The narratives of the European Union's AI strategy: continued domination and structural bias?

R.T. (Renske) ten Veen

Master Thesis

MSc European Studies

University of Twente

21 June 2021

20673 words

First supervisor: Dr. M.R.R. Ossewaarde (University of Twente)

Second supervisor: Dr. C. Matera (University of Twente)

Abstract

Although AI has been praised for its benefits, gender studies and postcolonial studies scholars have criticised its development, stating that AI continues domination of certain entities over others and is prone to have biases. Nevertheless, within the field of public administration, AI has become a topic of notice in the past few years. This went along with the creation of new myths, idea-building and narratives in the policy domain. After several narrative studies have been done on the AI policy narratives of EU Member States begs the question whether the EU's narrative will make a difference for the direction of policies on AI and whether it will really cater to themes that do not include technological and economic advancement, but also more 'soft' concepts such as democracy, diversity, and gender. Employing a content analysis, this master thesis analyses to what extent concepts of domination and bias manifest itself in the EU policy documents on AI.

Keywords: European Union, Artificial Intelligence, gender studies, postcolonial studies, content analysis

Foreword and acknowledgements

When I started the master of European Studies at the University of Twente, I did not expect that I would graduate with a thesis on AI. In fact, I knew very little about AI prior to starting this project. However, one of my rolling goals is to always keep on learning new things, which is also why I decided to do this master. While trying to initially shape this thesis, I first spent several weeks reading academic papers on AI and its implication on society. The more I read into the topic, the more I got interested to look at it from a gender and postcolonial perspective. Dissecting the EU's AI strategy was a fascinating but concentration-heavy experience. It taught me a lot on how strategies are written and made as well as it showed the incredible amount of attention that AI receives within the EU. Now, when I read newspaper articles on AI or hear something about it on the news, I feel much more knowledgeable on the topic.

I would like to thank Dr. Ringo Ossewaarde for the many insights on AI and the constructive feedback during our meetings. Next to that, I would like to thank my second supervisor and second reader Dr. Claudio Matera for taking part in the project too and helping to shape the thesis with additional feedback. I would also like to thank my mother for her blind encouragement and good dinners. Finally I would like to thank my friends, colleagues at work and my boyfriend Deven for showing interest in my college work and for supporting me in different ways.

Renske ten Veen,
20 June 2021

Table of contents

Abstract **2**

Table of contents **4**

List of abbreviations **6**

1. Introduction **7**

2. Theoretical and conceptual framework **12**

2.1. Introduction 12

2.2. The European 'politics of AI' 13

2.3. Critical perspectives on AI 15

2.3.1. Postcolonial perspectives on AI 16

2.3.2. Gender perspectives on AI 17

2.4. Domination and bias 19

2.4.1. Concepts of domination 19

2.4.1.1. Eurocentrism 19

2.4.1.2. Androcentrism 20

2.4.2. Concepts of bias 20

2.4.2.1. Racial/ethnicity bias 21

2.4.2.2. Gender bias 21

2.5. Conclusion 22

3. Methods **23**

3.1. Introduction 23

3.2. Case and background 23

3.3. Method of data collection 24

3.4. Methods of data analysis 26

3.4.1. Content analysis 26

3.4.2. Praxis and coding scheme 28

3.4.2.1. Recap of concepts or themes 28

3.5. Conclusion 33

4. Analysis **35**

4.1. Introduction 35

4.2. The Eurocentric behaviour towards AI 35

4.3. AI and its problems with androcentrism 40

4.4. The (under)discussion of potentially racist AI 42

4.5. Women as AI's weakest link or an area of opportunity 46

4.6. Conclusion 49

5. Conclusion **51**

6. References **56**

Appendix A **59**

AI Watch 59

Studies, strategies and reports 59

Council of Europe 60

Recommendations 60

Declaration 60

Guidelines 60

Studies, strategies and reports 60

European Commission 61

Brochures 61

White paper 61

Recommendations 61

Guidelines 62

Studies, strategies and reports 62

OECD 63

Recommendations 63

List of abbreviations

AI	Artificial Intelligence
AI HLEG	Artificial Intelligence High-Level Expert Group
AWS	Automatic weapon systems
CEPEJ	European Commission for the Efficiency of Justice
CoE	Council of Europe
DG CONNECT	Directorate-General for Communications Networks, Content and Technology
EC	European Commision
EU	European Union
OECD	Organisation for Economic Collaboration and Development
STEM	Science, technology, engineering and mathematics

1. Introduction

The amount of scholarship on Artificial Intelligence (AI) is immense and its growing body of scholarly literature is hard to define in 2021. With the continuous development of technology, AI is increasingly used in our daily lives. Logically, as part of this development, or 'transition' as several scholars in the field have named it, AI has started to become a topic of notice in public administration and politics. In the 2010s, the European Union as well as its Member States started producing a large amount of policy documents, brochures and strategies on the development and deployment of AI. With this, a large share of new storytelling, narrative-building and myth creating was initiated on their behalf. In Europe, several EU Member States have released their own AI strategies, clearly denoting different narrative paths and interests when it comes to the development of AI. In existing scholarship, Ossewaarde & Gülenç (2020) make an analysis of these local 'myths' and find that local strategies are mostly focused on economic and technological benefits that come with the development of AI. However, as Fischer & Wenger (2020) note, the amount of scholarship that focuses on the content of AI narratives in Europe is still rather small. While Ossewaarde & Gülenç (2020) have found that most Member States' narratives focus on roles of leadership and utopianism, it begs the question whether the EU's narrative will make a difference with their own narrative on AI. It is curious to pinpoint whether the EU's narrative will reach further than just technological and economic advancement, and therefore also discuss AI's risks and disadvanagement as a result include more 'soft' concepts in its narratives, such as democracy, diversity, and gender.

After all, when it comes to AI, some do not celebrate its many advantages to society. We should therefore start to see AI as a social construction in which we humans have influence while constructing it, instead of a kind of deterministic technological force (Dignam, 2020). Sharma & Sarangi (2019) for example have noted that one of the most important issues in AI is its impact on human rights. Meanwhile, Monea (2020) underlines that society 'cannot have egalitarian or democratic technology if we hard code pre-existing regimes of marginalisation in our AI systems' (2020, p. 203). In addition, there has been criticism from feminists and postcolonial theorists on the technology, who argue that this development is harmful and biased negatively towards a certain part of the population, showing signs of domination issues (Dignam 2020; Monea 2020). Dignam (2020) notes how poorly designed AI, mainly driven by a 'small group of unaccountable technocrats', can lead to socially dangerous developments, creating the potential to discriminate. Continuing this from a feminist point of view, Ciston (2019) has argued that AI should be created an evaluated from multiple perspectives to abstain from creating these inequalities. After all, from research, it has been concluded that most people that work on AI are

predominantly white male engineers with a maths background and mindset, while women of colour are severely underrepresented (Dignam, 2020). Dignam (2020) logically points out that we run a risk to create bias and domination if we stay on the current AI development trial. Therefore, issues with democracy, diversity and gender should be assessed thoroughly while discussing AI development and deployment. In his chapter, Monea (2020) gives an example of the consequences of poorly designed AI policy, as he analyses the use of AI algorithms with regards to establishing image tags to visual data in for example Google Photos and ImageNet, showing that certain already-existed algorithms are biased and racist. Dignam (2020) has also pointed out several already existing features in which AI can be seen to contribute to potential harmful practices. Research for example proved that AI image recognition was calibrated with dominant images of white men (Buolamwini and Gebru, 2018, in Dignam, 2020). With regards to gendered AI, we see a similar problem. Research by Garcia (2016, p. 113, in Dignam, 2020) showed that men using a search engine to find jobs were shown higher paying jobs than women looking for the same role.

As shown above, a lot of research has been done on AI policy and the potential discrimination and diversity issues that comes with AI, including domination and bias. However, there is still a lack of research on this topic when it comes to the EU's policies. This has led to the following research question in this master thesis:

> To what extent concepts of domination and bias manifest itself in the existing EU policy documents on AI between 2011 and 2021?

As this research will mostly follow and add to a vastly moving body of literature, the research question will therefore be of a more descriptive nature and will not only pinpoint concepts, but also seek to analyse these concepts in its narratives. Focusing on the AI strategy and its subsequent EU documentation, this thesis logically employs a content analysis following the nature of the data, which is text. Following the critical literature on AI, this thesis will have two different focuses which each have a defined set of sub questions. Half of the sub questions of this thesis will address to what extent the EU's AI policies discuss the concept of domination. In the other half of the sub questions, this thesis will analyse to what extent the same policies address and include concepts of bias. As a result of that, this thesis will answer four different sub questions that focus on concepts with regards to domination and bias. First, this thesis will analyse how the domination concept of 'Eurocentrism' manifests itself in the EU policy narrative, which is followed by an analysis of the manifestation of domination concept of

‘androcentrism’.Subsequently, this thesis will show how racial or ethnicity bias is discussed in the narrative, followed by an analysis of gender bias. The exact sub questions are:

- SQ1: How does the domination concept of Eurocentrism manifest itself in the EU AI policy narrative?

To answer this sub question, the master thesis will employ a content analysis which is set to analyse the EU’s AI strategy from a postcolonial perspective. From this perspective, there will be a focus on the concept of ‘Eurocentrism’, which is one of the main critiques of postcolonial scholars on the AI agenda of policy-makers in Europe. What constitutes ‘Eurocentrism’ exactly is explained in Chapter 2 of this thesis, which focuses on the theoretical and conceptual framework. The exact way in which this sub question is further executed is explained in Chapter 3, which deals with the methods of this thesis.

- SQ2: How does the domination concept of androcentrism manifest itself in the EU AI policy narrative?

To answer the second sub question of this thesis, a content analysis will again be employed. Within this regard, there is a stronger focus on the domination question again, similar to the first sub question. However, this time, it is rather discussed through a gender lense. This thesis therefore opted for the term of ‘androcentrism’ to describe a phenomenon in the world in which the AI development, deployment and general policy-making is centred around one gender, namely males. The second sub question is explained and employed the same way as the first and will therefore be further explained in Chapter 2 and 3..

- SQ3: How does the concept of racial/ethnicity bias manifest itself in the EU AI policy narrative?

To answer the third sub question, this thesis again focuses on the postcolonial view of AI and aims to show how racial or ethnicity bias manifests itself in the EU narrative. Therefore, this sub question will again employ a content analysis to find an answer. The concept of bias is widely discussed in the literature and remains an important topic to be studied. The third sub question is also explained and employed through the same measures as the previous two.

- SQ4: How does the concept of gender bias manifest itself in the EU AI policy narrative?

To answer the fourth and last sub question, this thesis will employ a similar method to the previous, third question. The only difference is that instead of focusing on the postcolonial perspective, this thesis in this question rather focuses on the gender view of bias in AI. After all, the literature that is presented in Chapter 2 will point out that there is also a lot to discuss and say about the concept of gender bias in AI. Therefore, this thesis includes a sub question on the matter with regards to the EU's AI strategies. This question follows the same strategy of answering as the other questions presented in this thesis, using a content analysis as the main tool.

The contribution of this master thesis to the academic debate is threefold. First, this thesis will build on the studies such as Ossewaarde & (2020), Jimenez-Gomez et al. (2020) and Dignam (2020), which all dealt with the implementation and making of AI policies. As a result of this, the thesis is situated within the field of the 'politics of AI', seeing AI as a highly politicised development. This thesis will add an initial narrative study, showing the goals and shortcomings of the European Union's AI strategy. Second, this thesis will contribute to the current discussion in scholarship about diversity and AI policies, adding to the large field of current scholarship including Leavy (2018) and Monea (2020) with their individual analysis of the creation of AI has implications on people of different colours and gender identities. Currently, there are very few concrete studies through the lenses of postcolonial and gender studies with regards to AI. Although there has been enough theorised in the field about the implication of AI on policy, there is still a gap to fill with studies that show concrete examples of this. This master thesis aims to fill such a gap by employing a content analysis on the EU narrative with regards to their AI strategy. Third, this thesis aims to shed light on the European Union's process of policy-making and the specific function of the European Commission in the agenda-shaping phase. After all, this thesis studies the narrative of a developing policy. In its aim it will therefore show how the policies are (in)coherent in their presentation and in their content. With a slowly growing amount of literature on the EU's AI policies, this narrative perspective will add something else to the earlier research on supranational policy research done by Reis et al. (2020) and Hildebrandt (2020) on the implementation of AI policies in the EU.

This thesis analyses a qualitative data set within the framework of a qualitative interpretive content analysis. The research undertaken in this thesis will be of a purely qualitative nature and will only use existing qualitative data. As a result, a textual content analysis of the EU's policy documents was

thought to be the most appropriate method for this particular master thesis. The findings that come from content analysis can be useful 'building blocks' on the theoretical framework (Neuendorf, 2017). In content analysis, the goal is to develop 'generalisations' about the phenomena or topics that are analysed (ibid). The function of theory within the realm of content analysis is then to provide 'roadmaps' to create these generalisations (ibid). The reason why this method is indeed most appropriate for this thesis lies in the goal of this method, which is to 'undertake a close reading of the text to provide insight into its organisation and construction, and also to understand how texts work to organise and construct other phenomena' (Philips & Hardy, 2002).

This master thesis will start with a theoretical framework in Chapter 2, in which it will summarise and reflect on the most important findings in the literature with regards to AI. First, this chapter will include a state of the art on the 'politics of AI', showing the most important findings on the application of AI in public administration and in the social sciences. Then, this chapter will show an exploration of the use of critical perspectives and AI, focusing on an theoretical framework that will encompass postcolonial perspectives of AI, describing the potential racial bias that AI can potentially have followed by a theoretical framework that will discuss the gender research that has been done on AI. As a result of this, the chapter will continue to show how the concepts of domination and bias are important to the study of AI in public administration. After the theoretical framework, there will be an elaboration on the research design, which will include a content analysis of EU documents. Chapter 3 will therefore first start with a recap of the information with regards to AI in EU policy-making. This will be followed by the data collection method and the method of analysis of this thesis. Furthermore, this chapter will include a reporting section, in which the steps of the research are shown. Subsequently, this thesis will have an analysis section in which the two sub questions will be discussed. First, this master thesis will address the analysis of the extent to which concepts of domination are prevalent in the EU's AI policies. Then, the same will be done for the concepts of bias. This will then help to answer the main research question of this master thesis, which will be addressed in the conclusion section, and will sum up and reflect on the findings of the content analysis.

2. Theoretical and conceptual framework

2.1. Introduction

Artificial Intelligence (AI) is a computer science concept, but with its application to practical problems has found its ways to different fields, such as medicine, economy and the general social sciences. This specific thesis will solely focus on the social science and public administration application of AI. There has been an open question in the application of AI in public administration and the challenges it brings forward. How will AI affect democracy and society as a whole? And how can entities make robust regulation for AI that respects human rights? These questions are discussed within the dimension of the 'politics of AI' and form the framework of this chapter.

In the first part of this chapter, I will discuss the current praxis and academic consensus on the forming of AI-related policies. This first part aims to give an overview of the literature that has been written specifically on the application of AI in governance, including the term 'politics of AI'. Then, in the next subchapter of this paragraph, two relevant theories on the application of AI in public administration are discussed.

