should be passed from the provider to the user or the entity that customized the system and brought it to market. The original provider must cooperate with the new provider to make the AI system compliant, except the provider offers his AI system as a service or accompanying services commercially then the liability and compliance obligations remain with the provider. (Amendment 14, Recital 57b). In course of the JURI opinion OS was also suggested to be incorporated in the definition of a GPAI "irrespective of the modality in which it is placed on the market or put into service including as OSS" (Amendment 36, Article 3, 1a) and specifies in the definition of OSS AI systems that it is attributed not only to the AI models released under the open licenses but also the underlying training and test data (Amendment 37, Article 3, 1b) (*Opinion of the Committee on Legal Affairs*, 2022). Later that year the Council adopted the definition of GPAIs, which includes OSS, in its General Approach. However, it did not adopt the suggested alteration by the Legal Affairs Committee regarding the classification of those modifying GPAIs as providers. Also, the proposal to limit the scope of this legislation exclusively to OSS-based AI systems available for purchase was not accepted (General Approach, 2022). Instead the Council specifies for GPAI systems that "irrespective of whether the general purpose AI system is placed on the market or put into service as a pre-trained model and whether further fine-tuning of the model is to be performed by the user of the general purpose AI

system" (General Approach, 2022, art. 4a (1)) the proposed requirements and obligations shall apply. The adopted text by the Parliament in the first reading explicitly states in Article 2 that "AI components provided under free and open-source license" (Texts Adopted, 2023, art. 2 (5e)) should be excluded from the scope of the legislation. However, GPAIS do not fall under this exemption. Further do the newly introduced recitals provide a more detailed description what placing on the market entails. Collaborative work and use of OSS is not considered "placing on the market" as long as any commercial transaction is omitted, which is defined as any payment for the component or an accompanying service, any data not collected for purely security or functional purposes, or any service. Article 12a also highlights that this exemption has been amended to encourage the adoption, development, and use of AI by startups, SMEs, and individuals (Texts Adopted, 2023).

5.2.2 Private use

Originally, the European Commission's proposal for AI regulations included an exemption for "personal, non-commercial" use of AI in the definition of "user." (AI ACT, 2021, art. 3(4)). Later, the Committee on Transport and Tourism (TRAN) recommended that the definition of "user" be expanded to include persons responsible for the use of AI technology, while removing the earlier exception for personal use. Instead, they added a new provision to the obligations section (Article 9) that would end users' requirements when AI is used solely for personal and private purposes (*Opinion of the TRAN*, 2022).

When the General Assembly reviewed the proposal, it adopted the amended language from the Committee on Transportation and Tourism and moved the exception for personal use to Article 2 under the heading of Scope. Despite this change, the requirement for transparency of AI use remains, including cases involving purely personal and non-commercial uses (General Approach, 2022).

5.2.3 Research

The European Union's proposed AI act emphasized that the Act shall not negatively affect research as long as conducted under the common research ethics, but concerns arose regarding potential negative impacts on research. In line with this sentiment, the Committee of the Regions called for improved definitions of users and providers, especially during joint research efforts among industries, governments, and academia. As late as December 2021, calls were made to ensure that the regulation would not hinder scientific inquiry [1].

By the end of 2022, various parliamentary committees had echoed this call and called for an explicit exclusion of research from the scope of the regulation. The Information Technology, Research, Education, and Agriculture (ITRE) Committee argued that AI systems and their outputs should only fall under the purview of the regulation once they are placed on the market or put into service (*Opinion of the Committee on Industry, Research and Energy*, 2022). Similarly, the Culture and Education (CULT) and Transport and Tourism (TRAN) Committees both sought better delineations and exemptions for research activities to prevent suppression of academic investigation and the development and testing of innovative products (*Opinion of the Committee on Culture and Education*, 2022; *Opinion of the TRAN*, 2022).

In March 2022, the Legal Affairs (JURI) Committee synthesized these submissions, recommending that the regulation should not apply to AI systems or their outputs developed primarily for scientific research until such time that they become available commercially or functionally (*Opinion of the Committee on Legal Affairs*, 2022). Furthermore, the committee believed that member states ought to possess the ability to adapt regulations pertaining to research within the bounds of current EU statutes.

Finally, the European Union’s Institutions -the Council and the Parliament- reached an agreement on the wording of the Artificial Intelligence Act in

November 2022 *(General Approach, 2022),* which directly cites the importance of ensuring freedom for research activities while preserving fundamental rights and privacy. Accordingly, the adopted Act includes clauses specifically targeted at AI systems created exclusively for scientific research and development, thus freeing them from compliance obligations until placed on the market.

2 (6) "This Regulation shall not apply to AI systems, including their output, specifically developed and put into service for the sole purpose of scientific research and development."

2 (7) "This Regulation shall not apply to any research and development activity regarding AI systems."

The adopted text of the Parliament in June 2023 follows a similar trajectory, excluding "research, testing, and development activities" from the scope of the AI Act until they are either brought to market or tested in a real-world setting (Texts Adopted, 2023).

5.2.4 Data

The commission demands providers to ensure training and testing of high-risk Ais is only done on high quality data sets (AI ACT, 2021, art. 10). To support developers in their compliance efforts the Commission also proposes to afford data collection and sharing to create common data spaces for "non-discriminatory access to high quality data" (AI ACT, 2021, p. 29). The danger of possible corrupted datasets and the importance of reliable open data was underlined in the Councils Impact Assessment from the 23rd April 2021 and the Opinion of the Czech Council Presidency on the 14th June 2021 and finally this paragraph got adopted in the General Approach of the Council. While this provision does not specify who gets access to this common data pools, this could affect the open-source community positively by having access to reliable open data sets to work with.

