A Systematic Assessment of National Artificial Intelligence Policies: Perspectives from the Nordics and Beyond

Niels van Berkel
nielsvanberkel@cs.aau.dk
Aalborg University
Aalborg, Denmark

Eleftherios Papachristos
papachristos@cs.aau.dk
Aalborg University
Aalborg, Denmark

Anastasia Giachanou
angia9@upv.es
Universitat Politècnica de València
Valencia, Spain

Simo Hosio
simo.hosio@oulu.fi
Oulu University
Oulu, Finland

Mikael B. Skov
dubois@cs.aau.dk
Aalborg University
Aalborg, Denmark

ABSTRACT

Echoing the evolving interest and impact of artificial intelligence on society, governments are increasingly looking for ways to strategically position themselves as both innovators and regulators in this new domain. One of the most explicit and accessible ways in which governments outline these plans is through national strategy and policy documents. We follow a systematic search strategy to identify national AI policy documents across twenty-five countries. Through an analysis of these documents, including topic modelling, clustering, and reverse topic-search, we provide an overview of the topics discussed in national AI policies and contrast the differences between countries. Furthermore, we analyse the frequency of eleven ethical principles across our corpus. Our paper outlines implications of the differences between geographical and cultural clusters in relation to the future development of artificial intelligence applications.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; *HCI theory, concepts and models*; • **Social and professional topics → Government technology policy**.

KEYWORDS

Artificial Intelligence, AI, policy, ethics, topic modelling, strategy, national guidelines

ACM Reference Format:
Niels van Berkel, Eleftherios Papachristos, Anastasia Giachanou, Simo Hosio, and Mikael B. Skov. 2020. A Systematic Assessment of National Artificial Intelligence Policies: Perspectives from the Nordics and Beyond. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20), October 25–29, 2020, Tallinn, Estonia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3419249.3420106

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
NordiCHI '20, October 25–29, 2020, Tallinn, Estonia
© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-7579-5/20/10…$15.00
https://doi.org/10.1145/3419249.3420106

1 INTRODUCTION

Following an increase in computational power, opportunities for the collection of *in situ* data, as well as the theoretical development of better algorithms, applications of Artificial Intelligence (AI) have considerably expanded over the past years. Such applications range from everyday services (*e.g.*, multimedia recommendation, writing support) to highly impactful domains (*e.g.*, clinical support, criminal risk prediction [21]). Given this increasing impact of AI on societies, as well as the competitive advantages and potential risks introduced by AI technology, governments across the globe have expressed the need for increased regulatory oversight and policy development (see *e.g.*, [29, 41]). In May 2019, forty-two countries adopted the first set of intergovernmental policy guidelines on AI as set forward by the OECD (Organisation for Economic Co-operation and Development) [34]. These guidelines promote concepts such as accountability, fairness, and transparency – as has been previously advocated within the field of HCI [1, 44, 48] and the wider Computer Science community [2].

With inherent cultural, economic, and moral differences between countries [24], perceptions on how AI applications should operate do not follow one common or global mindset [6, 25]. As such, various governments (primarily those located in the Global North) have set out to draft national policies for AI [25] – outlining the opportunities and threats as perceived for their respective country. Various geographically and/or culturally aligned countries have furthermore developed local guidelines on AI. For example, leaders of the Nordic-Baltic region have signed an agreement to reinforce their cooperation on AI [29]. Similarly, the European Union (EU) set out a coordinated approach to AI, stressing the values of the EU as a differentiating factor; "*the EU's sustainable approach to technologies creates a competitive edge, by embracing change on the basis of the Union's values*" [14]. However, exactly how much these national outlooks on AI differ between countries remains unclear.

In this paper, we set out to identify differences in perspectives on AI between countries and geographical regions through a systematic comparison of national policy documents. First, we identify relevant national policy documents through a systematic review process following the PRISMA framework [42]. This resulted in a total of 25 relevant national policy documents. Second, we perform a quantitative content analysis of the texts, including topic modelling to discern the issues discussed and a frequency analysis of

various ethical principles. Our analysis concludes with an initial qualitative assessment on the ethical principles discussed in order to obtain richer insights into the policy's ethical discussions. We find that geographically related countries show a large degree of semantic similarities, highlighting an overlap in AI policy between countries. Our analysis furthermore reveals significant differences between the focus on ethical principles, as well as contrasts between geographical clusters on how these ethical principles can be attained.