This thesis uses critical perspectives to address and discuss AI in a social science context. Specifically, this thesis uses postcolonial theory and gender theory as these two perspectives have been found especially relevant to current research on the application of AI to our daily lives. Finally, through the critical perspectives focus, this thesis conceptualises a few concepts with regards to AI within the field of social sciences. In order to specifically analyse the direction of the development of AI within the European Union, four different concepts are useful to look at when doing so. These four concepts are distilled from the postcolonial and gender perspectives that were dealt with in the previous subchapter on critical perspectives and AI. These four concepts were not randomly chosen. They all deal with the most important two issues with regards to the development of AI in society. The two most pressing problems are perhaps the risk of 'bias' and the development of a certain type of 'domination'. Both in the postcolonial and gender literature, these terms come back quite often. In this subchapter, two concepts of 'domination' are described as well as two concepts of 'bias', creating a conceptual framework for the analysis conducted in this master thesis. In this, AI will be conceptualised along postcolonial and gender perspectives and focus on the concepts 'ethnical bias' and 'Eurocentrism' within the postcolonial analysis and 'gender bias' and 'androcentrism'.

2.2. The European 'politics of AI'

There are several reasons why governments and governmental actors are so suddenly moving forward with the development of AI policies, creating the European version of the 'politics of AI'. First, because within the evolution of AI, there is an increasing understanding that AI is not only a technological development, but also interlinks with elements that adhere to governance, legal systems and societal problems (Jimenez-Gomez et al., 2020). The public administration oversight over AI development in both the private and public sectors, both at the national, international and supranational level has been established for some time already (Schippers, 2020). Kane (2019) notes that technology, including AI, should 'be monitored as any powerful person would be'. After all, as Kane (2019) states, these types of technology provide very suitable contexts in which powerful people can attack and undermine democratic principles (ibid). While AI is further developed, the protection and strengthening of democratic practices, processes, and institutions is essential (ibid). Schippers (2020) states that there is a 'real need' for international cooperation and collaboration with regards to AI regulation.

One of the main implications on governance from an AI perspective are ethical issues according to Jimenez-Gomez et al. (2020), including the protection of human rights and the use of artificial instead of human intelligence. Djeffal (2020, p. 256) writes that in order to understand the relationship between AI and democracy, AI should be understood as a broad technological concept first. That demands the question: 'Can systems solve complex problems independently?' (ibid). There is also a fear that AI takes over decision-making in several contexts, where a human being would have taken the decision before (ibid). Djeffal writes that 'the contingency of the internet means that 'like every medium before it, from the alphabet to television, [it] is shaped by the ways that society chooses to use its available tools'. (p. 260). Schippers (2020) underlines this view, as they write that AI has a wide-ranging impact on democratic politics and on society as a whole.

Some of the work is already on its way in Europe. The Council of Europe has installed a committee on AI for instance that has to examine AI in a human rights context (Schippers, 2020). In the EU, there have been efforts to install regulation too. The EU has established ethical guidelines on the use of AI, and, with its General Data Protection Regulation, has introduced a supranational framework for AI regulation in the area of data protection (ibid). However some scholars disagree with this view of AI receiving widespread attention from policy-makers, such as the EU. Rather, they argue that 'AI governance' and 'AI politics' are in their infancy. Fischer & Wenger (2021, p. 4), for example, state that the body of literature with regards to AI governance is still rather narrow and has focused only on several areas of application, especially within the security field. Haner & Garcia (2019) underline this interest of

the European Union to invest in AI from a military point of view, stating that the EU is among the top investors in AI-drived AWS. Rather than using AI in all parts of society with responsible frameworks and ethics guidelines, Haner & Garcia state that with combined knowledge of all its Member States, the EU could potentially become the dominant actor in the AI arms race too (2019, p. 335).

Meanwhile, in the eyes of Girasa (2020), the EU is 'cognizant' of the need for the development and growth of AI, but also the importance of regulating it in a manner that protects the technology from becoming abusive. With regards to these ethical issues, Reis et al. (2020) state that the EU has been 'a pioneer in defining good practices and creating new regulations for the use of AI', in their comparison of EU AI regulation to Portugal. The European Commision, the Member States, and in addition Norway and Switzerland, aim to cover a cross-border investment in AI, and increase trust among parties to the development of AI (Girasa, 2020). In addition to that, European Commission appointed a 52-member High Level Expert Group on AIto support any implementation and design of AI policy in Europe, and to, in the first place, give advice on the long-term challenges and opportunities with regards to AI (ibid). The High Level Expert Group has already published an AI Ethics Guidelines. These practices are also necessary according to AI scholars, but at the same time, they have highlighted that not in all cases, these ethical frameworks are among the top priorities of governmental actors. Buhmann & Fieseler (2021) question the implication of AI in governance, while AI is still suffering from poor transparency, explainability and accountability.

Outside of the scope of the EU, other actors in the policy-making field, for example the United Kingdom and the United States, have taken a more 'libetarian approach' towards AI policy (Dignam, 2020). This essentially means that they give the private sector more opportunity to develop. If we see governments as users of technology, it is not strange that they might choose to give more freedom to the development of AI. Jimenez-Gomez et al. (2020) states that governments might look to AI, finding it interesting as a way to further digitalise the public administration, using the best practices from the private sector in the public sector. Jimenez-Gomez et al. (2020) even underline that especially governments that are interested into further developing the use of data are right to consider AI as AI can help to further create a 'data-driven digital government', giving the example that a similar thing also happened with interoperability in governments. Djeffal (2020, p. 276) states that there is 'an active choice' for institutional actors, such as the EU, to use AI in certain contexts and at the same time, to inquire to talk about the 'ethics' or 'politics of AI' to discuss the relationship between AI and humans. Jimenez-Gomez et al. (2020) write that actors within the scope of the 'politics of AI' should themselves deliver public services to citizens with enough guarantees that these protect certain values and rights.

This literature review showed that there is a lot of agreement in the field about the use of and employment of AI in society and how public actors and institutions such as the EU should deal with the technology. At the same time, it is clear that there are no clear recommendations that are specific to the EU as an actor in the AI policy field. Values such as 'human rights' and 'democracy' are named, but understanding these values needs some further in depth discussion. In this thesis, there is a main focus on two types of value-based approaches to AI. The first one is the postcolonial perspective, which is going to be deepened first. Second, there is a gender perspective on AI, which will be deepened out after, which mainly focuses on the implications of gender on AI.

2.3. Critical perspectives on AI

This thesis offers a theoretical analysis from a critical theory perspective. Using critical theory is relevant to studying AI as critical theory focuses on the changing of society as a whole (Bohman, 2005). It is of no doubt that AI will heavily influence our lives and futures in the coming decades. Scholars have been discussing the implementation of AI in society and public administration quite widely, but have not seem to have come to a consensus whether the current development is harmful or harmless to democratic values for instance, leaving a lot of room for criticism on the current development of AI. Djeffal (2020) has written, for example, that democratic values should be included in the design process of AI. However, it is also noted that it is difficult to operate as there are different ideas about what is democratic and what is for example diverse (ibid, p. 270). However, Dignam keeps on stating that if the 'humans designing the project are not representative of society, have explicit and/or unconscious world views, this can strongly bias the outcomes'. According to Dignam (2020), the biggest challenge, unlike Djeffal (2020) puts it, lies not in identifying the exact risks and problems of AI, but to create an accurate public governance response to it. Kane (2019) underlines this view as well, stating that societal values need to be reasserted above technological and commercial imperatives, especially with regards to democracy. Sarangi & Sharma (2019) add that when technological progress becomes faster, as is the case with the recent development of AI, reacting should be more proactive in confronting certain ethical and moral ramifications, especially in situations that could be riskful to society. McQuillan (2020, p. 166) even goes as far as saying that AI should not be applied to any part of complex social and cultural problems, because otherwise certain AI elements, such as 'deep learning' cannot function when it comes to a certain degree of societal and political complexity. Meanwhile, postcolonial and gender scholars give a deeper interpretation to this last statement.

2.3.1. Postcolonial perspectives on AI

The field of postcolonial studies engages itself in both historical and contemporary inequalities, built on historical conditions (Bhabha, 1992, cited in Bhambra, 2009). In its roots, a postcolonial perspective of AI deals specifically with the relations between AI and the world that is still constructed by certain relations that were formed in the time of colonialism. With Europe being a former 'colonial power', it is important to address this specific perspective on AI with regards to the later analysis on European Union documents. With its monetary union, strong focus of the development of Europe in the current global power shift and aspirations for creating a hegemony, the European Union has been under critique from a postcolonialist perspective (Onar & Nicolaïdis, 2013). Postcolonial criticism bears witness not only to contemporary inequalities, but also to their historical conditions (Bhabha 1992, cited in Bhambra, 2009). Dirdlik (2002, cited in Bhambra) writes that the continued focus on developing Europe as a pseudo-centre of the world in terms that relate to the economy and development in general is a clear sign of Eurocentrism in the EU's policies. Meanwhile Onar & Nicolaïdis (2013) write that Europe is going through some kind of 'existential crisis' and as a result of that is long to re-establish its 'regional hegemony' in a time of a global powershift (p. 285). With that, in policies, you will find a more pragmatic and normative turn, which will aim at 'reinvigorating' the EU in a world that is increasingly 'less European' (ibid.).

A part of the AI scholars have however agreed that the postcolonial perspective, which includes questions about race and ethnicity too, is very important for the development of AI. As Monea (2020) states, questions of how AI intersects with pre-existing practices of racial marginational have become central as AI is continuously treated as the future of our world. We are expected to see the use of AI in our economies, in the military and in our governance systems in the near future. In their study, Ferrando (2014) specifically focused on attitudes towards the importance of including a 'race' or 'ethnicity' framework within the AI research field. Asking a set of Computer Science students across different stages of their academic career questions on the development of AI, Ferrando (2014) also asked them 'Do you think that concepts such as race and ethnicity will be significant in the development of AI?'. In this questionnaire, Ferrando (2014) notes that half of all the students responded with 'no' on the question. Meanwhile, about a third chose 'maybe' and a fifth agreed the most with 'yes'. At the same time, it becomes clear from the results that the further the students were in their stages of education, the amount of respondents answering 'maybe' increases too. However, at the same time, the margin of students answering outright 'yes' stays small.

When focusing on bias, Sharma & Sarangi (2019, p. 71) write that no machines can be expected to be free of bias as they are created by humans and humans are not free of bias as well (ibid). Campbell (2020, p. 121) writes that since machines were invented, humankind has tried to find the 'vox machina' and even as early as Turing's 'imitation game', scientists have tested whether and if machines can think on their own independently. Believing the idea that using AI to make decisions at the institutional level can lead to fairer decision making is therefore not valid and can potentially even lead to more institutional bias (Sharma & Sarangi, 2019). Monea (2020) writes that we cannot have an egalitarian or democractic type of technology with AI if we have hardcoded bias into our systems (p. 203).

2.3.2. Gender perspectives on AI

With regards to gender and AI, already back in the 1990s, the feminist debate on science produced outstanding approaches on AI, labelled under the encompassing term of Feminist Epistemology (Ferrando, 2014). The Standpoint Theory, which arose amongst theorists such as Dorothy Smith, Donna Haraway, Sandra Harding and Patricia Hill Collins, emphasised the starting point of this knowledge production, including more non-male knowledge production (ibid). In 1993, Adam (p. 313) writes that feminist standpoint theories offer the concept of a successor science to replace masculinist science but we need actually to see instances of a successor science to know if we can incorporate such ideas into AI systems. This connects to Ferrando (2014)'s view of that time in AI gender theory, and writes that technology and science, including AI, are never free from sexist biases. Ferrando (2014) then builds on Haraway's definition of the feminist approach to AI, which is that 'feminist objectivity means quite simply situated knowledges'. Ferrando (2014) takes the view that those who are in the centre of the hegemony, which can be a male-centred hegemony, do not have to reckon with the differing views that might be outside of that certain created hegemony. According to Adam (1993, p. 313), the AI and general science debate tends to focus on specific claims made by men only, an argument that can then degenerate into certain psychological analyses, for example the fragile male ego proving masculinity by mastering women; or functionalist arguments, or 'male reasoning', promoting the interests of men where such arguments are offered as causal explanations for the existence of particular ideas. While Ferrando (2014) takes a stance that the developed Feminist Epistemology sets the constitutive frame for the development of posthuman epistemological approaches of AI, feminist scholars from that period such as Adam (1993) and Halberstam (1991) disagree with this idea. Much rather, Feminist postmodernist views of AI reject the notion of a single truth about reality, for example the male-centric

truth (ibid). Adam (1993) argues that it is important to look at the 'situatedness' of each individual's viewpoint with a given cultural context to determine whether this viewpoint is biased in a sexist manner.

Then taking a posthuman or postmodernist stance of AI creation is not useful according to both Adam (1993) and Halberstam (1991). Halberstam (1991) writes that although in the framework of theories, postmodernism or posthumanism and feminism often seem to mirror each other, it is debatable whether two approaches are in dialogue or opposition and whether one takes precedence over the other. Generally, Halberstam (1991) contends the idea that feminism and postmodernism enjoy a mutual dependence within the scholarly and theoretical field of science and AI. Adam (1993, p. 314) argues that using postmodernism into bias perspectives on AI will lead to a certain kind of relativism towards the biases, which might lead to a disregard to the bias problems that exists, as postmodernists refuse to accept that anything is actually real and objective. However, postmodernism is not completely used as it aims to create pluriformity in views, which is also something that gender scholars want (ibid.). More modern scholars also agree with this view. Ciston (2019, p. 3) argues that AI should be created as well and critiqued and evaluated from multiple perspectives and methodologies in order to address the social inequalities that the technology can reinforce. Dignam (2020) states that gender bias has influenced AI systems that were mostly designed by white men. Dignam (2020) here names examples that include face recognition systems and word-embedding techniques. Lutz (2019) has named ‘poor women’ as the ‘test subjects’ of surveillance technologies. Meanwhile, Dignam (2020) also has described that a large company such as Amazon has employed types of gender discriminatory AI to recruit candidates for positions within their company. Cirillo et al. (2020, p. 2) on the other hand argue that the most common reason for these types of undesirable bias lies in the lack of creating a representative enough sample of the population while creating the AI programme.

In the Feminist Epistemology of the 1990s, there was an active question raised on the way forward (Adam, 1993, p. 321). Adam (1993) actively called for a way forward, stating that scientists needed to commit and face up to the issues that gender brought in AI. Cirillo et al. (2020) state that in recent years, the awareness of gender biases have increased and become more widespread at the same time. However, at the same time, the biases precede to still exist in AI, which has led Ciston (2019) to conclude that AI needs to be further re-imagined in the coming years and that this can be done by fostering communities that activate the necessary voices that are now left unheard.

2.4. Domination and bias

2.4.1. Concepts of domination

First, this thesis deals with the concepts of domination. After all, critical perspectives have shown that AI can be emancipatory and repressive towards human beings diagonally (Susen, 2009). Fuchs (2017) argues that the basic issue of domination itself is grounded in global capitalism, and therefore the development of technological inventions that can be used for economic gains, such as AI. This can lead to an albeit exploitative and exclusionary character which interacts with specific forms of domination including patriarchy, racism and nationalism (ibid). Susen (2009, p. 85) on the other hand calls 'domination' the opposite of emancipation, saying that human beings, rather than the economy, are the source of creating domination as we continuously construct so-called 'systemic imperatives'. This subchapter will deal with two particular concepts of domination, 'Eurocentrism' and 'androcentrism', which will be explained in the next two paragraphs.

2.4.1.1. Eurocentrism

When we are dealing with the concept of 'domination' within the scope of postcolonial analysis, we have to look at the concept of 'Eurocentrism'. Franzki (2012) has defined Eurocentrism as a concept that was defined by its critics, who say that Eurocentrism is prevalent in knowledge and power structures across the world. Franzki (2012) states that 'Eurocentrism' itself means 'a world-view (...) helping to produce and justify Europe's dominant position within the global capitalist world system'. Keita (2020) has argued that one of Europe's most important goals is to stay the economic centre of the world, creating 'Eurocentric economic dominance' over the world versus the Global South (p. 27). Ossewaarde & Gülenç (2020) have an analysis of local myths and regulations, finding that in Europe, local strategies are mostly focused on economic and technological benefits that come with the development of AI. These strategies are found to have found a clear path of Eurocentrism of AI. In Germany, they find that the political development of AI is heading into the direction of creating German leadership when it comes to EU progression of AI technology (ibid, p. 57). The Dutch approach, on the other hand, focuses on a so-called 'digital utopianism', in which AI will contribute to a growing economy and new corporate strategies coming with that (ibid, p. 59). With all these different approaches, it is therefore the question whether the European Union's approach will make a difference for the socialisation of policies on AI and whether it will really cater to themes that include democracy, diversity, and gender rather than just blatant Eurocentric economic growth as the local approaches did.