6 Expectations associated with the effect of the AI Act on open-source.

The Thematic Analysis was applied separately for the positive and negative categorized data. Additionally, since the TA aimed to answer the question of what expected impacts were expressed on open-source efforts (1) and regarding attributed benefits to open-source development (2), statements were grouped into (1) primary impacts and (2) secondary impacts before coding.

In the first cluster, as also suggested by Kerr et al. (2020), forward-looking statements with the phrases "would," "could," and "might" (Apha Kerr et al., 2020, p. 6), but also terms such as "will" and "runs the risk to" were primarily considered. On the other hand, secondary effects were also mentioned by listing the beneficial features of open source and suggesting that these will be lost if the development and adoption of OSAI are hindered. The second category, therefore, also includes statements in the present tense. For example

> *Open-source AI is worth protecting. First, open-source AI promotes safety through transparency…;*
>
> *Second, open-source AI promotes competition. Small to medium enterprises across Europe can build on open-source models to drive productivity;*
>
> (LAION, 2023, p. 1)
>
> *To that end, we respectfully urge the European Parliament to:*
>
> *1. Ensure that open-source R&D can reasonably comply with the AI Act.*
>
> (LAION, 2023, p. 2)

The thematic analysis of negative categorized expectations arguing for the partial or total exclusion of OS from the scope of the AI ACT revealed two prominent themes in the primary effects cluster and four in the secondary effects cluster with multiple codes, presented in Table 5. Overall, the impact of the AI ACT on open-source developments is expected to stifle or completely halt them. The themes are structured around the inquiry into why compliance obligations have this effect on open-source productions. One theme consisted of statements

expressing the belief that open-source actors' requirements are *difficult or unattainable*. The other theme emphasized the idea that the open-source community would become *demotivated* to share their AI components or models due to the burden of compliance. Although both themes are closely related, a distinction is made since the first theme posits that open-source developers are typically individuals with limited resources who contribute to open-source projects without financial incentives. In the second cluster, the most common theme mentioned by six different stakeholders was *dependence* non-European developers or well-resourced BigTech companies as AI providers and the associated room for manoeuvre. That would emerge if AI research, development and dissemination progress is stifled. Three stakeholders also emphasized the importance of *safe & secure AI* that might be lost if open-source as a transparent provider or checking mechanism on commercial AI is hampered, directly connected to *common goods* as in public knowledge and public-interest oriented AI, which go hand in hand with the former.

Table 5: Positive category; List of themes, codes and exemplary excerpts

Cluster	Themes	Codes	Examples
primary effect	Difficult & impossible	No means to comply	"consider the case of a single student developing an AI capability; they cannot afford to comply with EU regulations and may be forced not to distribute their software
		Very difficult to comply	"and this means that open source developers in the EU will face a significant regulatory burden that will be very difficult to comply with"
	demotivated	Disincentivize R&D	"the Council risked creating a chilling effect on the open source development of foundation AI models/systems"
		Disincentivize sharing	"This would disincentivize publishing open-source AI models"
secondary effect	dependence	Well-resourced corporations	"open-source models to drive productivity, instead of relying on a handful of large firms for essential technology."
		Foreign providers	"Europe may cross a point-of-no-return, falling far behind in AI development, and being relegated to a consumer role without its own decision-making on critical technologies that will shape our societies."
	progress	Hamper AI dissemination	"open-source R&D is essential to safely develop, study, and deploy large AI models in Europe"
		Hamper AI innovation	"if you already place overly heavy burdens on openly released features at the top of the AI innovation

			stream you risk hindering incremental innovation, product differentiation and dynamic competition"
	public good	hamper public interest-oriented AI	"open-source R&D is essential to safely develop, study, and deploy large AI models in Europe, and to ensure these technologies serve the public interest."
		hamper public knowledge creation	"Second, they enable critical research, and thus public knowledge, on the function and limitations of GPAI models."
	safe & secure AI	Risk less safe AI in EU	"Open-source AI is worth protecting. First, open-source AI promotes safety through transparency"
		Hamper AI transparency	"allowing for more open-source GPAI provides more transparency in their development."
		Hamper fair and trustworthy AI	"this ability of open source development to serve as a check on corporate closed-source practices has become increasingly clear."

Positive categorized text, also due to the small sample size, has much lesser themes and codes, and some codes are only found once. The themes detected in the primary were the worries of *less safe AI models,* specifically for OS GPAIs or OS codes used for research, and to *hamper downstream development* of OS but also proprietary AI models if the source models get excluded. This might translate into secondary effects similar found in the positive category, like *power asymmetry* between original providers and downstream providers, as the former would be off the hook and the latter would depend on their benevolence to comply with the AI ACT. The most prominent topic was *safe & equitable AI.*

Table 6: Negative category; List of themes, codes and exemplary excerpts

Cluster	Themes	Codes	Examples
Primary effect	less safe source models	Design decision	"data and design choices made at the stage of original development would result in both unfair and ine~ective regulatory outcomes"
		loopholes	"would have to make sure the high-risk AI system is in line with the AIA, but this is almost in all cases merely a self-assessment obligation."
secondary effect	power asymmetry	Dependence to comply on big tech	"it places the latter in a fully dependent position, without having the appropriate means to seek redress when the general purpose AI system causes damage"
		AI adoption	"As a result, it could lead to a limited uptake of general purpose AI systems on the one hand, and a (further) concentration of AI innovation with general purpose AI developers on the other"
	safe & secure AI	amplifying risk	"The fact that these models are being deployed at scale means that they can become singular points of failure that can radiate harms (e.g., security risks, inequities) to countless downstream AI applications."
		Unintended unsafe	"One of the scary parts of open-source AI is how intensely easy it is to use," he says. "The barrier is so

	& unethical use	low… that almost anyone who has a programming background can figure out how to do it, even if they don't understand, really, what they're doing."