In line with this year's NordiCHI theme of 'shaping experience, shaping society', we provide an in-depth analysis of national AI policy documents. Through this analysis, we contribute to a better understanding of the different perspectives on the opportunities and threats of AI to society. Our work highlights that not one uniform perspective on 'AI' exists. These findings can be used to inform the design of future Human-AI interactions tailored to the needs and desires of end users across different cultures and localities.

2 RELATED WORK

Recent literature shows an increasing interest in exploring the future real-world implications of (near-future) AI systems. Pargman et al., in a 2018 NordiCHI workshop, asked participants to imagine the future of computing and wisdom by writing a 'fictional abstract' of research papers that will be written in 2068 [35]. These fictional abstracts highlight a number of concerns and questions regarding the implication of future AI technology[1]. A number of studies have aimed to quantify the desired behaviour of domain-specific AI systems. One of the most large-scale and well-known studies on this topic is the 'Moral Machine Experiment', in which participants are asked to choose the most favourable outcome of two traffic accident scenarios (*e.g.*, killing three pedestrians or two drivers) [6]. This information could be used to inform the behaviour of self-driving cars in an accident scenario. As participant responses are collected from a global audience, the authors are able to identify three global clusters of moral viewpoints towards the behaviour of automated vehicles [6]. These clusters, labelled as 'Western', 'Eastern' and 'Southern', highlight differences in moral preferences in terms of machine ethics. Awad et al. also find sub-clusters among the three aforementioned clusters, *e.g.* Scandinavian countries and Commonwealth countries within the Western cluster [6].

These differences in preference between cultures are not exclusive to the operation of AI systems. The HCI literature has explored the effect of culture in a variety of studies. Haddad et al. compare two different interface designs among older adult Caucasians and East Asians and find preference differences in interface design [22]. Social media usage has also been investigated cross-culturally, highlighting differences in self-presentation [53], self-disclosure [52], as well as tasks such as brainstorming [47]. These studies have typically contrasted American and Chinese user populations. The extensive work by Geert Hofstede on cultural dimensions does, however, highlight that there are vast differences between the culture and values of societies that go beyond a US-China contraposition [24]. The aforementioned study by Awad et al. highlights how these differences in culturally shared values can come to light in the preferences of individuals towards technology [6].

[1] https://futuresnordichi.wordpress.com/participants-and-abstracts/

In line with these academic studies, governments have begun to stress the need for AI to reflect their cultural values, as can be found in both national and supra-national documents. For example;

- The White House's website on AI describes the need for an "*AI with American values*" [41], reflecting principles around freedom, human rights, law, privacy, and opportunity for all.
- A whitepaper by the EU's European Commission stresses the need for the EU to support AI development as based on "*European values and rules*" [15].
- Nordic-Baltic collaboration on AI stating the development of shared "*ethical and transparent guidelines, standards, principles and values*" [29].

In this paper we aim to uncover, through an analysis of national AI policy documents, the focus area(s) of individual national AI policies and the similarities and dissimilarities between national documents.

2.1 Analysing Perceptions on Artificial Intelligence

Our current work builds most prominently on prior work by Jobin et al., who analysed emerging principles surrounding 'ethical AI' [25]. Following a trend among governmental and commercial entities to stress the need for ethics in AI, Jobin et al. analyse and compare 84 documents describing ethical principles or guidelines. Their results highlight a total of eleven unique ethical principles, with emerging convergence around five principles (transparency, justice and fairness, non-maleficence, responsibility, and privacy). Other principles, such as solidarity, dignity, and sustainability, are mentioned considerably less [25]. Despite this emerging convergence among a section of the principles, the authors also highlight significant semantic and conceptual differences between the documents in terms of interpretation and recommendations of the areas concerned.

Researchers in field of HCI have also conducted meta-analyses on specific applications of AI. Völkel et al. analyse the proceedings of IUI (International Conference on Intelligent User Interfaces) to identify what is deemed as 'intelligent' in the context of user interfaces [46]. In their analysis, Völkel et al. investigate intelligence across three different levels; as an entity (*i.e.*, what is considered as intelligent), the co-descriptors of this entity, and the actions of this intelligent entity. Their results reveal that 'adaptation', 'automation', and 'interaction' are the most commonly used co-descriptors when referring to intelligence in user interfaces [46]. Abdul et al. applied topic modelling to analyse a large corpus of papers from the HCI and Machine Learning communities [1]. By analysing this topic across different venues and research fields, their results highlight both the overlap and gaps in interdisciplinary research between HCI and AI research [1].