2.4.1.2. Androcentrism

Another important concept to consider within the framework of 'domination' is 'androcentrism' within AI development. Ferrando (2014) takes the view that those who are in the centre of the hegemony do not have to reckon with the differing views that might be outside of that certain created hegemony. According to Adam (1993, p. 313), the AI and general science debate tends is androcentric and the 1990s AI debate was focused on the view of men only, an argument that can then degenerate into certain psychological analyses, for example the fragile male ego proving masculinity by mastering women; or functionalist arguments, or 'male reasoning', promoting the interests of men where such arguments are offered as causal explanations for the existence of particular ideas. In 2021, it can still be argued that the concept of 'androcentrism' is very prevalent in the praxis of AI development. By some more recent feminists and gender scholars, such as J. Ann Tickner, (AI) security is described as mostly being steered and employed by men and therefore it being androcentric (Tickner & True, 2018, p. 221; Kappler & Lemay-Hébert, 2019). Gender scholars have also taken note of this. Hegarty (2007) has for example recalled that the discussion about gender, the role of females and transgenders, have been prevalent already since early AI experiments, such as the famous Turing's test. However, Hegarty (2007, p. 13) argues that the normativity that has developed around sexuality, gender and AI have become signs of the masculinist definition of intelligence (Hegarty, 2007, p. 13). The discussion of androcentrism in AI is therefore still alive and relevant to the research of certain narratives (Tomalin et al., 2021).

2.4.2. Concepts of bias

In order to speak of 'concepts of bias', we first need to define what is meant with 'bias'. In this, the view as proposed by Girasa (2020) is followed, which states that there are numerous biases found in AI-based technology. First, there is sample bias, then there is prejudice or stereotype bias (ibid). Girasa (2020) writes that sample bias occurs when the statistics do not accurately represent the true values of parameters as exemplified when the average of the set being studied inaccurately reflects the true average value of the studied target. According to Girasa (2020), prejudice or stereotype bias is the belief that certain attributes, characteristics, and behaviors replicate the typical qualities of a particular group of people. Sarangi & Sharma (2019, p. 84) have also written about prejudice and stereotype bias in AI and use the definition of a stereotype as a 'a set idea that people have about what someone or something is like, especially an idea that is wrong'. In the next two paragraphs, two concepts of bias will be discussed, namely 'racial/ethnicity bias' and 'gender bias'.

2.4.2.1. Racial/ethnicity bias

The first of the two concepts of bias is 'racial bias' or 'ethnicity bias'. Monea (2020) has written that the problem of bias rather lies at the heart of the development of AI, namely the computer scientists that build the programmes. Building on Megan Carcia (2017)'s argument, Monea (2020, p. 203) states that AI systems need to tackle these bias issues better if they keep being flagged up. On top of that, Silicon Valley, where most of the AI is made, needs more diverse computer programmers (ibid), so no 'racial bias' or 'ethnicity bias' is built into the systems. In this, Djeffal (2020), Monea (2020) and Sarangi & Sharma (2019) also underline that transparency is key as well in developing AI. Hernández-Orallo (2017, p. 398) therefore concludes that AI is an artifact that should be under constant surveillance and that there is a need to evaluate whether they do their tasks well in any kind of approach it takes, including social ones, making sure that there is no 'racial' or 'ethnical bias' in any of the systems.

2.4.2.2. Gender bias

The second of the two concepts of bias is 'gender bias'. Twine (2018, cited in Dignam, 2020) reckons that women, especially black women, are particularly underrepresented by a type of 'gender bias' in AI. Ferrando (2014) finds in their survey that gender is not quite seen as an important topic for computer science students, while they also note that the amount of students starting to find it a more important topic in AI grows while the level of education advances. Dignam (2020) noted that most people that work in tech companies are male. From the four largest tech companies, Amazon, Facebook, Apple and Microsoft, around thirty percent of its workforce was female (Reuters 2018, cited in Dignam 2020). When you zoom into the roles within these companies that deal specifically with the creation of new AI programmes and models, the amount of female employees is even less. Between the four companies, the percentages of women in these roles ranged from 23% to 19% (ibid). Dignam (2020) notes this as deeply problematic, as the AI systems that are created in this environment will reflect the flawed values of its tech designers, following Frischmann and Selinger (2018, cited in Dignam 2020)'s view that this is just as problematic for our perception of what is just, as in many cases we simply do not know the basis of AI decisions, which has become known as the 'black box' proposition. Cirillo et al. (2020) also reckon that the 'black box' can be the main problem at times, as it can introduce 'gender bias' by obscuring discriminatory practices. In sum, this shows that gender bias is an important concept to analyse.

2.5. Conclusion

This theoretical and conceptual framework focused on the application of AI in society. Scholars are more and more frequently writing about the use of AI in public administration, especially with regards to legislation and policy-making. The main question in the literature stays the extent to which legislators and policy-makers have to 'protect society' against AI and the extent to which they should invest in the future of AI development.

Next to that, AI has been discussed from a critical theory framework as early as the 1990s. Both postcolonial and gender scholars have shown the negative implications that AI can impose on parts of humanity due to existing structures in (mostly) Western societies. As a result of these conclusions, this chapter went further to conceptualise two different themes that came up quite frequently in the literature. These two issues or themes could be summarised as 'concepts of domination' and 'concepts of 'bias'. Within these two, four specific concepts were described, which included Eurocentrism, male domination, racial/ethnicity bias and gender bias. The four specific concepts will be used in the upcoming methodology chapter to analyse the current state of EU AI legislation and policy-making with regards to domination and bias.

3. Methods

3.1. Introduction

In this chapter, the main methods of this master thesis will be discussed. This chapter will start with a broad overview of the case and will demarcate the main case of this study. Following that, this thesis will describe the method of data collection. In this subsection, it will be explained how the data was gathered and which data will be included to analyse the case. In the last section of this chapter, there will be a focus on the methods of data analysis. First, the method, which is qualitative content analysis, will be explicitly described. After that, it will be elaborated on how this method is going to be used in the thesis.

3.2. Case and background

AI narratives are a relevant topic to study with regards to the field of EU studies, because the EU has written a lot about AI and has released a lot of documentation on the developing technology, supporting its decision with scientific studies, brochures and recommendations. In this, the European Commission has played a major role, which is not surprising as it is the main organ within the framework of the EU to prepare legislation. The policy development of AI shows that it is not only a relevant, but also very timely and important topic to the EU at the moment. After the European Commission announced that the EU would create an AI strategy, further steps were laid by the AI HLEG. Between then and the White Paper of 2020, the AI HLEG produced several different documents that were used by the EU as a background for writing its policies. Interestingly, while the development and academic study of AI has been around since the twentieth century, it has only been in the last decade that the EU has stepped up its mark to actually write something on the topic. It is important to understand the strategy that the EU has and its implication on amongst others more-often marginalised groups. In the current scholarship, there is a lot of focus on the benefits of AI in the EU, but a concrete focus on this development from a gendered and postcolonial perspective lacks.

Text is the main source for the EU to describe its AI strategy. The current available textual data on the EU's AI policy narratives is large and can be sorted in three different categories: national policies and strategies of EU Member States on AI, EU/European Commission policy documents on AI and documents on EU AI narrative written by other European institutions. In this thesis, it was opted to choose for the latter two categories, as national policies are relevant, but do not tell the whole tale about the EU's AI policies. 'Other' documentation would include documents of the Council of Europe and

the OECD on the development of AI policy in the European Union, because they do analyse the EU's policies rather well. Next to that, these two institutions actively give recommendations on AI policies. Therefore, documents from these two organisations were also included to see how far the EU follows up on the advice given by these organs.

The case that is analysed naturally also has some boundaries. The case includes all documentation on EU policy developments within the AI section and includes EU policies on AI specifically. This thesis understands policy development as the way how the current and future AI policies are being studied and how it is reported on the progress of AI policy in the EU. Therefore, this thesis does not quite study the application of AI policy or the local implementation of AI within the EU. Instead, it keeps the focus on the analysis of policy narratives as found in the documents. With regards to policy development, it can sometimes be hard to describe what would be included in this term. After all, policy development of the EU can also be studied from outside the EU as we see with the studies and reports from the Council of Europe and the OECD. The implementation of AI and political opinion on AI are not included in this thesis.

3.3. Method of data collection

In this section, I will explain the reasoning behind the gathering of suitable data for the study as well as making sense of this data. Then, the process of data gathering will be presented as well as the gathered suitable data for the analysis. Furthermore, it will be explained why this data is suitable for the analysis. By completing this content analysis as outlined above, this master thesis aims to give an analysis of common diversity issues in AI and into how far they are addressed in the EU's strategy.

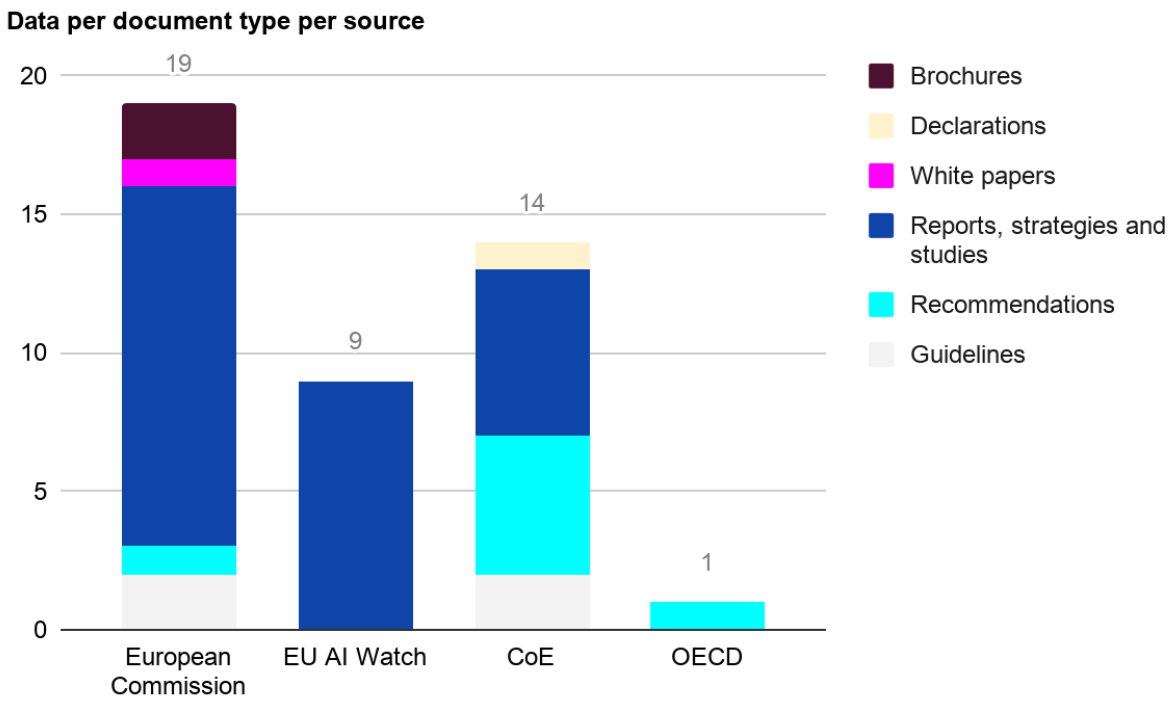
The data that has been gathered for the analysis of this master thesis is deemed the most appropriate type of data for this study and is believed to give an as complete as possible to be able to construct the EU's current development of AI narrative and regulation. Moreover, this master thesis uses 'within-case sampling'. This means that this research enables the researcher to thoroughly immerse into the available data on a single case study, which is in this case, the European Union (Mills et al., 2010). The units of analysis of this particular master thesis is the European Union's AI narrative.

Documents on the EU's AI policies were retrieved through the databases that are available on official web sites of the European Union's Commission, the Council of Europe and the OECD. The reasoning behind the inclusion of both the Council of Europe and OECD in the data collection lies in the fact that these two have done extensive research and provided recommendations on AI strategy developments within the EU, which makes the study of these documents relevant. The nature of the

documents that were available was quite broad, ranging from policy recommendations to specific research on strategy elements to actual AI strategies. The intention of this master thesis is to include as much as a diverse range of documents in order to get the broadest policy discourse possible. Within this, there was however a decision made to not include documents that announced or described events that had taken place within the sphere of AI development in the EU as these do not contain policy narratives.

After a large inventorisation, a total of 43 documents on the development of AI narrative in the EU were found to be suitable and available on AI within the EU sphere. The documents are retrievable in Appendix A of this master thesis. The documents were written between 2011 and 2021 and together add up to 1,869 pages. These documents are trying to conceptualise what should be included in the European Union's policy-making process. So far, the policy-making in the EU has reached up towards a White Paper, which shows that the AI strategy is still in its early design phase. Therefore, it is an interesting moment to make an analysis of the current state of AI strategy in the EU, although it is not very far developed yet.[1]

Figure 1



[1] Note: during the writing of this master thesis, in 2021, the EC has published several new documents on AI, including a proposal regulatory framework that addresses potential AI dangers, a coordinated plan on AI and a communication on the creation of a European approach towards AI.

As can be seen in Figure 1, most of the documents that are used come directly from the official websites of the European Union, 28 in total. On the European Commission's special website on AI[2], the section that deals with the published reports on AI within the European Union, called 'European Union AI policies', was primarily used[3]. There, in its backlog, the 22 documents were collected. Within this, 'call documents' and 'event reports' were neglected, as they were likely not going to show a policy discourse unlike the other documents. Included in the 22 documents are the four documents from the AI HLEG. All the available studies (9 documents) in the EU's special science hub AI Watch[4] were collected, albeit neglecting 'activities reports' for the same reason as 'call documents' and 'event reports'. These nine studies were seen to be in a different category as the EU, as they serve more as a basis for the policy-makers to deal with AI in Europe rather than a direct research requested by the Commission on the narrative. Of the other documents from the 43, 14 are from the Council of Europe. These fourteen were retrieved from the Council of Europe's special site on Artificial Intelligence[5] and then from the web page 'work in progress', showing all the available documents on AI in Europe. From this web page, I took all available documentation. The last remaining document is 1 from the OECD, taken from their website oecd.ai, which is a recommendation on the use of AI in OECD Member States, which is also relevant to EU AI narrative.

In Figure 1, all the documents are sorted per type of document and per source. This thesis has distinguished between six different types of documents: brochures; declarations; white papers; reports, strategies and studies; recommendations; and guidelines. On top of that, this thesis distinguishes between four different types of sources: the European Commission; EU's AI Watch; the Council of Europe; and the OECD. Most of the documents are reports, strategies or study results on the development of AI, 28 in total. This thesis further includes 7 on policy recommendations. 4 documents were (draft) guidelines on AI, 2 were brochures, 1 was a white paper and 1 was a declaration.