7 Discussion

Overall, very noticeable was the strong connection between open-source and General Purpose AI. 10 out of 11 mentioned GPAIs or foundational models in connection with open-source. While some authors like LAION and Mozilla mentioned a distinction between GPAIs being open-source and those being only open-access but proprietary, others were discussing GPAIs only in the light of open source, conveying the notion that all or most of these models are open-source, e.g.:

> *"While some major technology companies open-source their models, such as Google's BERT or Open AI's GPT~2, the corporate incentives to release these models will diminish over time as they become more commercialized".*

(Engler, 2022)

> *providers of systems like open source large language models (e.g. GPT-3)*

(Wiggers, 2022)

As already pointed out in the literature review are, the LLMs models GPT3 and four only freely accessible for the end users but do not have access to the codes or core models. However, the above-shown example from Brookings CEO Albert Engler might mislead the reader to conclude that most foundational models are open-source. The "misuse of openness" and the association of GPAIs to open-source were also already noted and criticized by AINow (Amba Kak & Sarah Myers West, 2023). Therefore, some of the statements considered in this TA relate directly to GPAIs but have only been incorporated into the analysis when the author considers the inclusion of open-source models or the association of GPAIs with open-source developments. Such a case, for example, is the analysis reports by AINow (2023) and ALLAI (2022), which call for including all GPAIs in the AI Act regardless of origin. This effect also works vice versa; not only can it

confuse the audience that all GPAIs are open-source productions, but all open-source productions must comply with the same obligations as GPAIs.

7.1 Expectations regarding the AI Act and its impact on open-source efforts

Expectations circulated to legitimize partial or complete exemption of AI from the scope of the AI Act highlighted the stifling effect on OSAI productions and are attributable to the high compliance burden. Either developer may get demotivated to share their projects, or obligations imposed by the AI Act exceed the available resources at the developer's disposal. The second assertion used the narrative that open-source developers are typically private individuals who engage in OS projects solely driven by intrinsic motivation, and limited means. However, it is crucial to acknowledge certain aspects that were consciously or unconsciously overlooked in shaping this perspective. Firstly, the extensive requirements exclusively apply to high-risk AI systems. GPAIs in the Council's General Approach would be classified as such, but as already mentioned above, the notion conveyed that all open-source productions are universally affected. Secondly, developers can indeed establish lucrative businesses based on their open-sourced code. Lastly, it is essential to note that numerous contributors, particularly in the United States, are well-resourced information and communication technology (ICT) corporations. Connecting this to the sociology of expectations, do these informal expectations exclude critical reflections and are, therefore, from a constructivist perspective, less robust (Van Lente, 2012)

On the other hand, the research organizations AINow and ALLAI, advocating for compliance obligation specifically for OS GPAIs, emphasize that currently, only well-resourced developers can train and run such power-intensive models, as already reflected in the literature review. If they were excluded, less safe AI models could be the result, as design decisions of architecture and training data taken by the original provider would not have to consider the AI Act and significantly impact the AI systems built upon it. Positive assertions attributed to open source outlined in chapter 3.4, like constant peer revision and improvement

of code, full transparency through accessibility and rigorous documentation on authorship and versioning, which also lead to safe and equitable AI, were in this discussion not reflected. However, if OSAI had these expected characteristics, it would already be largely conformable with the high risk-requirements (see chapter 7.1.1 high-risk systems). One author stated:

> *"The latest draft proposals include a get-out clause for developers, they simply need to include a usage limitation in the instructions for use. Frankly, it would be unusual to download open-source software without legal documentation already being included – so this is an extra line or two in there – not a huge burden upon developers."*

(Adam Leon Smith quoted by Morrison, 2023)

There are already examples of open-source products, such as the Logj4 (Korn, 2021), that have not lived up to these expectations, therefore the inclusion statement made by the proponents of OS is more robust than that of the opponents.

7.2 Expectations regarding the AI Act and its impact on characteristics attributed to open-source development

7.2.1 Progress and Power Relations

The detected secondary effects shed light on the statements used by the authors to underscore the significance of preventing the anticipated primary effect. In the case of the negative cluster, the expected inhibition of OSAI is foremost connected with *progress* and *dependencies* on BigTech and foreign providers, which go hand in hand. Progress is expected to be hampered on several levels encompassing research, innovation, development, and adoption of AI. That resonates with a widely shared perception of OS development being integral to those activities. A surfaced document in May 2023 written by an Alphabet worker discloses that the open-source community is demonstrating remarkable agility and efficiency in tackling complex issues and coming up with solutions at speed exceeding what large companies like Alphabet or OpenAI can achieve through their limited networks of experts working independently (Patel & Azfal Ahmad,

2023). While this opinion statement with an unconfirmed source is an informal mechanism to shape expectations, formal research reports on OSS support these statements and show, especially amongst SMEs, OSS uptake is very high (European Commission. Directorate General for Communications Networks, Content and Technology., 2021; Open:UK, 2022; RedHat, 2022). UK:Open reports based on a study conducted in 2022 that 100% of UK businesses established in the last three years are using OSS (Open:UK, 2022).