3.4. Methods of data analysis

3.4.1. Content analysis

As the previous paragraphs already showed, most of the strategy of the EU's AI manifests itself through text. As a result of that, a textual analysis, such as a content analysis, is a logical method of data

[2] https://digital-strategy.ec.europa.eu/en/related-content?topic=119
[3] Note: during the conduct of this study, this page was migrated to a new URL. Previously, this web page was available as https://ec.europa.eu/digital-single-market/en/artificial-intelligence, which now redirects to the web page linked in the footnote above
[4] https://knowledge4policy.ec.europa.eu/ai-watch/key-reports-analysis_en
[5] https://www.coe.int/en/web/artificial-intelligence/work-in-progress

analysis for this thesis. As Neuendorf (2017) has written that the findings that come from content analysis can be useful 'building blocks' on the theoretical framework. Therefore, this method of analysing the European Union's AI narrative is thought to be the most appropriate method for this particular master thesis. In content analysis, the goal is to develop 'generalisations' about the phenomena or topics that are analysed (Neuendorf, 2017). The function of theory within the realm of content analysis is then to provide 'roadmaps' to create these generalisations (ibid). Understanding validity as the accuracy of a method and the accuracy of understanding a method, a textual analysis is most appropriate for this thesis as the the goal of this method is to 'undertake a close reading of the text to provide insight into its organisation and construction, and also to understand how texts work to organise and construct other phenomena' (Philips & Hardy, 2002), which matches well with interpreting the concepts that are used in this thesis. With regards to analysing and measuring the employment of AI in public administration, several strategies and methods have been used. These have mostly incorporated a method related to strategy analysis. This thesis, however, will use a qualitative content analysis as a tool to analyse the data with regards to the concepts mentioned in the conceptual framework. Qualitative content analysis can be used in either an inductive or a deductive way. This master thesis uses a deductive way to conduct its analysis. Deductive content analysis processes involve three main phases: preparation, organisation, and reporting of results (Elo et al., 2014). The preparation phase consists of collecting suitable data for content analysis, making sense of the data, and selecting the unit of analysis (ibid.). In deductive content analysis, the organisation phase involves categorisation matrix development, whereby all the data are reviewed for content and coded for correspondence to or exemplification of the identified categories (Polit & Beck, cited in Elo et al., 2014). The categorisation matrix can be regarded as valid if the categories adequately represent the concepts, and from the viewpoint of validity (Schreier, 2012, cited in Elo et al., 2014). In the reporting phase, results are described by the content of the categories. (Elo et al., 2014).

The content analysis consists of two parts. In its first part, the analysis will focus on the concepts of domination, while in the latter part of this thesis, the analysis will zoom in on the concepts of bias. Both parts of the thesis will use the same data set, namely EU strategy documents on AI. This thesis will include forty-three different documents that are at the core of the EU's steps towards comprehensive AI legislation. In the next section, there is a short recap of the concepts that this content analysis will use, namely 'Eurocentrism', 'androcentrism', 'racial/ethnicity bias' and 'gender bias'. These concepts have been diluted from the theoretical framework. For example, when looking at domination from a postcolonial level, Eurocentrism is quickly the main concept that forms itself. Similarly, when looking at

domination from a gender perspective, androcentrism is the concept that takes the focus. The concepts that come from looking at bias flow a bit more logically. From a postcolonial perspective, this is obviously 'racial/ethnicity bias' and from a gender perspective, this is 'gender bias'.

3.4.2. Praxis and coding scheme

The way to describe the keywords in this master thesis is inspired by Weber et al. (2017) and Erlingsson & Brysiewicz (2017), who have opted for a table system that sorts the keywords per category and per concept per category. The key words naturally flow from the conceptual framework of the past chapter, that has described the most common keywords as well when it comes to describing a certain concept. The table in Figure 2 is designed as a way to sort the information in the documents.

3.4.2.1. Recap of concepts or themes

To discuss Eurocentrism in detail, this thesis has made a division between three different categories: economic development, technological development and 'Euroculture'. The codes that correspond with each theme are in italic and a dark grey.

- **Economic development**:
    - Europe aiming to benefit from AI *economic*ally (Keita, 2020);
    - letting the *economy grow* with the help of AI (Dirdlik, 2002, cited in Bhambra, 2009; Ossewaarde & Gülenc, 2020);
    - the economic potential and (dis)advantages of AI, increasing *capital* and *budgets* and stimulating *innovation* (ibid.).
- **Technological development**:
    - Europe as the centre of AI research; employing new *technologies*/AI/*cyber* in European society and economy (Bhambra, 2009);
    - Europe being the AI world leader and world leading producer in AI *development* (ibid; Ossewaarde & Gülenc, 2020).
- **Euroculture**:
    - Europe as the steering party or *leader* to develop 'ethical guidelines' to streamline AI in the *future of Europe* (Onar & Nicolaïdis, 2013);
    - European *norms* and *values* implemented in AI (ibid., Franzki, 2012);
    - European *identity* of AI (ibid.).

To discuss androcentrism in detail, this thesis has made a division between two different categories: defence and male reasoning. The codes that correspond with each theme are in italic and a dark grey.

- **Defence**:
    - AI as a new concept in the *defence* and *security* domain (Tickner & True, 2018; Kappler & Lemay-Hébert, 2019);
    - AI as an improvement for weaponry and the *army* in general (ibid.);
    - AI *war* (ibid.).
- **Male reasoning**:
    - the use of *male-centric* language, such as *'mankind'* instead of 'humankind' (Adam, 1993; Ferrando, 2014) and failing to acknowledge gender problems (Cirillo et al., 2020);
    - too much focus on the *functionality* of AI instead of sociological implications of current structures (Adam, 1993).

To discuss racial/ethnicity bias in detail, this thesis has made a division between two different categories: diversity and transparency. The codes that correspond with each theme are in italic and a dark grey.

- **Diversity**:
    - *diversity* at the core of AI development to combat bias (Dignam, 2020);
    - combatting racism in AI development (Monea, 2020);
    - discussing the current *power structures* leading to potential racist/ethnicity bias in AI (ibid.).
- **Transparency**:
    - the discussion of the creation of racist AI *algorithms* and the potential racist bias problem of the *black box* (Monea, 2020; Dignam, 2020);
    - the use of AI in *surveillance* systems and the risks of racist bias (ibid.);
    - the discussion of the existence and lack of *transparency* in AI (Djeffal, 2020; Monea, 2020; Sarangi & Sharma, 2019).

To discuss gender bias in detail, this thesis has made a division between two different categories: gender and equality. The codes that correspond with each theme are in italic and a dark grey.

- **Gender**:
    - the discussion of the concept of *gender* with relation to the development and deployment of AI (Ciston, 2019; Dignam, 2020);

- the inclusion of *women* and *other genders* in the narrative (Cirillo et al., 2020);
- the discussion of the concept of bias with regards to *genders* that were mentioned in the narrative (Dignam, 2020).

- **Equality**:
  - the catering of *equality* to combat bias in AI (Ciston, 2019; Dignam, 2020);
  - the existing structures leading to bias in AI such as *gaps* between people (Dignam, 2020);
  - the current and future situation on the *labour market* that will either contribute or help to decrease gender bias (ibid.; Ferrando, 2020);
  - the use of *quotas* to ensure the decrease of gender bias in AI (Cirillo et al., 2020).

Figure 2

Theme	Category	Codes
Eurocentrism	Economic development	Budget
		Capital
		Economic
		Economy
		Growth
		Innovation
	Technological development	Cyber
		Development
		Technological
	Euroculture	Future of Europe
		Identity
		Leadership
		Norms
		Values
Androcentrism	Defence	Army
		Defence
		Security
		War
	Male reasoning	Functionality
		Mankind
		Males
		Men
Racial/ethnicity bias	Diversity	Diversity
		Power
		Racism
		Structures
	Transparency	Algorithms
		Black box
		Surveillance
		Transparency
Gender bias	Gender	Females
		Gender
		Non-binary people
		Transgenders
		Women
	Equality	Equality
		Gap
		Labour market
		Quota

This thesis takes inspiration from the content analyses done by Kassarijan (1977), and Taylor (2003) on gender stereotypes and Weber et al. (2017) on a public administration topic. Both content analysis focus on the examination of certain key words per category or theme in the data. Normally, this can be done in both a quantitative or qualitative way. In this thesis, a qualitative content analysis is central. This means that the use of these keywords will not be quantified but rather qualitatively analysed. First, there will be a search for these keywords or 'codes' in each text. This thesis will use sheets or tables for the notation of these quotes before continuing further research. An example can be found below in Figure 3:

Figure 3

Policy document	**Narrative**	**Code(s)**
European Commission. (2018, 24 April). *Communication Artificial Intelligence for Europe*. European Commission. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51625	*"To further strengthen trust, people also need to understand how the technology works, hence the importance of research into the explainability of AI systems. Indeed, in order to increase transparency and minimise the risk of bias or error, AI systems should be developed in a manner which allows humans to understand (the basis of) their actions."* (p. 15)	transparency

The quotes are further analysed and interpreted to see how the word is essentially used in the context of the themes that are at state in this master thesis. As is clear from this, this content analysis is of an interpretive nature. With regards to qualitative social research, Rosentahl (2018, p. 18) reckons that interpretive methods can achieve the investigation of phenomena that have been little studied. This fits to the master thesis as there has been little research done before on strategy documents on AI from both a postcolonial and gender perspective, including concepts of domination and bias. Rosentahl (2018) writes that interpretive social research seeks to understand subjective meaning and reconstruct latent meaning, and the implicit knowledge of the actors in their social world. In practice, in this master thesis, this means that the keywords will be used in a way to find out how the documents means deal with more abstract topics that are relevant to such as 'eurocentrism', 'androcentrism', 'racial/ethnicity bias' and 'gender bias'. Coming back to the main research question of this thesis, it will essentially answer a 'simple' question as 'do these documents address them and if so, how?'.

However, before we continue with the main body of this thesis, the analysis itself, it is important to first address reliability. Understanding reliability as the consistency of a method to measure a certain

concept, this thesis again underlines that it is of an interpretive nature. As we know, interpretations can vary among different people. However, this does not mean that this thesis is impossible to replicate. The term 'subjective meaning' for instance does not refer to a completely private mental process of the individual, as Rosentahl (2018) writes that the subjective meaning is based on how the actors involved ascribe meanings to their actions, and to their social reality. Interpretive social research is partially on 'the collective stocks of knowledge acquired in the course of their socialisation' that are already established prior to the analysis (Rosentahl, 2018, p. 18). With the method and coding system that is presented in this subchapter, this master thesis aims to state that the same result would have been achieved if another researcher had used the same method underneath the same circumstances. Therefore, this thesis will code closely alongside its concepts and will retrieve its conceptual framework from the theoretical framework. By doing this, this thesis aims to further solidify the analysis of the four concepts that are at stake.

3.5. Conclusion

To sum up this chapter, it can be concluded that the development of EU AI strategy can best be studied from a policy narrative perspective. This policy narrative can be best found by analysing reports, studies and other documentation on EU AI strategy. These documents do not necessarily have to come from the EU directly. Also documents from the OECD and the Council of Europe can be deemed useful in the purpose of analysing AI policy development in the EU. As set out in the method of data analysis section, this thesis will use 43 different documents to analyse the concepts at stake. The analysis will be employed via a content analysis as explained in the method of data analysis section.

In its central research question, this thesis aims to show to what extent certain concepts can be found in EU AI policy development. Within these policy documents that will be of interest, especially the study of their narratives is central to this thesis. Their narratives can tell a deeper meaning of the text, going to the core of the intention of the documents and the EU AI strategy as a whole. Within a narrative, concepts are also likely to be displayed too. This thesis aims to answer four sub questions, which on the one hand deal with the question of how the EU's AI strategy deals with concepts of bias and on the other hand, concepts of domination.

This thesis will continue as follows. In the first part of Chapter 4, there will first be a focus on how the documents address concepts of domination. In this first half, which consists of two chapters, two sub questions will be answered. These sub questions are: 'How does the domination concept of Eurocentrism manifest itself in the EU AI policy narrative?' and 'How does the domination concept of

androcentrism manifest itself in the EU AI policy narrative?'. The sub questions will be answered by the employment of the content analysis presented in this chapter. The content analysis will look for the most telling quotes and will try to give a sense of what the texts that are analysed are aiming to say, looking for agreements as well as contradictions between different documents. By looking from a domination perspective, this means that there will be an explicit focus on the appearance of certain codes, which include for example 'development', 'future of Europe' and 'men'. Meanwhile, in the second half of Chapter 4, there will be a focus on bias, which will answer the sub questions 'How does the concept of racial/ethnicity bias manifest itself in the EU AI policy narrative?' and 'How does the concept of gender bias manifest itself in the EU AI policy narrative?'. Using the same method as in the first half of Chapter 4, this part will have an explicit focus on the appearance of codes such as 'black box', 'equality' and 'gender'.

4. Analysis

4.1. Introduction

This chapter analyses the EU narrative on the development and deployment of AI in our economy, healthcare and society as a whole. The first part will focus on five different themes: economic development, technological development, and norms and values (the so-called 'Euroculture' as explained in Chapter 3) from the wider perspective of Eurocentrism and defence and male reasoning from the perspective of androcentrism. The second part of this chapter focuses on the discussion of bias in the EU's AI narrative. This part will focus amongst others on themes such as transparency, racism, feminism and gender.

In this chapter, the narrative that was found with regards to the codes will be discussed in detail, trying to make sense of what was important and dominant in the narrative on AI and what categories were not.

4.2. The Eurocentric behaviour towards AI

As Franzki (2012) and Keita (2020) explained it in Chapter 2 of this thesis, one of the directories of Eurocentrism in practice is the continued focus of Europe's position in the global society and especially the further development of this in terms of 'economic growth'. In the EU's AI strategies, growth is an important theme. A large proportion of the EU narrative with regards to AI focuses on the potential economic gains that AI will bring to Europe and the EU economy. The importance of AI in EU society and, as a result, the economy is clearly named by the Coordination of Digital Strategy and Green ICT unit (2020), which state that AI has become a 'mature technology' and an increasingly important part of our daily lives. Meanwhile, Misuraca & Van Noordt (2020) write that AI has a large potential for the economy of the EU and conclude that one of the main aims of the European AI strategy is to increase the AI uptake across the economy. Misucara & Van Noordt (2020) write amongst others that AI has become 'an area of strategic importance' with the possibility to become a 'key drive of economic development' (p. 2). This statement was also underlined in the communication by European Commission (2018) and repeated by Samoili et al. (2020a, p. 7), who wrote that from its market perspective, AI has a 'strong focus on industrial development' and the 'assessment of the economic value of future market predictions'. It is therefore not surprising to see the term 'economic benefit' to be discussed in the whole EU narrative on AI as the motor for economic growth as well. Citing Lee (2018), Samoili et al. (2020a, p. 5) argue that the main benefits of AI actually will come from its wider application across the economy

much rather than from the continued development of the technologies around it. Samoili et al. (2020a) argue that this is also very much the view that the European Union pertains too when it comes to AI:

> *"The Artificial Intelligence for Europe Communication of the EC sees Artificial Intelligence (AI) as a technology that could boost the European competitiveness (EC, 2018a cited in Samoili et al., 2020a). The Coordinated Plan on Artificial Intelligence outlines how EU Member States could coordinate their strategies, efforts and investments to maximise the benefits of AI for Europe (EC, 2018b, cited in Samoli et al., 2020a)"* (Samoili et al, 2020a, p. 4).

The narrative concludes that an integral part of developing AI for it to aid the European economy depends on investments. Samoili et al. (2020a, p. 5) writes that a successful transition to an 'AI-driven economy' relies heavily on the investments of the public sector in AI. This line of thinking is especially found in narratives that have a particularly strong focus on economic development, as they tend to also have a strong focus on investment and financial support of the development of AI. Tsinaraki & Kostic (2020), for example, focus on the historical development of AI, also in the economic sense, and therefore showed a strong focus on the amount of investment and budget that has been spent on AI throughout the years. De Nigris et al. (2020) meanwhile conclude that while Europe and the EU are strong on the basis of research on AI, but that the investments supporting the development need to be scaled up (p. 5). Nepelsi & Sobolewski (2020) write about 'increasing capital' as rather increasing 'organisation capital' and 'human capital' as opposed to 'capital' just being about the financial aspect. Economic development was also connected with words like 'opportunity' and 'benefit', also 'progress'. De Nigris et al. (2020) take a term as 'innovation sandbox' as a way to operate, invest in and use AI to practically experiment. Interestingly, European Commission (2021) also strongly focuses on the 'economic benefits' of AI and the use of 'innovation', saying that the EU should strive towards 'a new legislation that mitigates risks but does not limit innovation'. Not every part of the EU narrative however focuses on 'innovation' and 'economic development' as interlinkages. For example, the European Commission (2020) discusses 'economic' and 'growth' extensively, while 'innovation' is not mentioned at all.

Meanwhile however, in the parts of the narrative that deal with economic growth, there is very little attention to the social and normative consequences to AI in comparison to the economic gains. In much of the narrative, there was nothing to be found about 'norms' and 'values', or ethical issues with regards to economic growth or economic development. This shows that although there has been a narrative in the scholarship that the EU is trying to lead with regard to the ethical side of AI as well, that the development of this ethical framework is not yet quite on the front of the AI policy narrative. However, at the level of the Council of Europe, an organisation which mostly focuses on human rights,

the creation of an ethical framework is seen as vital. For example, the *Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes* of the Council of Europe mentions 'economic' issues rather in a way that economy leads to inequality as it mentions terms as 'economic segregation'. 'Innovation' is mentioned in the way that it needs to be structurally diverse so it cannot do harm to people. Writing that the EU completely excludes discussions about norms and values in its narrative would be unfair. However, DG CONNECT (2019a) shows that the AI narrative is not solely centred around 'economic growth' as a benefit to society but it also includes more social problems such as 'economic exclusion' too occasionally. A similar narrative can be found in Servoz (2019), which speaks about the opportunities of AI in the working world. Besides discussing the implications that AI will have, both positively and negatively, on the labour market, the potential risks that AI brings to the labour force are also touched upon. Servoz (2019) notes:

> *"Advanced automation technologies can bring about many opportunities but also risks. These risks and opportunities are not equally distributed across society or the economy. Gender, ethnic and socioeconomic diversity are important because careers in AI are well remunerated and are an area of rapid growth."* (Servoz, 2019, p. 6).

Aiding to that general economic development of the EU is the technological development of AI too. This becomes clear through the narrative as there is a clear interlinkage between technological development and economic development. For example, Nepelsi & Sobolewski (2020) state that specifically, 'technological development' is often coined alongside notions such as 'economic progress' and 'economic growth'. At some times, they use a grouped notion of 'technologically-driven economic growth' or 'technologically-driven economic development'. Using the example of Silicon Valley, Nepelski & Sobolewski (2020, p. 4) argue that the history of AI shows that when the public sector, such as the EU, invests in technological development, it will benefit the economic sector. This strong interlinkage between 'technological development' and 'economic development' can also be found in Martinez Plumed et al. (2020), which assesses the readiness of certain sectors with regards to implying AI within the EU. Among other things, Martinez Plumed et al. (2020) argue that the massive impact of AI can only be reached when the technology really ends up being that transformative as previous technologies such as electricity also have been and there needs to be an extra investment in the development of expertise in AI. In much of the narrative, it becomes clear that there is rather a strong focus on using AI for the benefit of technological advancement and research. However, the Council of Europe (2019) has also argued that consideration should be given to social trends and that AI also raises novel questions about autonomy, privacy and transparency. This latter view is not always so clearly present in the narrative that

directly comes from the EU. One of the places where this is not so clearly visible is in AI HLEG (2020)'s narrative, where it states that the international competition is strong and that Europe's leading position in AI can only be guaranteed in the case that forces are joined at the European level in order to exploit the potential of AI (p. 3). It is clear that the EU is also very invested in becoming a world leader in AI development. In the European Commission (2020)'s *White Paper on Artificial Intelligence*, there is a focus on the creation of 'technological leadership' in a race for 'global leadership'. Writing that the aim for the EU is to become a 'global data hub', European Commission (2020) argues that the EU should therefore use its 'technological strengths' and 'technological capabilities'. However, unlike AI HLEG (2020), with regards to 'development', European Commission (2020) argues that the ethical side of AI should not be forgotten either and that the EU can also become a global leader in that respect.

> *"Europe is well positioned to exercise global leadership in building alliances around shared values and promoting the ethical use of AI. The EU's work on AI has already influenced international discussions. When developing its ethical guidelines, the High-Level Expert Group involved a number of non-EU organisations and several governmental observers. In parallel, the EU was closely involved in developing the OECD's ethical principles for AI25."* (European Commission, 2020, p. 8)

Adding to that, DG CONNECT (2019a) also states that Europe must aim for technological leadership, but at the same time, this responsibility and research should focus on all aspects of AI, including social and ethical ones. It says that 'trustworthy AI solutions that address societal challenges should be promoted' (DG CONNECT, 2019a, p. 21). The Council of Europe meanwhile states that democracy should come first with regards to developing the technology rather than aiming for technological leadership as the main goal of the EU (*Declaration by the Committee of Ministers*, 2019). It states that particular attention should be brought to those who develop AI to monitor them to see whether they apply with the ideas of democracy and fairness (ibid, p. 2). Already at an earlier moment, the Council of Europe (2018) stated that there is a lot of emphasis on the role of the state as a 'rule maker' to protect citizens in any case. With regards to the combination of 'technological development' and 'social development', it is visible that there is often discussion at the Council of Europe level about the need to not forget the 'human component' of AI development, as for example is said by the Council of Europe Commissioner for Human Rights (2019). At the EU level, DG CONNECT (2021) shows however that a strong focus on fostering 'development' is pertained, although DG CONNECT also writes that this development should be 'human centric', stating that machines should have similar ideas as humans.
With this view in mind, it is not surprising that part of the narrative focuses as much on soft implications as on hard benefits of AI, including the ethical side and values. Misuraca & Van Noordt (2020) state that

the overall goal for the EU is to become a world-leading region in developing and deploying 'cutting edge, ethical and secure AI, promoting a human centric approach at the global level' (Misuraca & Van Noordt, 2020, p. 7). This is in line with the Council of Europe's *Recommendation CM/Rec(2010)13*, which dealt with the protection of the individual in the development of data and AI, there is more of a focus on creating a more wholly social view on AI too. Also in *Recommendation 2102 (2017)1* from the Council of Europe, which dealt with technological convergence, AI and human rights, there was no mention of economic development and rather, the contents of this text focused on how AI will affect society. In the 2019 OECD Recommendation on AI, a clear balance is also struck between economic benefits and development as a result of deploying AI. This recommendation denotes the implications that AI might have for democracy and human rights, two very important topics in Europe. For example, it states:

> *"RECOGNISING that, at the same time, these transformations may have disparate effects within, and between societies and economies, notably regarding economic shifts, competition, transitions in the labour market, inequalities, and implications for democracy and human rights, privacy and data protection, and digital security;"* (OECD, 2019, p. 6).

Analysing the ethical component of AI in these documents is thought-provoking, because part of the narrative is very focused on the ethical side of AI, while in other parts of it, this component completely lacks. This latter situation is for example the case in Samoili et al. (2020b) as they focus on the definition of AI and analyses a set of 55 documents from different sources. In its introduction, theory section and methodology, it aims to create an idea of what AI is, while the ethical side is barely addressed at all. Values are among the most mentioned soft tools with regards to AI development from a more social-cultural perspective. While, the AI HLEG (2018) recurringly says that 'trustworthy AI' is the 'Northern star' of the EU, the Council of Europe (2016b) repeats that policy-makers should create an inclusive space for ethical discussion of AI and how the development of the future situation is in the hands of the public administration. Here, things such as 'principles', 'norms' and 'values' also come into play. Also with regards to the use of AI within the legal framework, the EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ) (2018) discusses a similar idea. 'Economic development' is mentioned there as a contextual factor to be taken into account by a judge when making a decision. The Council of Europe (2019) simply only coins 'European values', while in the European Commission (2018)'s Communication, it only talks about 'European Union' values alongside 'fundamental rights'. On page 20 of the communication, it says that 'the Union's values and fundamental rights are at the forefront of the AI landscape'. In Van Noordt et al. (2020), 'EU values' and 'fairness' were mentioned to be important

types of ‘values’. In Misuraca & Van Noordt (2020),‘European values’ was coined, specifically stating that this concerned ‘democratic values’, ‘human values’ and ‘human rights’. Also in the Council of Europe’s Recommendation CM/Rec(2019)10, ‘democratic values’ were seen as the most important values to be considered with regards to the creation of education on technological developments such as AI. The Council of Europe’s Recommendation CM/Rec(2020)1 gave a more superficial narrative as it stated that:

> *“[The] impact [of algorithmic systems] will also depend on the broader organisational, thematic, societal and legal context in which they are used, each of which is associated with specific public and ethical values ”.* (Recommendation CM/Rec(2020)1, p. 5).

At the EU level, however, it is not always clear what kind of values are at the stake of the development and which values are exactly more important than other values. DG CONNECT (2021) states that machines should have the same set of values as humans, not specifically mentioning which values. When speaking about ‘values’, meanwhile, the EU’s Robotics and Artificial Intelligence Innovation and Excellence (Unit A.1) (2021) talks about ‘European values’ and ‘European Union values’, although not clearly stating what these values entail. Digital Economy Unit et al. (2021) name a wider context of ‘values’, including ‘democratic values’, ‘citizen values’ as well as ‘European values’. Interestingly, here the values are narrowly specified, something that clearly lacked in most of the sections that dealt with values. Here, ‘European values’, also includes ‘privacy, transparency, accountability, fairness, and trust’ (Digital Economy Unit et al., 2021 p. 2). These narratives show that the EU is still in the need of defining what exact values are at stake and at the heart of discussion with regards to AI. In its current position, the ‘values’ narrative shows clear incoherences between the different actors within the EU and their policies.

4.3. AI and its problems with androcentrism

As shown in the theoretical framework, the Feminist Epistomology of the 1990s had as one of the strongest arguments that the development of AI was mostly organised in a male-centric way and with a lot of malecentric language. With regards to the EU, this malecentric language use, such as using terms as ‘men’ and ‘mankind’ was almost wholly absent in the policy narrative. Following the theoretical framework, it is hard to declare why this is the case as these results were rather unexpected. Reasons for the results can be found in the fact that the world and also the world of science and policy-making has moved on significantly since the 1990s. Additionally, a lot of documents that were analysed for this master thesis were at least co-written by people of other genders than men.

Does this mean that the EU narrative is not malecentric at all? As said, none of the texts have proven to be particularly androcentric. None do have the classic 'male reasoning' components in its textual body as stated in the Feminist Epistomology of the 1990s, including terms as 'mankind' and focusing extensively on 'security' and 'functionality' issues. However, on the flipside, most texts are actually not quite fighting the status quo when it comes to potential androcentrism or when it comes to actively opposing androcentrism. This contrasts the Council of Europe's Recommendation CM/Rec(2019)1 which deals with the implications of sexism in society and technology, including AI. It is interesting that little of the EU narrative so far has addressed these issues of a sexist and female-unfriendly type of economy and technology that AI can create. It is thought-provoking how there is a discussion about terms such as 'economic violence against women' and 'ending cyberdiscrimination' in the Council of Europe's recommendation, while the EU narrative has not included these terms at all.

A very small fraction of the narrative with regards to security contains fragments of a form of more 'militarised speech', using terms as 'army' and 'war', you often see that there are multiple examples in one document or there are no examples at all. In general, these terms' were only rarely used and mostly as a metaphor. For example, Craglia et al. (2020) coin terms as 'security' and 'war' (e.g. 'cyberwar') in a military sense.

> *"In recent years, [AI-powered psychographics have] become a new tool and target for social influence and control, from tailored advertising and nudging consumer's habits to manipulating political orientation, and it is linked to new modalities of 'digital aggressions' and even (cyber)war".* (CB Insights, 2018, cited in Craglia et al., 2020, p. 14).

Meanwhile, at the EU side of things, DG CONNECT (2018) focuses on 'European leadership' as well and there is talk of Europe being in a global 'war' with regards to AI talent. Similarly, the European Commission's Microelectronics and Photonics Industry (Unit A.3) (2020) talks about the opportunities ahead with regards to new technologies and compares the budget of the EU with the budgets of the US and China and concludes that the EU needs to step up its game, talking about an 'economic war' with regards to AI.

Most of the AI narrative that has to deal with defence, has to deal with 'security', usually connoting 'protection' and 'safety' of people against machines, rather than machines helping to create that feeling of security themselves. For example, in Misuraca & Van Noordt (2020), 'security' in itself is meant more in a softer feeling of 'safety' than 'safety' in a militarised or hard sense. Other narratives, such as Samoili et al. (2020a), Council of Europe (2016b), the Council of Europe's Recommendation CM/Rec(2020)1, and DG CONNECT (2021) and European Commission (2020) agree with this view, coining

‘security’ mostly with terms as ‘protection’ and ‘safety’. Nevertheless, some of the narrative rather talks about using AI as a measure to actually create security itself. Samoili et al. (2020b) use ‘security’ to describe an example of a potential domain in which the EU can use AI technology to its benefit. Van Noordt et al. (2020) coin ‘security’ alongside the term ‘opportunity’ for AI to be used at. The Robotics and Artificial Intelligence Innovation and Excellence (Unit A.1) (2021) mentions ‘security’ in the context of applying AI to the security sector. Martinez Plumed et al. (2020) specifically say that AI could be used for law enforcement by applying facial recognition technology. H

However, on the contrary, part of the narrative about ‘security’ is also about the fact that people should rather be protected from AI itself rather than be protected by it. For example, in the Council of Europe's Recommendation CM/Rec(2010)13, security was also mentioned as an important topic, but it is discussed within two different lights. First, it is discussed in the context of ‘state security’ and second, in the context of the ‘protection of [personal] data’. Meanwhile, the Directorate General of Human Rights and Rule of Law (2021) states that when ‘security’ is part of the development of AI, the focus should be on the human rights aspect and on potential dangers. In other words,: people should not become a victim of AI systems in a ‘security’ situation. Adding to that in the Council of Europe Commissioner for Human Rights (2019) states that ‘security’ of AI should refer to protecting the people against potential harmful biases that AI can have when inappropriately used. On top of that in the OECD Recommendation, it is said that transformations that come with AI ‘may have disparate effects within’, and then obviously between ‘societies and economies, notably regarding economic shifts, competition, transitions in the labour market’ and and as result of that, employing AI might to inequalities and implications for ‘digital security’ (OECD, 2019, p. 6). The focus on ‘security’ is the only part of the EU’s AI narrative that can be named somewhat androcentric. This is not strange, given that the security domain is one of the most androcentric sectors in Europe and the EU, mostly employing males, which makes it interesting to analyse in the context of feminist studies and AI.

4.4. The (under)discussion of potentially racist AI

Among the policy narratives of the EU, a low level of attention is dedicated to the potential harmful effects AI can have if it is biased in a certain way. While gender bias is addressed in several policy narratives, it became clear that especially bias influenced by racial or ethnical motives is barely discussed in the policy narrative. It is surprising to see that in some narratives, AI is rather seen as a way to counter potential biases, which is of course, looking at the current scholarship by Dignam (2020), Monea (2020) and Sarangi & Sharma (2019), not quite the case. In this line of thinking, AI is, for example, discussed as a

possible replacement in humans in healthcare processes, by more strongly automated processes (DG CONNECT, 2020). Matters such as visual inspections can be replaced by sensors and algorithms as well as using AI for the processing of communications (ibid., p. 35). DG CONNECT (2020) concludes that AI can reduce costs, optimise processes and can reduce waiting times in healthcare. In the same publication, DG CONNECT also argues that algorithms can systematically do certain things better than humans can (2020, p. 58).