Prominent players in the open-source community, including LAION, Mozilla, GitHub, as well as policy researchers Paul Keller of Open Future (EU) and Alex Engler of Brookings (USA), have highlighted the importance of open-source artificial intelligence (OSAI) in enabling SMEs and startups to close the gap with large corporations, particularly in the information and communications technology (ICT) sector. LAION founded in Germany, and Thomas Dohmke from Germany, explicitly emphasized that hindering the spread of OSAI within the European Union would lead to further falling behind in global competition in artificial intelligence. This statement alingns with the picture provided when considering the number of new patents and research publications mainly stemming from US and China, with IBM, Microsoft and Toshiba being the leaders in AI-related patent applications. (Irene Kitsara, 2022) Moreover, it resonates with EUs concerns and efforts to establish digital sovereignty(Tambiama, 2020). The competition with US and China was already detected by Kerr et al. (2020) as an effective expectation to legitimize investments in costly AI research in Europe. What is, however, not conveyed in the voiced statements concerning progress and dependence is that big ICT companies are profiting from the open-source community just the same. The anonymous author, an Alphabet employee, sums this up nicely in his conclusion about Meta's recent fully open-sourced LLaMA:

> *"Paradoxically, the one clear winner in all of this is Meta. Because the leaked model was theirs, they have effectively garnered an entire planet's worth of free labor. Since most open source innovation is happening on top of their architecture, there is*

nothing stopping them from directly incorporating it into their products. The value of owning the ecosystem cannot be overstated."

(Patel & Azfal Ahmad, 2023)

Additionally, the opposing voices suggest quite a contrary view on dependence. If OS providers get exempt and produce unsafe and inequitable AI to the AI ACT standards, it would fall on the downstream provider to comply with the AI Act making and, subsequently, putting the downstream developer under the dominant power of the original one. The former would have to rely on the latter's cooperation to provide the necessary documentation and security of the core model to bring the customized versions into compliance with the requirements for high-risk systems. In addition, the original developer cannot be held legally accountable for its created AI system. In such a case, AINow expects downstream developers to be discouraged from using and customizing such systems. Thus, in the context of GPAIs, SMEs would be deprived of the opportunity to innovate on such systems and leave the field to the original provider, i.e. mainly BigTech.

The statement that a lack of OSAI systems also slows progress in research and development is also shared by several formal sources. Hence, this statement has a high degree of robustness. However, the relationship between progress and predicted dependency is less robust, as other experts believe that this predicted progress could be used equally, if not better, by established ICT companies. Several facts also indicate that private developers and small companies would continue to depend on the major players because they own the ecosystem.

7.2.2 Safety and Security

Both parties argued in their communication that regulating or not regulating OS under the AI Act would negatively impact safe and fair AI. The parties that wanted to exclude open-source referred to the transparencies of open-sourced systems and the peer revisions this allows. This statement aligns with the common practice of rigorous open-source documentation, which allows

transparency about what was changed, when it was changed and by whom, and includes the author's contact details for accountability (Perrin 2021) and the security and quality of code achieved through forking. (Knut Blind et al., 2021). Further, a 2021 study by Red Hat showed that more than 80% of IT leaders surveyed worldwide rated open source as equal to or more secure than proprietary (RedHat, 2022). While the statements about security and safety are robust in that many formal studies examine the positive impact of open-sourcing systems on these two attributes, it can be misleading to the reader and create the expectation that open-source systems provide 100% security. The case of log4j, a widely used OSS component, highlights the security issues with OS. When some security vulnerabilities were discovered, log4j was already used in many downstream OS and proprietary programs. Hence, the positive cluster's argumentation seems again more robust, highlighting the potential escalation of security vulnerabilities if the freely accessible and low-barrier baseline model becomes widespread. In this scenario, any existing problems with the source model could become more widespread.

The anticipated decline in open-source production was also linked to a negative impact on civil society. The AI Act in such a form as proposed in the General Approach would have the effect that AI developments and research would not be oriented to the needs of the public, and thus essential knowledge assets and public-oriented AI applications would be missing in the future.

The latter two themes of common good and security are reminiscent of the expectations regarding ethical issues related to AI, which have been promoted in recent years by the EU's strategic communication and can be connected to Kerr et al.'s (2020) findings. As they describe that the EU’s used specific terminologies like “good and for all” and “transparency and accountability” were found in the strategies and communication of consulting firms and businesses. It appears that proponents against the inclusion of open source are targeting the EU's strategic points with their arguments, creating the narrative that by legislating OSAI, the EU would be acting contrary to its stated strategy.

7.3 The reflection of circulating expectations in the AI Act

The objective of this chapter is to determine whether the amendments adopted by legislative bodies align with the expectations that have been circulated (RQ3). It is important to emphasize that correlation does not imply causation, but this study aims to highlight to what extent the formulated future visions correspond with the amendment process.

In the Commission's original proposal (AI ACT, 2021) of April 2021, OS are not literally mentioned. However, the wording of the definition of placing on the market and thus liability under the AI ACT is ambiguous and could be interpreted to apply to open-source systems. OS could only be clearly excluded if they are used exclusively in the context of research and development activities or if they are applied only for private use. In 2021, there was no media coverage of discussions of open source in the context of the AI Adt, except the statement by the IEEE professional association calling for the inclusion of OSAI as early as August 2021.

In response to the Slovenian Presidency, which advocated the exclusion of GPAI, ALLAI, whose members are also represented in the EU High-Level Expert Group on Artificial Intelligence, drafted a statement on these AI systems. That statement does not explicitly mention that open source should be included, but it does call for no GPAI system to be excluded, even for research. Furthermore, it denounced the freedom granted to researchers and claimed that freedom in research should not be absolute and should be regulated under the AI Act. Only after the JURI opinion of 12.09.2022 (see Figure 1) is there no indication that this view is reflected in any of the opinions adopted by the Parliament or the Council. The JURI opinion for the first time GPAIs are mentioned in an officially adopted document of a Parliamentary body and includes for the first time OSAI.