Strangely enough, two pages past that particular statement that AI is better capable enough in some respects than humans, DG CONNECT also argues that algorithms can be led to make mistakes when supplied with incorrect information and that these errors can have considerable consequences (2020, p. 60). AI HLEG (2018) also acknowledges that discrimination with AI can occur as those in control of the algorithms may intentially try to create biased outcomes in order to exclude certain groups of people (p. 16). This can be done by manipulating the data and employing algorithms that are biased from its core (ibid.). AI HLEG (2018) also notes that discrimination can also occur unintentionally due to bias as a result of incompleteleness or simply 'bad governance models', but does not further elaborate on the exact implication of that latter term (ibid.). If the initial data set is biased, AI HLEG argues that machine learning algorithms may use these biases to follow certain patterns (ibid.). Servoz (2019) writes that it is important that the developers of AI are therefore drawn from diverse backgrounds, differing in an ethnic background too, creating a better representation of society in developing AI as representativeness is at the core of the development too (p. 113). Meanwhile, AI HLEG notes that this will not solve all the problems and argues that data will always have some kind of bias, but that it is important to build identification systems of this bias along with further developing AI (2018, p. 16).

One of the important key points of the development of AI has therefore become the importance of transparency in the development process. In its Communication, the European Commission (2018) stated that 'building trust' is essential in the long term for developing AI. It calls for the EU to promote the innovation of AI but also for the protection of principles such as accountability and transparency (ibid., p. 3). In the 2018 Communication, it was stressed that the way how AI is built should be understood by the average citizen and that nothing about the AI systems should be hidden or not be able to be traced.

> *"To further strengthen trust, people also need to understand how the technology works, hence the importance of research into the explainability of AI systems. Indeed, in order to increase transparency and minimise the risk of bias or error, AI systems should be developed in a manner which allows humans to understand (the basis of) their actions."* (European Commission, 2018, p. 15)

However, in AI theory, the parts of AI algorithms and technologies that are 'hidden' or not 'retrievable' anymore are in a so-called 'black box'. Only a small fraction addresses the black box situation, but most documents do not mention this concept at all and only focus on the good components of AI. However, acting as the 'black box' situation is only a potential future risk was not quite the right approach to take back in 2018. Since then, there has been an agreement that the 'black box' situation is already something that is existing in the already-developed AI too. DG CONNECT (2020) correctly acknowledges that 'most popular AI systems today are of such a complexity that [it] is not possible for humans to explain how they are generated' (p. 24). At the same time, it should be noted that DG CONNECT in the same publication also writes that it does not 'need to present an obstacle' to the working of an AI product (2020, p. 8). In a 2021 publication, the European Commission, the Digital Economy Unit and Joint Research Centre do not seem to agree with this view anymore. They rather note that in democratic societies, the results of using AI should be explainable and should 'not come out of a black box' (Digital Economy Unit et al., 2021, p. 3). Servoz underlined this view already in 2019, by writing that 'opening up black boxes' would be needed before AI tools were used in public services (2019, p. 139). Servoz (2019, p. 150) writes that in its then current form, 'machine learning' was nothing more than a black box and that there was no way to retrace how decisions and recommendations were made. Servoz (2019) continued to warn that the effect would be that people, including those employed in public administration, would 'follow machine recommendations blindly'. Servoz writes:

> *"In its current form, machine learning is unexplainable, like a black box. There is no way to track back from teaching data provided to recommendations/decisions made. As soon as decisions about humans are involved, this is problematic - how do you obtain redress, for example?"* (Servoz, 2019, p. 150).

Within national governments of the EU, we have already seen that this has led to a negative situation. The Netherlands employed a deep-learning algorithm in its tax system to mass surveillance the whole tax-paying population and to find potential fraudsters on the basis of deep-learning algorithms (Wassink, 2021). However, these algorithms were found to be deeply racially biased (Schellevis, 2021). As a result of this, tens of thousands of Dutch households were wrongly marked by this algorithm to be fraudsters. A heightened number of people of non-Dutch descent, born in a foreign country or two with double citizenship were marked as fraudsters (ibid.). Upon researching the cause of the creation of this deeply racist algorithm, it was hard for the researchers to actually find the cause as this was hidden in a black box after all (Wassink, 2021). With all the knowledge that has been available from science about the black box situation and the risks of employing AI-powered mass surveillance, it is nothing less than

worrying to see that this discussion is still alive at the EU-level too. This is for example visible in the argument that DG CONNECT. (2019a) presents, where it is argued that it is not important whether the design of AI is 'beneficial or harmful' but whether the algorithm that is proposed can achieve the desired outcome that is wanted by the designers, giving the opportunity for creators to experiment (ibid., p. 41).

The discussion about employing AI in surveillance systems can perhaps be seen as one of the biggest paradoxes found in the narratives about AI within the EU. A large proportion of the debate around AI focuses on AI-empowered surveillance as one of the most important features of future security. Nevertheless, the way how this should be employed, especially with the involvement of AI remains a widely discussed issue, often neglecting a considerable amount of criticism with regards to racism and the black box situation. Although the EU seems to have reached a consensus that AI offers new possibilities for public safety and that these can be seen as more beneficial over using human capacities in the same situations, not all policy narratives show this consensus, especially within the framework of bias. As said in the paragraph about safety in the subchapters of technological development and defence, technologies such as facial recognition and the use of deep-learning algorithms have been suggested to help with surveillance. However, with regards to surveillance, the legitimacy of using these techniques has been under question, with the statement that they might have potential biases in them due to the backgrounds of their creators. These potential biases are widely discussed in the narratives, but not always exclusively with regards to surveillance techniques. DG CONNECT, however, takes a firm stance with regards to employing AI in surveillance technologies and the risks of doing so. They write that governments should withstand the temptation of wanting to build a so-called 'secure society' by employing pervasive mass surveillance systems (27, p. 11). DG CONNECT reckons that this would be 'extremely dangerous if pushed to extreme levels' (ibid.). On top of that, it says that there should be 'very clear and strict rules' when it comes to using AI in some kind of form of surveillance at all (ibid., p. 20). According to DG CONNECT, commercial use of AI-empowered surveillance techniques should be countered as much as possible (ibid., 11). Interestingly, AI HLEG states that AI will enable governments and states with an 'ever more efficient identification of individuals' (28, p. 11). However, AI HLEG states that AI-powered control techniques need to be sufficiently controlled and monitored insofar that facial recognition is only used to identify European citizens and not to track or follow them (ibid.). AI HLEG states that there should be a clear difference between 'targeted surveillance' and 'mass surveillance' (29, p. 34). In the 2020 White Paper of the European Commission, it is interesting to see that the line of thinking of AI HLEG is followed rather than the line of thinking that DG CONNECT presented in 2019. The White Paper goes even further in their promotion of AI technology with regards

to surveillance and states that AI can perform surveillance tasks that were previously done by humans (26, p. 11). However, here it does again follow some of the thinking of AI HLEG on the employment of AI in mass surveillance. The White Paper states that AI might not work in mass surveillance as it would breach EU data protection and privacy laws and that its widespread use might have negative consequences for the freedom of expression and political freedoms of individuals (ibid.), again showing the many inconsistencies with regards to racial/ethnicity bias in the EU's AI strategy.

4.5. Women as AI's weakest link or an area of opportunity

Next to racial/ethnicity bias, one of the most frequently named types of potential harmful biases in AI is gender bias. The Council of Europe, for example, has warned that AI poses new challenges with regards to gender equality and gender stereotypes. It stated that the use of AI-empowered algorithms can even strengthen existing gender stereotypes and even contribute to an increasing amount of bias and sexism in society. The Council of Europe meanwhile has encouraged policy-makers to add a gender equality perspective in AI policies, including raising the issue of potential gender bias in AI. The Council of Europe Commissioner for Human Rights. (2019) underlines that algorithmic systems often contain bias with relation to gender among other things. The CoE's CEPEJ (2018, p. 57) adds that believing that algorithms can be neutral right now without any biases is the same as believing in a 'myth'.

In 2020, AI HLEG stated that the European Commission should indeed not lose sight of themes such as diversity and inclusion in developing AI and that Member States should require certain 'gender competences' for those who work in AI and in STEM in general (2020, p. 8). DG CONNECT (2019, p. 32) also mentioned these 'gender competences' and mentioned that Member States should set 'incentives to offer gender sensitivity training'. However, what this exactly entails is not elaborated on in any of these documents and therefore, its meaning stays rather vague. The Council of Europe meanwhile has a clear view of what AI and gender stereotypes do to the implementation of AI in society. It states that sexism and sexist behaviour are actually rooted in the reinforcement of these gender stereotypes. Stereotyping lies in the evaluation of other's capacities and needs. For example, if you believe that women need more help with something, it might create the stereotype that women are weaker. Therefore, it is logical that when talking about gender in AI policies, those who face stereotypes should not be portrayed as in need of help or as weak. Rather, the focus should be on the core of the problem, namely the existing structures. Therefore, it is important that those who get marginalised are not portrayed as the 'weaker sort'. Unfortunately, this happens at rare occasions in the narratives. For example:

> *"Vulnerable demographics (e.g. children, minorities, disabled persons, elderly persons, or immigrants) should receive greater attention to the prevention of harm, given their unique status in society. Inclusion and diversity are key ingredients for the prevention of harm to ensure suitability of these systems across cultures, genders, ages, life choices, etc."* (AI HLEG, 2018, p. 9)

Rather than that, DG CONNECT (2019a) recommends public administration to create a systematic evaluation of the extent in which the currently existing structures as we know them ensure the best standards possible to combat potentially harmful practices (p. 40). What remains a fact and is mentioned in this document is that only a small margin of those who work in AI are not a male. DG CONNECT (2019a) mentions that only 12% of the leading machine-learning researchers are female. DG CONNECT calls this rather a loss of talent than something that can cause harm with its existing structures.

> *"Furthermore, the prevalence of male norms and networks in science and technology professions disadvantage women in informal selection and promotion processes. As a consequence, Europe is currently not exploiting its talent pool, causing an annual productivity loss."* (DG CONNECT, 2019a, p. 34)

Servoz (2019, p. 130) argues that public authorities have a key role in monitoring the social impacts of AI. This does not only mean that they have to look at the amount of people of different genders working on AI but also monitor the AI systems they create (ibid.). Rather than just 'nurturing talent', Servoz argues that the 'natural option' is to employ observational techniques that look over how AI is created, looking at different perspectives including non-discrimination and gender (ibid.). With DG CONNECT, this type of view is quite limited. DG CONNECT (2019a) argues that there should be a priority to increase gender equality and 'gender mainstreaming' in all AI policies, but the focus here stays rather on the nurturing of talent of different genders (2019a, p. 35). DG CONNECT please to address gender bias in algorithmic decision-making and says that institutional biases are going to be hard to retrace (ibid.). On top of that, it argues that AI will affect women in the labour market, calling for more research 'to understand the impact on traditionally "female" jobs' (ibid.). Several EU policies with regards to AI go into this direction, putting a clear difference between male and female with regards to AI and the role of women in AI.

Gender balance is a recurring topic in parts of the EU narrative, but it is sometimes not clear what are the push and pull factors that disturb and contribute to this balance. In its Communication, the European Commision (2018) argued that 'no one' should be left behind in the digital transformation that

is initiated by the development of AI (p. 3). Servoz (2019, p. 103) writes that AI presents new challenges for 'gender equality', pointing out that the development of AI might lead to a 'growing gender pay gap' and 'algorithmic and data biases'. Servoz (2019, p. 49) names this type of AI biases especially harmful for women as he argues that the implication of increased use of AI will likely also reinforce more algorithmic bias against women. Servoz (2019, p. 49) names this a 'cause of concern' as about 85% of the machine-learning working force was male in August 2018. By 2020, the European Commission stated that it is needed to upscale the efforts and skills of the workforce when it comes to AI. In this there should be undertaken extra effort to especially increase the amount of women employed in the sector (2020, p. 6). This line of thinking is also found in other EU narratives, including DG CONNECT (2019a), which argues that the EU must invest in the development of human capital when it comes to AI and the fostering of the 'best scientific minds'. The role of women in this is seen as important and as something that needs to be supported an extra bit more:

> *"A doubling-down of efforts is needed to increase the number and proportion of women in science and technology. This will not only increase the available talent pool, but also foster the relevance and quality of research and innovation of AI systems for society as a whole."* (DG CONNECT, 2019a, p. 26)

In 2018, the European Commission also argued in its Communication that it is important to employ more women in the development of AI as well as people of diverse backgrounds, including those with disabilities. However, it stated that it does not end there when it comes to promoting diversity and gender balance in AI (European Commission, 2018). It said that it is also important that those who work in AI get education and training on diversity 'in order to ensure that AI is non-discriminatory and inclusive' (ibid., p. 14).

In some parts of the narrative, when it comes to the 'gender balance' in AI, it is not clear whether this refers to a balance between the women and the rest or between all different genders. From the EU's narratives and studies, it becomes clear that there is very little attention given to what gender actually is and what gender as a concept actually means to AI. Central to the gender question, the EU's main directives as discussed in 26, encompassing the Race Equality Directive, the Directive on equal treatment in employment and occupation and the Directives on equal treatment between men and women in relation to employment and access to all goods, all focus on the classic women vs. men dichotomy (p. 13). In none of the studies, reports and other documents analysed in this study, there was any mention whatsoever to the implicit meaning of 'gender' in the context of AI and when this was given, it specifically referred to the dichotomy between 'women' and 'men'. However, as we know from

the field of gender studies, gender is much more broad than just this dichotomy between men and women, and ecompasses also other groups of people, such as non-binary people and transgenders. Unfortunately, this type of thinking is not embedded yet in the EU's AI narratives, which has at no point given a mention towards non-binary people or transgenders at all.

4.6. Conclusion

This chapter aimed to analyse the EU narratives with regards to domination concepts such as Eurocentrism and androcentrism and bias concepts such as gender bias and racial/ethnicity bias.

One of the ways how domination can manifest itself in the AI policy narrative in the EU is through the manifestation of Eurocentrism. Eurocentrism was quite prevalent in the EU policy narrative through various ways. First, Eurocentrism manifested itself through the continuous focus on economic benefits that AI could bring to the EU. This especially showed itself through the continuous focus of helping to boost and grow the economy. Secondly, this showed itself through the focus of creating a AI world leadership for the EU in which the EU would be the technological centre of the world with regards to AI research. In none of the cases was there ever a mention of helping third-world countries or the Global South with the development of AI.

A second part of the domination narrative often lies in androcentrism. In this study, it was found that androcentrism does not manifest itself so clearly, or at least not as clearly as Eurocentrism. The research found that there were none to very few uses of androcentric language, which includes the use of words such as 'men' instead of 'people'. Androcentrism could only be potentially found in two different ways. First, in the discussion of the use of AI in the security domain as the security domain remains very androcentric in the EU. Secondly, following some recommendations of the Council of Europe with regards to equality and human rights, one can make an argument that the EU's AI narrative is androcentric by the mere fact that it does not clearly oppose androcentrism either.

One of the two ways in which bias was analysed in this master thesis was through the lense of analysing gender bias. The analysis found that the EU narrative did touch on the topic of gender bias and often, this was used as a way of discussing the negative impact that AI can have on specifically women. In much of the narrative, concrete examples of this bias were lacking. The discussion of the exact meaning of gender stayed rather superficial in much of the discussion. Next to that, there was very little discussion of the impact of AI on other genders besides women and men. In fact, there was no mention of any other different genders than these two.