In the same year, while the inclusion of GPAIs was still under discussion, the probably most prominent proponent of OS exemption made first headlines with his blog post in August. The JURI aligned partially with his view, as they

included OS GPAIs in the law text but explicitly excluded any other OSAI. This decision is supported by a recital highlighting the contribution of OS in research and economic contribution in Europe. Other OS exclusion advocates later that year can be divided into those that sought to exclude OSAIs from compliance altogether, like BSA and those that called for excluding OSAIs under certain conditions or reducing obligations for OS actors; the latter was mainly OS providers. However, non of their views seemed to be reflected in the adopted position of the Council, as OSAI is only mentioned to ensure the inclusion of all GPAIs.

AINow, which published its report in April, refers to the negative statements made concerning OS and denounces the misuse of the term "open" and, concerning the GPAI, emphasizes that even if they would meet the original definitions of open source, their size still makes them copyable and executable only for companies with extensive resources. It further reiterates ALLAI's warning that excluding GPAIs under any pretext would jeopardize European security and innovation if the end user were held solely responsible for meeting the systemic requirement. The Parliament represents A similar view in the finally adopted Text on June 14th 2023. Extensive requirements are introduced for GPAIs (now referred to as foundational AI models in the legislative text), and, most importantly, providers (meaning developers) are required to bring the system into compliance. In the Commission's original text, the provider would no longer be responsible as soon as a downstream user changes his AI system. In addition, it includes a more precise definition of open-source that incorporates the licensing structure as a determinant of the definition of OSAI systems. Also, the suggestion that research should not be completely free of restrictions is partially reflected in the adopted text of the first reading, as it now adds that the exception for research does not apply under real-world conditions.

However, the importance of promoting the development of OSAIs is also reflected in the text outlining the positive impacts on security and progress like the JURI. Furthermore, these expectations are also performative and effective as

the Parliament introduces provisions on the development of OSAIs to exclude them from the scope of the AI Act, except for certain exceptions.

All in all, the importance of open source developments is reflected in the most recently adopted text, but only to the extent that it does not run counter to the expectations of the parties that point to the risks of an OSAI exclusion.

8 Conclusions

In summary, this study shows that open source was not explicitly considered in the position statements of the EU legislative bodies until September 2022 in the adopted position of the JURI. However, by not mentioning it, the legal text can be interpreted in a way that the legal text does concern OSAIs but also the other way around. A disagreement of interpretation, in this case, could arise mainly from the definition of "making available on the market". Since "commercial activity" against payment or free of charge was not further specified, and several business models of open source SaaS or code maintenance by the community generate value, however, as in the former case, not directly by the offered AI model or as in the second case, no monetary value. (AI ACT, 2021) However, users in general, and therefore also users of OSAIs, would be excluded from regulation in the case of private use, just as OSAIs used for research. Finally, with the consideration of GPAIs, open-source GPAIs were also explicitly included and obliged to comply to the high-risk requirements. However, they only briefly mentioned in the added article concerning GPAIs. In the last adopted version of the AI Act, a more detailed description of open-source developments has been added. In the most recent iteration of the AI Act, which has been formally accepted, there is now a more comprehensive definition of open-source developments. It specifies that OSAI is generally not encompassed within the provisions of the AI Act unless it pertains to highly classified models carrying significant risks or falls under the classification of foundational models as outlined in the final text.

A notable aspect was that the inclusion of open-source within the liability scope was limited to GPAI models. Commentators who advocated excluding OSAIs from the AI Act often framed it as if all OSAIs were covered under the Act's provisions and subsequently created the image that all open-source activities would be negatively affected. Further was a narrative created that it is the open-source community against BigTech, and subsequently, by hampering open collaboration BigTech would become more powerful. In contrast, the general public would become dependent on BigTech, without OSAI as essential knowledge creation and innovation source. Interestingly supporters of OSAI inclusion utilized a similar narrative, portraying all GPAIs, including OS, as produced by wealthy corporations. As compliance burdens would be moved to actors further down the value chain, small providers and the general public would lose the GPAIs to innovate upon them. Both parties provided a future picture of power imbalances that could arise between small and large providers and between Europe and the leading international countries in AI research and development if OSAI was not according to their suggestion

Furthermore, this study shows that concerns about expected ethical implications were mainly addressed from the user's side. While opponents of open-source regulations highlighted the characteristics of transparency and accessibility to translate into safety and security, proponents used the same arguments to highlight the danger of scalable risks. Ethical and sustainable implications arising from training, specifically GPAIs, were addressed by neither side except one proponent when voicing their expectations.

The results also suggest that expressed expectations in favour of including OSAI were more robust and more widely reflected in the AI Act, despite being less widespread. The fact that they circulate less, could be problematic, especially if we consider Van Lente's (Van Lente, 2012) theory of the legitimacy of strategic foresight. If the justifications for particular pieces of legislation lack the

reinforcing backbone of a jointly developed vision of the future, EU legislative bodies could face resistance from member states when enforcing the law.

This study was limited in its sample size and reflects only the expressed expectations of experts who disseminated their statements via blog posts and online news articles. However, it does not show how these expectations were adopted and disseminated by the civilian population of the EU. Therefore, this study can only indicate what future images are used and how they are reflected in the adaptations of the AI Act but further research has to be done to compliment this picture.

LIST OF TABLES

REFERENCES

About. (2006, September 19). Open Source Initiative. https://opensource.org/about/

Ajeesh, A. K., & Rukmini, S. (2023). Posthuman perception of artificial intelligence in science fiction: An exploration of Kazuo Ishiguro's Klara and the Sun. *AI & SOCIETY*, *38*(2), 853–860. https://doi.org/10.1007/s00146-022-01533-9

ALGORITHM WATCH (Ed.). (2023). *SustAIn*. *2*. https://algorithmwatch.org/en/wp-content/uploads/2023/03/SustAIn-magazine-issue-2.pdf

Allan M. Turing. (1950). Computing Machinery and Intelligence. *The Philosophy of Artificial Intelligence*, 41–66.