Secondly, within the bias domain, this master thesis asked itself the question of how the discussion of racial/ethnicity bias does show itself in the EU policy narrative. The discussion of racial and ethnicity bias shows itself in the EU narrative, but remains quite undercooked in comparison to the gender bias one. Although practical evidence has shown that AI can be biased racially and ethnicity-wise, there was little attention devoted to that in the EU policy narratives. The EU narrative stays superficial on the content, rather stating that discrimintation should be avoided and equality should be guaranteed. However, it again lacks discussion about bias and problems as racism in general.

5. Conclusion

This master thesis aimed to understand how discussions about bias and domination with regards to AI manifested itself in the EU's policy narratives between 2011 and 2021. Employing a content analysis, this master thesis tried to answer this question by looking at the EU's AI policy narratives through four different concepts, two focusing on bias and two on domination. The concepts did widely manifest itself in the EU policy narrative, but unsurprisingly, some concepts appeared more in depth and were more frequent than others. The concept that appeared the most in the EU policy narrative was 'Eurocentrism', as the EU's policy narrative widely discussed themes and ideas that could be seen as Eurocentric. There was a strong focus on developing the EU with AI in both an economical and technological sense, focused on creating EU world leadership. The concept of bias was also widely discussed, but very clearly not as widely as the benefits of AI in the economical and technological sense. However, to say that there is a strong focus to counter bias in AI in the EU narrative would be misleading. There is some discussion, but it stays rather superficial and inexplicit with regards to the extent of knowledge that there is on the matter in the field.

Answering the sub questions 'How does the domination concept of Eurocentrism manifest itself in the EU AI policy narrative?' and 'How does the domination concept of androcentrism manifest itself in the EU AI policy narrative?', the first insight coming from the analysis of this master thesis showed that domination manifested itself mostly in Eurocentrism, rather than clear domination of men in the AI policy field. This master thesis found that the EU policy on AI focuses on the potential economic gains that AI will bring to Europe and the EU economy. The EU's core economic growth was an important theme and at some points, this focus on growth and the development of AI in Europe could clearly be named part of a specific EU focus on Eurocentrism. A part of this Eurocentrism also manifested itself in the fact that the EU keeps on wanting to be the centre of AI development in the world, becoming a 'technical hub' and a 'world leader' and taking part in a pseudo 'AI war' with other countries that it has to win.

Another part of the domination paradigm meanwhile focused on the androcentric perspective in which AI is still dominated by males, who mainly create the technology. Building on that, the second insight of this master thesis is that with regards to the narrative, this malecentric language use was almost wholly absent. Reasons for that can be that the world and also science and policy making has moved on significantly since the Feminist Epistimology on AI in 1990s that was discussed by Adam (1993) and Ferrando (2014). However, this does not mean that there is no case to be made to state that the EU's policy narratives cannot be called androcentric. The Council of Europe for instance argued that part

of not being androcentric is fighting against potential androcentrism, which the EU seems to be lacking in doing as became clear from the master thesis.

Answering the sub questions 'How does the concept of racial/ethnicity bias manifest itself in the EU AI policy narrative?' and 'How does the concept of gender bias manifest itself in the EU AI policy narrative?', this master thesis found, as its main insight on this topic, that the policy narratives of the EU dedicated a low level of attention to the potential harmful effects AI can have if it were biased in a certain way. Another insight was that while gender bias is addressed in several policy narratives, it became clear that especially bias influenced by racial or ethnical motives remains under-discussed in the policy narrative. It was surprising to see that in some narratives, AI is rather seen as a way to counter potential biases, which is of course, looking at the current status quo in the AI scholarship, not quite the case as Dignam (2020) and Haner & Garcia (2019) have pointed out. The discussion about gender bias is however much stronger and more prevalent in the policy narratives. The EU narrative seems to follow some of the earlier Council of Europe recommendations, which warned that AI could pose new challenges with regards to gender equality and gender stereotypes, potentially creating biases in the AI systems. However, at the same time, a last and clear insight was that the EU usually showed a very superficial explanation of the implicit meaning of 'gender bias' in the context of AI. When this a clear definition was given, it specifically referred to the dichotomy between being biased towards or against either 'women' and 'men'. However, as we know from the field of gender studies, gender is much more broad than just this dichotomy between men and women, and encompasses also other groups of people, such as non-binary people and transgenders. Unfortunately, this type of thinking is not embedded yet in the EU's AI narratives, which has at no point given a mention towards non-binary people or transgenders at all.

The first part of the theoretical framework of this master thesis found its roots in the field of gender studies. The gender studies field has been evolving rapidly in the past few decades, increasingly being used in more and more different fields as a theoretical sub-basis. Most recently, the friction and support between the sub fields of feminist studies, queer studies and gender studies as a whole have only grown (Truman, 2020). With new insights about gender, sexuality and identity, it has become increasingly challenging for political scientists, politicians and policy makers to talk about the topic of gender (Cavanaghan, 2017). If you talk about gender, who do you include exactly? Although the many policy narratives of the EU do in fact talk about the importance of respecting the different genders, it is questionable whether they really do understand the gender questions that are at stake currently. And whether they understand they will have an impact on the development of AI and vice versa. This for

example manifests itself in the clear male-female distinction that is going on with regards to gender equality in the policies. However, as the field of gender studies has shown us researchers, the concept of 'gender' is much more diverse than just the terms 'male' and 'female' (Mazzuca et al., 2020). At the same time, this idea in gender studies also creates friction with part of the more radical feminist scholars, who say that the status of women should not be forgotten in the wider gender debate (ibid). From the narratives, it became clear that the EU is not submerged in the current debate, let alone the current friction between different subfields of the gender studies paradigm. There is in fact little to no interaction with this debate at all in the policy documents, which is shown by the lack of discussion and the lack of term-dropping with regards to gender and sexuality. The reasons for this can only be found in further research. After all, although the EU is a supporter of gender diversity and acceptance and is clearly against discrimination of any kind, the policies do not engage themeselves in the discussion of the gender studies paradigm of AI. Though, it should be pointed out that the gender discussion in the individual Member States of the EU is divisive. While some Member States' politics have adopted the theme of gender diversity and have pleaded for the recognition of non-binary and transgender people, other Member States are still behind on this development.

A second part of the theoretical framework of this master thesis dealt with postcolonial studies on racist AI. Unfortunately, practical situations with the deployment of AI have shown us that algorithms and deep-learning systems can be racist (Dignam, 2020; Monea, 2020; Wassink, 2021). What does this mean for the development of AI when looking towards it from a postcolonial perspective? First, we keep on seeing that the practical situation of the deployment of AI in its current form and fashion is leading towards situations that are undesirable in public administration (Monea, 2020). The childcare benefit scandal in the Netherlands is a prime example of this (Wassink, 2021). Perhaps due to the fact that AI is a new topic to policy narratives at the EU level and still little employed in EU public administration, we are yet to see any example of AI leading to undesired situation with regards to racism and discrimination on the basis of ethnicity. Similarly to the gender situation, the EU narrative keeps on talking about non-discrimination and equality for all its citizens. Nevertheless, the EU seems to not incorporate the wider discussion that goes on about AI within the postcolonial field. From a postcolonial perspective, as already pointed out in the earlier sections about the use of the concept of domination in the EU narrative on AI, the general line of thought cannot be called postcolonial at times. As the master thesis showed, the EU remains very focused on the AI-driven economic growth of the Union. What is noteworthy is that the EU is very focused on maintaining its core position in the world and is not interested in developing AI in other, less affluent parts of the world. In order for the EU to do so, it would

need to first be able to properly address problems with racism, bias and domination at the policy front with regards to AI. Only after that, we could potentially see the EU moving towards an AI policy narrative that could be considered somewhat postcolonial.

This master thesis contributed one of the first policy narrative researches on the EU's AI policies, contributing to a set of insights that have confirmed the assumptions of the theoretical framework in some ways, while the insight on the use of masculine language showed that some of the assumptions coming from the feminist field of analysis did not stand ground when it came to the analysis of the EU's AI strategy. As written in the introduction, this thesis contributed to the current research in three ways. This thesis added to the general study of 'politics of AI' as it found out which elements are important for policy-makers by trying to comprehend the narratives of policy documents, therefore building on early studies, including such as Ossewaarde & Gülenç (2020), Jimenez-Gomez et al. (2020) and Dignam (2020). Second, this thesis contributed to the current discussion in scholarship about diversity and policies by adding a study based on perspectives and ideas coming from the scholarship of gender studies and postcolonial studies. With studying concepts rooted in those two fields, this master thesis added to the large field of current scholarship including Leavy (2018) and Monea (2020) with their individual analysis of the creation of AI has implications on people of different colours and gender identities. Third, this thesis aimed to shed light on the European Union's process of policy-making and the specific function of the European Commission in the agenda-shaping phase. Although the master thesis did not have many findings within that respect, general trends within the EU policy narrative with regards to the general AI policy were still found. In that sense, this thesis added another perspective to the earlier European supranational policy research done by Reis et al. (2020) and Hildebrandt (2020) on the implementation of AI policies in the European Union and to Sarangi & Sharma (2020) on the agenda-setting with regards to AI in governance in general. Meanwhile, as Fischer & Wenger (2021) wrote, the use of content analysis on policy narratives with regards to AI has only been done by a few researchers.

In conclusion, this thesis tried to close a wide knowledge gap in the field of EU and AI studies, but at the same time, as is painted in this master thesis, the study of AI as a policy phenomenon in the EU is a large and fastly moving subject. Therefore, such a 'knowledge gap' can never be fully closed, and as a result, this thesis aims to be an addition to the understanding of the functioning and content of the EU's AI strategy. As this master thesis showed, it helped to underline what is currently being discussed within the AI domain in the EU narrative and which parts do need some further attention. As this master thesis showed, especially a more fine tuned discussion about the implications of AI technology on our society would be welcomed. Especially in the area of bias, a more dedicated discussion would not miss

for policy-makers. The research showed that although bias is being discussed, the discussion in the narrative stays superficial. As a practical implication of this master thesis, it is recommendable for policy-makers to not abstain from focusing on the current gender and postcolonial debates and concerns that continue to exist with regards to developing and deploying AI technology in Europe. However, what remained striking was the large amount of the narrative that tended to be prone to vagueness, inexplicitness and superficiality.

As a result, this master thesis has several recommendations for further research. First, a possible avenue for future research can lie in the analysis of the actual implementation of AI in public administration, focusing on it from a gender studies or postcolonial perspective. Especially one from a gender perspective would be welcome as this thesis found several inconsistencies between the theory that argued that most of the language around AI is malecentric, while that was not the case with regards to the EU AI strategy. As this thesis has shown, there is still a lot to achieve with regards to research on androcentrism and EU (AI) policies as this master thesis was not able to conclude specific things on it. The EU's general approach towards gender and androcentrism is very superficial, but to further pinpoint this, further research could be dedicated to this topic. From a bias perspective, it will especially be interesting to take further steps in policy implementation research with regards to that topic as most of the problems with bias rather manifest itself at the practical level of AI. The content analysis of this master thesis could find that bias is only addressed superficially, but to further analyse the implications that bias has on AI in the EU, there is a need for more practical examples and empirical research. Next to that, different and more refined narrative studies and content analyses would help to create a more coherent and from different perspectives than the gender and postcolonial studies' ones. This is also where the largest limitations of the master thesis lie, as the narrative study performed in this thesis rather stays elementary towards the content of the text. The EU's AI strategy shows a lot of different avenues and pathways which were irrelevant to the focus of this particular master thesis. The thesis was in its shape therefore only able to produce more superficial results on the whole body of the EU's AI strategy. Definite steps for further research using this master thesis definitely lie also in a further concretisation of the gender and postcolonial research on AI and on the EU as combined entities. No matter what, continued research on the societal impact of AI and the implications of gender and postcolonial thinking on AI will be necessary in the coming years.

6. References

Adam, A. (1993). Gendered knowledge — epistemology and artificial intelligence. *Ai & Society : The Journal of Human-Centred Systems*, *7*(4), 311–322. https://doi-org.ezproxy2.utwente.nl/10.1007/BF01891414

Bhambra, G. K. (2009). Postcolonial Europe, or understanding Europe in times of the postcolonial. *The SAGE handbook of European studies* (69-86).

Bohman, J. (2005). *Critical Theory*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/critical-theory/

Buhmann, A. & Fieseler, C. (2021). Towards a deliberative framework for responsible innovation in artificial intelligence, *Technology in Society*, *64*(February). https://doi.org/10.1016/j.techsoc.2020.101475

Bunz, M., Janciute, L., & University of Westminster. Communication and Media Research Institute. (2018). *Artificial intelligence and the internet of things: UK policy opportunities and challenges* (Ser. Camri policy briefs, 2). University of Westminster Press. https://doi-org.ezproxy2.utwente.nl/10.16997/book25

Campbell, M. (2020). Beyond conversational artificial intelligence. *Computer*, *53*(12). https://doi-org.ezproxy2.utwente.nl/10.1109/MC.2020.3024401

Cavaghan, R. (2017). *Making gender equality happen: Knowledge, change and resistance in EU gender mainstreaming*. Taylor & Francis.

Cirillo, D., Catuara-Solarz, S., Morey, C., Guney, E., Subirats, L., Mellino, S., … Mavridis, N. (2020). Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *Npj Digital Medicine*, *3*(1). https://doi-org.ezproxy2.utwente.nl/10.1038/s41746-020-0288-5

Ciston, S. (2019). Intersectional artificial intelligence is essential: polyvocal, multimodal, experimental methods to save ai. *Journal of Science and Technology of the Arts*, *11*(2 Special Issue), 3–8. https://doi-org.ezproxy2.utwente.nl/10.7559/citarj.v11i2.665

Dignam, A. (2020). Artificial intelligence, tech corporate governance and the public interest regulatory response. *Cambridge Journal of Regions, Economy and Society*, *13*(1), 37–54. https://doi-org.ezproxy2.utwente.nl/10.1093/cjres/rsaa002

Djeffal, C. (2020). AI, Democracy and the Law.*The democratization of Artificial Intelligence: net politics in the era of learning algorithms* (pp. 255-284), Sudmann, A. (Ed.). Transcript-Verlag. https://doi-org.ezproxy2.utwente.nl/10.14361/9783839447192

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Open*. https://doi.org/10.1177/2158244014522

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing a content analysis. *African Journal of Emergency Medicine*, *7*(3), 93-99.

Fischer, S. C. & Wenger, A. (2021). Artificial Intelligence, Forward-Looking Governance and the Future of Security. *Swiss Political Science Review*, *27*(1), 170–179. doi:10.1111/spsr.12439

Ferrando, F. (2014). Is the post-human a post-woman? cyborgs, robots, artificial intelligence and the futures of gender: a case study. *European Journal of Futures Research*, *2*(1), 1–17. https://doi-org.ezproxy2.utwente.nl/10.1007/s40309-014-0043-8

Franzki, H. (2012). Eurocentrism. *InterAmerican Wiki: Terms - Concepts - Critical Perspectives*.

https://www.uni-bielefeld.de/cias/wiki/e_Eurocentrism.html.

Fuchs, C. (2017). Critical social theory and sustainable development: The role of class, capitalism and domination in a dialectical analysis of un/sustainability. *Sustainable Development*, *25*(5), 443-458

Girasa, R. J. (2020). *Artificial intelligence as a disruptive technology : economic transformation and government regulation*. Palgrave Macmillan..

Halberstam, J. (1991). Automating gender: postmodern feminism in the age of the intelligent machine. *Feminist Studies*, *17*(3), 439–460.

Haner, J. & Garcia, D. (2019). The Artificial Intelligence Arms Race: Trends and World Leaders in Autonomous Weapons Development. *Global Policy*, *10*(3), 331-337.