Amba Kak & Sarah Myers West. (2023). *Confronting Tech Power* (2023 Landscape). AINow. https://ainowinstitute.org/wp-content/uploads/2023/04/AI-Now-2023-Landscape-Report-FINAL.pdf

Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, B., De, P. G., Feijoo, C., Fernandez, M. E., Gomez, G. E., Iglesias, P. M., Junklewitz, H., Lopez, C. M., Martens, B., Figueiredo, D. N. S., Nativi, S., Polvora, A., Sanchez, M. J. I., Tolan, S., Tuomi, I., & Vesnic, A. L. (2018, December 3). *Artificial Intelligence: A European Perspective*. JRC Publications Repository. https://doi.org/10.2760/11251

Apha Kerr, Marguerite Barry, & John D Kelleher. (2020). *Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance*. https://doi.org/10.1177/2053951720915939

Artificial intelligence act | Think Tank | European Parliament. (n.d.). Retrieved April 21, 2023, from https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792

Arturo Rosenblueth, Norbert Wiener, & Julian Bigelow. (1943). Behavior, Purpose and Teleology. *Philosophy of Science*, *10*, 18–24.

Balka, K., Raasch, C., & Herstatt, C. (2010). How Open is Open Source? - Software and Beyond: HOW OPEN IS OPEN SOURCE? *Creativity and Innovation Management*, *19*(3), 248–256. https://doi.org/10.1111/j.1467-8691.2010.00569.x

Berkhout, F. (2006). Normative expectations in systems innovation. *Technology Analysis & Strategic Management*, *18*, 299–311. https://doi.org/10.1080/09537320600777010

BigScience Research Workshop. (n.d.). Retrieved July 5, 2023, from https://bigscience.huggingface.co/

Borup, M., Brown, N., Konrad, K., & Van Lente, H. (2006). The sociology of expectations in science and technology. *Technology Analysis & Strategic Management*, *18*(3–4), 285–298. https://doi.org/10.1080/09537320600777002

Braun, V., & Clarke, V. (2019). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, *13*(2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

Bruce Malzlish. (1994). The Animal-Machine. In *The Fourth Discontinuity: The Co-Evolution of Humans and Machines* (pp. 15–31). Yale University Press.

Budde, B., & Konrad, K. (2017). *Governing Fuel Cell Innovation in a Dynamic Network of Expectations*.

Byrd, J., & Paquette, P. (2023). Frankenstein: A creation of artificial intelligence? *AI & SOCIETY*, *38*(1), 331–342. https://doi.org/10.1007/s00146-021-01298-7

Catelijne Muller, LL.M, Prof. Virginia Dignum, Maria Avramidou, & Mónica Fernández Peñalver. (2022). *AIA in-depth: 1 Objective Scope Definition*. ALLAI.

Chopra, S., & White, L. F. (2011). Artificial Agents and Agency. In *A Legal Theory for Autonomous Artificial Agents* (pp. 5–28). University of Michigan Press. https://doi.org/10.3998/mpub.356801

Corporate Europe Observatory. (2022, April 23). *Big Tech's last minute attempt to tame EU tech rules: Lobbying in times of trilogues*. Corporateeurope.Org. https://corporateeurope.org/en/2022/04/big-techs-last-minute-attempt-tame-eu-tech-rules

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts—General approach, 14954/22, 2021/0106(COD) (2022).

Dalglish, S. L., Khalid, H., & McMahon, S. A. (2021). Document analysis in health policy research: The READ approach. *Health Policy and Planning*, *35*(10), 1424–1431. https://doi.org/10.1093/heapol/czaa064

Developing and understanding responsible foundation models. (n.d.). Stanford CRFM. Retrieved April 11, 2023, from https://crfm.stanford.edu/

Digital sovereignty for Europe. (n.d.).

DIVISION E—NATIONAL ARTIFICIAL INTELLIGENCE INITIATIVE ACT OF 2020, DIVISION E, SEC. 5001 (2021). https://www.ai.gov/#:~:text=The%20National%20AI%20Initiative%20Act%20of%202020%20%28DIVISION,for%20the%20Nation%E2%80%99s%20economic%20prosperity%20and%20national%20security.

Ekkehard Ernst, A. (2022). Artificial Intelligence: Productivity Growth and the Transformation of Capitalism. In *Platforms and Artificial Intelligence: The Next Generation of Competences* (pp. 149–183). Springer International Publishing. https://doi.org/10.1007/978-3-030-90192-9

Engler, A. (2022, August 24). The EU's attempt to regulate open-source AI is counterproductive. *Brookings*.

https://www.brookings.edu/blog/techtank/2022/08/24/the-eus-attempt-to-regulate-open-source-ai-is-counterproductive/

Eric Raymond. (2002). The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary. *Choice Reviews Online*, *39*(05), 39-2841-39–2841. https://doi.org/10.5860/CHOICE.39-2841

EUR-Lex—Ordinary_legislative_procedure—EN - EUR-Lex. (n.d.). Retrieved May 10, 2023, from https://eur-lex.europa.eu/EN/legal-content/glossary/ordinary-legislative-procedure-codecision.html

European Commission. (2021). *ANNEXES to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*.

ANNEXES to the Propoasl fo a Regulation of the European parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, no. COM(2021) 206 final (2021). https://artificialintelligenceact.eu/annexes/

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLICENCE (ARTIFICIAL INTELLICENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, no. COM(2021) 206 final, EUROPEAN COMISSION (2021). https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

European Commission. (2018a, December 7). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS*. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0795

European Commission, D. F. (2018b). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Artificial Intelligence for Europe* (COM(2018) 237 final).

European Commission. Directorate General for Communications Networks, Content and Technology. (2021). *The impact of open source software and hardware on technological independence, competitiveness and innovation in the EU economy: Final study report.* Publications Office. https://data.europa.eu/doi/10.2759/430161

European Commission, Joint Research Centre (European Commission), Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). *AI watch – Defining artificial intelligence: Towards an operational definition and taxonomy of artificial intelligence*. Publications Office of the European Union. https://data.europa.eu/doi/10.2760/382730

European Parliament. (n.d.). *Proposal for a Regulation on a European approach for Artificial Intelligence | Legislative Train Schedule*. European Parliament. Retrieved May 10, 2023, from https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence

Opinion of the Committee on Culture and Education on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), (2022) (testimony of European Parliament).

Opinion of the Committee on Industry, Research and Energy on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), (2022) (testimony of European Parliament).

Opinion of the Committee on Legal Affairs on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), (2022) (testimony of European Parliament).

Opinion of the Committee on Transport and Tourism on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), (2022) (testimony of European Parliament).

Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1, P9_TA(2023)0236, first reading (2023).

Ex Machina (film). (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Ex_Machina_(film)&oldid=1162388335

F Indoannidou & V Konstantikaki. (2008, December). Empathy and emotional intelligence: What is it really about? *International Journal of Caring Science*, *1*(3), 118–123.

Free Software Foundation. (n.d.). *Free software is a matter of liberty, not price—Free Software Foundation—Working together for free software*. Free Software Foundation. Retrieved May 5, 2023, from https://www.fsf.org/about/

GitHub. (2022). *EU AI Act—GitHub Position Paper*. https://github.blog/wp-content/uploads/2023/02/GitHub_Position_Paper-AI_Act.pdf

Her (film). (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Her_(film)&oldid=1161985843

Hermann, I. (2023). Artificial intelligence in fiction: Between narratives and metaphors. *AI & SOCIETY*, *38*(1), 319–329. https://doi.org/10.1007/s00146-021-01299-6

High Performance, High Frequency, Bare Metal, Affordable Cloud Computing. (n.d.). Vultr. Retrieved July 5, 2023, from https://www.vultr.com/pricing/

High-Level Expert Group on Artificial Intelligence. (2019). *A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES*.

Irene Kitsara, A. (2022). Artificial Intelligence and the Digital Divide: From an Innovation Perspective. In *Platforms and Artificial Intelligence: The Next Generation of Competences* (pp. 245–265). Springer International Publishing. https://doi.org/10.1007/978-3-030-90192-9

Jääsaari, J. (2007). *Consistency and change in Finnish broadcasting policy: The implementation of digital television and lessons from the Canadian experience*. Åbo Akademis Förlag.

James Moor. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazin*, *27*(4).

Julia Zorthian. (2023, February 18). *Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer*. Time. https://time.com/6247678/openai-chatgpt-kenya-workers/

Karppinen, K., & Moe, H. (2012). *What We Talk about When We Talk about Document Analysis*. https://doi.org/10.2307/j.ctv36xvj36.12

Klaus Schwab. (2023, May 31). *The Fourth Industrial Revolution*. Encyclopedia Britannica. https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734

Knight, W. (2023, March 27). OpenAI's CEO Says the Age of Giant AI Models Is Already Over. *Wired*. https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/

Knut Blind, Mirko Böhm, Paula Grzegorzewska, Andrew Katz, Sachiko Muto, Sivan Pätsch, Torben Schubert, European Commission, & Directorate General for Communications Networks, Content and Technology. (2021). *The impact of open source software and hardware on technological independence, competitiveness and innovation in the EU economy: Final study report.* Publications Office. https://data.europa.eu/doi/10.2759/430161

Korn, J. (2021, December 15). *The Log4j security flaw could impact the entire internet. Here's what you should know | CNN Business*. CNN. https://www.cnn.com/2021/12/15/tech/log4j-vulnerability/index.html

LAION. (2023, April 28). *LAION in Open Letter to European Parliament Urge Call to Protect Open-Source AI in Europe—Unite.AI*. https://www.unite.ai/laion-and-a-group-of-27/

Le Scao, T., Workshop, B., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., Gallé, M., Tow, J., Rush, A. M., Biderman, S., Webson, A., Ammanamanchi, P. S., Wang, T., Sagot, B., Muennighoff, N., … Wolf, T. (2023). *BLOOM: A 176B-Parameter Open-Access Multilingual Language Model* (arXiv:2211.05100). arXiv. http://arxiv.org/abs/2211.05100

Li, L., Zheng, N.-N., & Wang, F.-Y. (2019). On the Crossroad of Artificial Intelligence: A Revisit to Alan Turing and Norbert Wiener. *IEEE Transactions on Cybernetics*, *49*(10), 3618–3626. https://doi.org/10.1109/TCYB.2018.2884315

MacKenzie, D. A. (2006). *An engine, not a camera: How financial models shape markets*. MIT Press.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qualitative Health Research*, *26*(13), 1753–1760. https://doi.org/10.1177/1049732315617444

McKinsey&Company (Ed.). (2018). *Notes from the AI Frontier: Modeling the impact of AI on the world economy*. McKinsey Global Institute.

Michael Atleson. (2023, February 27). *Keep your AI claims in check*. Federal Trade Commission. https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, *33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Morrison, R. (2023, April 19). EU to regulate 'general purpose' AI like ChatGPT. *Tech Monitor*. https://techmonitor.ai/technology/ai-and-automation/eu-ai-regulation-chatgpt

Nicolas Kayser-Bril. (2021, November 23). *European Council and Commission in agreement to narrow the scope of the AI Act*. AlgorithmWatch. https://algorithmwatch.org/en/eu-narrow-scope-of-ai-act/

Oliver Dimbath, Michael Ernst-Heidenreich, & Matthias Roche. (2018). *Praxis und Theorie des Theoretical Sampling. Methodologische Überlegungen zum Verfahren einer verlaufsorientierten Fallauswahl*. *19*(3), Art 34.