Hansson, S. (2018). Government Communication. In Wodak, R. & Forchtner, B. (Eds), *The Routledge Handbook for language and politics*. (pp. 326-341). Routledge.

Hegarty, P. (2007). From genius inverts to gendered intelligence: Lewis Terman and the power of the norm. *History of Psychology*, *10*(2), 132.

Hernández-Orallo José. (2017). Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement. *Artificial Intelligence Review: An International Science and Engineering Journal*, *48*(3), 397–447. https://doi-org.ezproxy2.utwente.nl/10.1007/s10462-016-9505-7

Hildebrandt, M. (2020). The Artificial Intelligence of European Union Law. *German Law Journal*, *21*(1), 74-79. doi:10.1017/glj.2019.99

Jimenez-Gomez, C. E., Cano-Carrillo, J., & Falcone Lanas, F. (2020). Artificial intelligence in government. *Computer*, *53*(10). https://doi-org.ezproxy2.utwente.nl/10.1109/MC.2020.3010043

Kane, T. B. (2019). Artificial intelligence in politics: establishing ethics. *Ieee Technology and Society Magazine*, *38*(1), 72–80. https://doi-org.ezproxy2.utwente.nl/10.1109/MTS.2019.2894474

Kassarjian, H. (1977). Content Analysis in Consumer Research. *Journal of Consumer Research*, *4*(1), 8-18.

Kappler, S., & Lemay-Hébert, N. (2019). From power-blind binaries to the intersectionality of peace: connecting feminism and critical peace and conflict studies. *Peacebuilding*, *7*(2), 160-177.

Keita, L. (2020). Eurocentrism and the Contemporary Social Sciences. *Africa Development / Afrique Et Développement*, *45*(2), 17-38.

Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In *Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14-16).

Lombardo, E. & Verloo, M. (2009). Contentious citizenship: Feminist debates and practices and European challenges. *Feminist Review*, *92*, 108-128. https://www.jstor.org/stable/40664035

Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. *Human Behavior and Emerging Technologies*, *1*(2), 141–148. https://doi-org.ezproxy2.utwente.nl/10.1002/hbe2.140

Mazzuca, C., Majid, A., Lugli, L., Nicoletti, R., & Borghi, A. (2020). Gender is a multifaceted concept: evidence that specific life experiences differentially shape the concept of gender. *Language and Cognition.*

McQuillan, D. (2020). Political Affinities of AI. *The democratization of Artificial Intelligence: net politics in the era of learning algorithms* (pp. 163-174), Sudmann, A. (Ed.). Transcript-Verlag. https://doi-org.ezproxy2.utwente.nl/10.14361/9783839447192

Mills, A. J., Durepos, G., & Wiebe, E. (2010). *Encyclopedia of case study research (Vols. 1-0*). Thousand

Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412957397

Monea, A. (2020). Race and Computer Vision. *The democratization of Artificial Intelligence: net politics in the era of learning algorithms* (pp. 189-208), Sudmann, A. (Ed.). Transcript-Verlag. https://doi-org.ezproxy2.utwente.nl/10.14361/9783839447192

Neuendorf, K. A. (2017). *The content analysis guidebook* (Second). SAGE.

Onar, N., & Nicolaïdis, K. (2013). The Decentring Agenda: Europe as a post-colonial power. *Cooperation and Conflict*, *48*(2), 283-303.

Ossewaarde, M.R.R. & Gülenç, E. (2020). National Varieties of Artificial Intelligence Discourse: Myth, Utopianism, and Solutionism in West European Policy Expectations. *Computers*, *53*(11), 53-61.

Phillips, N., & Hardy, C. (2002). Discourse analysis. *SAGE Publications, Inc*.

Reis, J., Santo, P. & Melão, N. (2020). Impact of artificial intelligence research on politics of the European Union Member States: the case study of Portugal. *Sustainability*, *12*(6708). https://doi-org.ezproxy2.utwente.nl/10.3390/su12176708

Rosenthal, G. (2018). *Interpretive social research - an introduction*. Universitätsverlag Göttingen. https://doi-org.ezproxy2.utwente.nl/10.17875/gup2018-1103

Sarangi, S., & Sharma, P. (2019). *Artificial intelligence: evolution, ethics and public policy*. Routledge. https://doi-org.ezproxy2.utwente.nl/10.4324/9780429461002

Schellevis, J. (2021, 1 February). *Fraude opsporen of gevaar voor discriminatie? Gemeenten gebruiken 'slimme' algoritmes*. NOS Nieuws. https://nos.nl/artikel/2366864-fraude-opsporen-of-gevaar-van-discriminatie-gemeenten-gebruiken-slimme-algoritmes

Schippers, B. (2020). *Artificial intelligence and democratic politics*. Political Insight, 11(1), 32–35. https://doi-org.ezproxy2.utwente.nl/10.1177/2041905820911746

Sudmann, A. (Ed.). (2020). *The democratization of Artificial Intelligence: net politics in the era of learning algorithms* (Ser. Ki-kritik, 1). Transcript-Verlag. https://doi-org.ezproxy2.utwente.nl/10.14361/9783839447192

Susen, S. (2009). Between emancipation and domination: Habermasian reflections on the empowerment and disempowerment of the human subject. *Pli: The Warwick Journal of Philosophy*, *20*, 80-110.

Taylor, F. (2003). Content Analysis and Gender Stereotypes in Children's Books. *Teaching Sociology*, *31*(3), 300-311.

Tickner, J. A., & True, J. (2018). A Century of International Relations Feminism: From World War I Women's Peace Pragmatism to the Women, Peace and Security Agenda. *International Studies Quarterly*, *62*(2), 221-233.

Tomalin, M., Byrne, B., Concannon, S., Saunders, D., & Ullmann, S. (2021). The practical ethics of bias reduction in machine translation: why domain adaptation is better than data debiasing. *Ethics and Information Technology*, 1-15.

Truman, S. E. (2020). *Feminist new materialisms*. SAGE Publications Limited.

Wassink, J. (2021, 20 January). *Who oversees Artificial Intelligence systems?*. Delta: Journalistic Platform TUDelft. https://www.delta.tudelft.nl/article/who-oversees-artificial-intelligence-systems

Weber, M., Caplan, S., Ringold, P., & Blocksom, K. (2017). Rivers and streams in the media: A content analysis of ecosystem services. *Ecology and Society*, *22*(3).

Appendix A

AI Watch

Studies, strategies and reports

1. Craglia, M., De Nigris, S., Gomez-Gonzales, E., Gomez Gutierrez, E., Martens, B., Iglesias Portela, M., Vespe, M., Schade, S. Micheli, M., Kostev, A., Mitton, I., Vesnic Alujevic, L., Pignatelli, F., Hradec, J., Nativi, S., Sanchez Martin, J.I., Hamon, R., Junklewitz, H. (2020). *Artificial Intelligence and Digital Transformation: early lessons from the COVID-19 crisis*. European Commission. https://knowledge4policy.ec.europa.eu/publication/artificial-intelligence-digital-transformation-early-lessons-covid-19-crisis_en
2. De Nigris, S., Craglia, M., Nepelski, D., Hradec, J., Gomez-Gonzales, E., Gomez Gutierrez, E., Vazquez-Prada Baillet, M., Righi, R., De Prato, G., Lopez Cobo, M., Samoili, S., Cardona, M. (2020). *AI Watch : AI Uptake in Health and Healthcare, 2020*. European Commission. https://knowledge4policy.ec.europa.eu/news/report-ai-uptake-health-healthcare-published_en
3. Martinez Plumed, F., Gomez Gutierrez, E., Hernández-Orallo, J. (2020). *AI Watch: Assessing Technology Readiness Levels for Artificial Intelligence*. European Commission. https://knowledge4policy.ec.europa.eu/news/ai-watch-assessing-technology-readiness-levels-artificial-intelligence_en
4. Misuraca, G. & Van Noordt, C. (2020). *AI Watch - Artificial Intelligence in public services*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/artificial-intelligence-public-services_en
5. Nepelski, D., & Sobolewski, M. (2020). *Estimating investments in General Purpose Technologies: The case of AI Investments in Europe*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/estimating-investments-general-purpose-technologies-case-ai-investments-europe_en
6. Samoili, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F. & Delipetrev, B. (2020). *Defining Artificial Intelligence*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/defining-artificial-intelligence_en
7. Samoili, S., Righi, R., Cardona, M., Lopez Cobo, M., Vasquez-Prada Baillet, M., & De Prato, G. (2020). *TES analysis of AI Worldwide Ecosystem in 2009-2018*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/ai-landscape-2009-2018-report_en
8. Van Noordt, C., Misuraca, G., Mortati, M., Rizzo, F., & Timan, T. (2020). *AI Watch - Artificial Intelligence for the public sector*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/peer-learning-workshops-reports_en
9. Van Roy, Vincent. (2020). *AI Watch - National strategies on Artificial Intelligence: A European perspective in 2019*. AI Watch. https://knowledge4policy.ec.europa.eu/ai-watch/national-strategies-artificial-intelligence_en

Council of Europe

Recommendations

10. Recommendation CM/Rec(2010)13 adopted by the Committee of Ministers of the Council of Europe on 23 November 2010 and explanatory memorandum. Council of Europe. https://rm.coe.int/16807096c3
11. Recommendation 2102 (2017)1 Technological convergence, artificial intelligence and human rights. Council of Europe. http://semantic-pace.net/tools/pdf.aspx?doc=aHR0cDovL2Fzc2VtYmx5LmNvZS5pbnQvbncveG1sL1hSZWYvWDJILURXLWV4dHIuYXNwP2ZpbGVpZD0yMzcyNiZsYW5nPUVO&xsl=aHR0cDovL2Fzc2VtYmx5LmNvZS5pbnQvbncveG1sL3hzbC1mby9QZGYvWFJlZi1XRC1BVC1YTUwyUERGLnhzbA==&xsltparams=ZmlsZWlkPTIzNzI2
12. Recommendation CM/Rec(2019)1 of the Committee of Ministers to member States on preventing and combating sexism. Council of Europe. https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168093b26a
13. Recommendation CM/Rec(2019)10 of the Committee of Ministers to member States on developing and promoting digital citizenship education. Council of Europe. https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=090000168098de08
14. Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. Council of Europe. https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

Declaration

15. *Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes*. (2019). Council of Europe. https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168092dd4b

Guidelines

16. Directorate General of Human Rights and Rule of Law. (2019). *GUIDELINES ON ARTIFICIAL INTELLIGENCE AND DATA PROTECTION*. Council of Europe. https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8
17. Directorate General of Human Rights and Rule of Law. (2021). *Guidelines on Facial Recognition. Council of Europe*. https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3

Studies, strategies and reports

18. Council of Europe. (2016a). *HUMAN RIGHTS AND BUSINESS*. Council of Europe. https://rm.coe.int/human-rights-and-business-recommendation-cm-rec-2016-3-of-the-committe/16806f2032

19. Council of Europe. (2016b). *Internet governance - Council of Europe Strategy 2016-2019 (2016)*. Council of Europe. https://edoc.coe.int/fr/internet/7128-internet-governance-council-of-europe-strategy-2016-2019.html
20. Council of Europe. (2018). *Guidelines to respect, protect and fulfil the rights of the child in the digital environment*. Council of Europe. https://rm.coe.int/guidelines-to-respect-protect-and-fulfil-the-rights-of-the-child-in-th/16808d881a
21. Council of Europe. (2019). *Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025)*. Council of Europe. https://rm.coe.int/strategic-action-plan-final-e/16809c3af1
22. Council of Europe Commissioner for Human Rights. (2019). *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*. Council of Europe. https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/16809046e64
23. EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ). (2018). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*. Council of Europe. https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c

European Commission

Brochures

24. DG CONNECT. (2021). *Artificial intelligence, real benefits*. European Commission. https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-real-benefits
25. Robotics and Artificial Intelligence Innovation and Excellence (Unit A.1). (2021). *Excellence and Trust in AI — Brochure*. European Commission. https://ec.europa.eu/digital-single-market/en/news/excellence-and-trust-ai-brochure

White paper

26. European Commission. (2020). *White Paper on Artificial Intelligence - A European approach to excellence and trust*. European Commission. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Recommendations

27. DG CONNECT. (2019a). *Policy and investment recommendations for trustworthy Artificial Intelligence*. European Commission. https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence

Guidelines

28. AI HLEG. (2018). *Draft Ethics guidelines for trustworthy AI*. European Commission. https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-
29. AI HLEG. (2019). *Ethics guidelines for trustworthy AI*. European Commission. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai Ai

Studies, strategies and reports

30. AI HLEG. (2019d, 14 September). *Sectoral Considerations on the Policy and Investment Recommendations*. European Commission. https://op.europa.eu/nl/publication-detail/-/publication/6d266f0f-f7c3-11ea-991b-01aa75ed71a1
31. AI HLEG. (2020). *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment*. European Commission. https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment
32. Coordination of Digital Strategy and Green ICT (Unit F.1). (2020). *Shaping the digital transformation in Europe - Working paper: Economic potential*. European Commission. https://ec.europa.eu/digital-single-market/en/news/shaping-digital-transformation-europe-working-paper-economic-potential
33. Delipetrev, B., Tsinaraki, C., Kostic, U. (2020). *Historical Evolution of Artificial Intelligence*. European Commission. https://knowledge4policy.ec.europa.eu/publication/historical-evolution-artificial-intelligence_en
34. DG CONNECT. (2018). *The European Artificial Intelligence landscape*. European Commission. https://ec.europa.eu/digital-single-market/en/news/european-artificial-intelligence-landscape
35. DG CONNECT. (2019b). *A definition of Artificial Intelligence: main capabilities and scientific disciplines*. European Commission. https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines
36. DG CONNECT. (2020). *Trends and Developments in Artificial Intelligence - Challenges to the Intellectual Property Rights Framework*. European Commission. https://ec.europa.eu/digital-single-market/en/news/trends-and-developments-artificial-intelligence-challenges-intellectual-property-rights-0
37. Digital Economy, Recovery Plan and Skills (Unit F.4) (2020). *European enterprise survey on the use of technologies based on artificial intelligence*. European Commission. https://ec.europa.eu/digital-single-market/en/news/european-enterprise-survey-use-technologies-based-artificial-intelligence
38. European Commission. (2018, 24 April). *Communication Artificial Intelligence for Europe*. European Commission. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51625
39. Digital Economy Unit, Joint Research Centre, European Commission. (2021). *Public Sector Modernisation for EU Recovery and Resilience*. European Commission.

https://ec.europa.eu/jrc/sites/jrcsh/files/policy_brief_public_sector_modernisation_for_eu_recovery_and_resilience_final_20210210.pdf

40. Microelectronics and Photonics Industry (Unit A.3). (2020). *Emerging Technologies in Electronic Components and Systems (ECS) - Opportunities Ahead*. European Commission. https://ec.europa.eu/digital-single-market/en/news/emerging-technologies-electronic-components-and-systems-ecs-opportunities-ahead-0
41. Nepelski, D., Barbas, T., Craglia, M., Dalla Benetta, A., De Nigris, S., Denin, N.I., Gómez-González, E., Hamon, R., Hoeveler, A., Hradec, J., Junklewitz, H., Munoz-Pineiro, M.A., Nicholl, C., Nordvik, J.P., Pakalin, S., Sanchez Martin, J.I. (2021). *How can Europe become a global leader in AI in health?*. European Commission. https://knowledge4policy.ec.europa.eu/publication/how-can-europe-become-global-leader-ai-health_en
42. Servoz, M. (2019). *The future of work? Work of the future!*. European Commission. https://ec.europa.eu/digital-single-market/en/news/future-work-work-future

OECD

Recommendations

43. Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449) (2019). OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449