Open letter on the proposed regulation of artificial intelligence. (2022, 11). [Letter to Ivan Bartos, Josef Sikela, & Mikulas Bek]. https://www.spcr.cz/images/Open_letter_on_the_proposed_regulation_of_artificial_intelligence_FIN20221107_125114.pdf

OpenAI. (2023). *GPT-4 Technical Report* (arXiv:2303.08774). arXiv. http://arxiv.org/abs/2303.08774

Open:UK. (2022). *State of Open: The UK in 2022 Phase Two: UK Adoption*.

Oxford Analytica (Ed.). (2023). EU AI law faces tough conflicting pressures. *Emerald Expert Briefings*, *oxan-db*(oxan-db). https://doi.org/10.1108/OXAN-DB279356

Patel, D. & Azfal Ahmad. (2023, May 4). *Google "We Have No Moat, And Neither Does OpenAI."* https://www.semianalysis.com/p/google-we-have-no-moat-and-neither

Paul Keller. (2022, December 13). *How will the AI Act deal with open source AI systems?* Open Future. https://openfuture.eu/blog/how-will-the-ai-act-deal-with-open-source-ai-systems

Perrin, A. S. (2021). THE GENERAL DATA PROTECTION REGULATION AND OPEN SOURCE SOFTWARE COMMUNITIES. *Cybaris: An Intellectual Property Law Review*, *12*(1), 77–108.

POLITICO Europe (Director). (2022, April 21). *Panel discussion - The AI package, one year on: From theory to practice | POLITICO AI & Tech Summit*. https://www.youtube.com/watch?v=uvaUQg_q2Ho

Press conference by Roberta METSOLA, EP President, Brando BENIFEI and Dragoş TUDORACHE, rapporteurs on the AI Act plenary vote. (2023, June 14). https://multimedia.europarl.europa.eu/en/webstreaming/press-conference-by-roberta-metsola-ep-president-brando-benifei-and-dragos-tudorache-rapporteurs-on_20230614-1400-SPECIAL-PRESSER

RedHat. (2022, February 22). *The State of Enterprise Open Source 2022*. https://www.redhat.com/en/enterprise-open-source-report/2022

Safety standards. (n.d.). Openai.Com. Retrieved July 3, 2023, from https://openai.com/safety-standards

Samuel Brown. (2023, January 16). *The EU AI Act's Risk-Based Approach: High-Risk Systems and What They Mean for Users*. Futurium. https://futurium.ec.europa.eu/en/european-ai-alliance/document/eu-ai-acts-risk-based-approach-high-risk-systems-and-what-they-mean-users

Sawers, P. (2023, February 3). GitHub CEO on why open source developers should be exempt from the EU's AI Act. *TechCrunch*. https://techcrunch.com/2023/02/03/github-ceo-on-why-open-source-developers-should-be-exempt-from-the-eus-ai-act/

Schys, C. (2023). *The Lobbying Ghost in the Machine: Big Tech's of defanging of Europe's AI Act*. Corporate Europe Observatory. https://corporateeurope.org/sites/default/files/2023-03/The%20Lobbying%20Ghost%20in%20the%20Machine.pdf

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In H. Sheikh, C. Prins, & E. Schrijvers (Eds.), *Mission AI: The New System Technology* (pp. 15–41). Springer International Publishing. https://doi.org/10.1007/978-3-031-21448-6_2

Shen, H., DeVos, A., Eslami, M., & Holstein, K. (2021). Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–29. https://doi.org/10.1145/3479577

Shevlin, H., Vold, K., Crosby, M., & Halina, M. (2019). The limits of machine intelligence. *EMBO Reports*, *20*(10), e49177. https://doi.org/10.15252/embr.201949177

Strategic foresight. (n.d.). Retrieved June 8, 2023, from https://commission.europa.eu/strategy-and-policy/strategic-planning/strategic-foresight_en

Tambiama, M. (2020). *Digital sovereignty for Europe*. European Parliamentary Research Service.

The Legal Side of Open Source. (2023, March 27). Open Source Guides. https://opensource.guide/legal/

Theben, A., Gunderson, L., & López-Forés, L. (2021). *Challenges and limits of an open source approach to Artificial Intelligence*.

Think Tank European Parliament. (2021, May 19). *Understanding trilogue: Informal tripartite meetings to reach provisional agreement on legislative files | Think Tank | European Parliament*. Europarl.Europa.Eu. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)690614

Van Lente, H. (2012). Navigating foresight in a sea of expectations: Lessons from the sociology of expectations. *Technology Analysis & Strategic Management*, *24*(8), 769–782. https://doi.org/10.1080/09537325.2012.715478

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-019-14108-y

Virginia Braun & Victoria Clarke. (2021). *Thematic Analysis: Apractical guide*. SAGE Publication.

Wiggers, K. (2022, September 6). The EU's AI Act could have a chilling effect on open source efforts, experts warn. *TechCrunch*. https://techcrunch.com/2022/09/06/the-eus-ai-act-could-have-a-chilling-effect-on-open-source-efforts-experts-warn/

Yochai Benkler. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press.

Zhou, L., Gao, J., Li, D., & Shum, H.-Y. (2020). The Design and Implementation of Xiaoice, an Empathetic Social Chatbot. *Computational Linguistics*, *46*(1), 53–93. https://doi.org/10.1162/coli_a_00368