These works show that, through analysis of existing literature, we can identify differences in the perception and use of artificial intelligence in specific domains. Here, we similarly aim to uncover how a specific group of experts, in our case national governments, differently position themselves on the topic of AI.

2.2 Grey Literature in Research

The grey literature consists of a wide variety of documents and sources (*e.g.*, blogs, industrial whitepapers) and is positioned outside

the scope of the literature that is produced in academic publication channels (referred to as 'white literature' [3]). Numerous definitions of the grey literature exist, with the widely used 'Prague definition' stating that grey literature "*stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories, but not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body.*" [36]. As such, grey literature is often unstructured, not archived in a structured manner, and not (independently) peer-reviewed.

Despite these shortcomings in comparison with peer-reviewed academic publications, analysis of the grey literature allows researchers to identify emerging trends and practices in industry [39] and government [25] on a global scale. As such, there has been a rising interest among researchers to study the grey literature [28], with relevant example application areas utilising the grey literature to study the uptake of a novel technology by software practitioners [40], identify concerns around cloud computing security and privacy [37], and – most relevant to our current work – present a thematic analysis of AI policy documents [25]. In this paper, we analyse the grey literature of national AI policy documents to obtain a better understanding of various national perspectives towards artificial intelligence. As national policy documents highlight potential future areas of investment, our analysis furthermore offers an insight into the focus of national governments in terms of future AI developments.

3 METHOD

As highlighted by prior work based on grey literature [25, 37, 39], the selection of the analysed document is key to enable a sound analysis. Therefore, we developed a structured search strategy based on the PRISMA-ScR framework (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [42], as is common in the identification of grey literature (see *e.g.*, [25]). We detail the identification, inclusion, and exclusion criteria in the following sections and provide an illustrated overview of our AI policy identification search strategy in Figure 1.

3.1 Policy Identification & Screening

With the goal of mapping the recent global AI policy landscape, our first step was to identify as many national policy documents as possible. Although AI as a research field has existed since the 1950s, it is only due to recent developments that AI is now at the forefront of public debate, highlighting both economic opportunities as well as ethical implications on a national level. Hence, in order to capture these perspectives from a governmental perspective we limit our search strategies to policies that have been published by official institutions in the last five years. The document search was conducted from the 17th of February to the 5th of March 2020 by two of the paper's authors. Discrepancies were resolved through debate and reach of consensus.

Given our focus on policy documents, considered as part of the grey literature, no official database exists which covers all global national policy documents on AI. Therefore, we devised a multistage screening strategy (see Figure 1), in line with the prior work by Jobin et al. [25]. We first conducted a manual search based on four relevant link hub pages [11, 17, 31, 50], from which we were able to collect 120 initial document sources. Subsequently, we conducted a Google web search to identify documents that could have been missed by our first approach. For each of the 62 OECD countries, we conducted individual searches by using the country name and the subsequent keywords: [AI Strategy], [Artificial Intelligence Strategy], [AI Policy], [Artificial Intelligence Policy], [AI Initiative], [Artificial Intelligence Initiative]. We reviewed twenty results for each keyword-country combination. Furthermore, we made sure that for each search we opened a new incognito browser session as to not affect our search result based on our prior browsing history. In total, we screened 7440 search results (62 OECD countries x 6 keywords * 20 results) and we identified 84 relevant document sources as based on the aforementioned criteria.

Out of the 84 documents retrieved through Google search, 52 were duplicates when compared to the link hub results and were subsequently removed. The remaining document sources were scanned for publication status (drafts were excluded), language (non-English were removed if no official translated version was provided), and broken links. We decided to not include unofficial translations or to make use of automated translation tools. Given our goal of a quantitative analysis of the corpus, the language use is an important factor in our analysis. Furthermore, we excluded numerous sources that were linking to press releases or news articles announcing the development of a national AI policy, but were still unpublished at the time of our search. Following this process, another 52 sources were excluded from our analysis – leaving a total of 100 documents.

3.2 Document Inclusion

We next executed an inclusion stage [42], in which we manually searched for additional documents. This search was based on information collected in the prior stage, *e.g.* identifying sources for which we encountered a broken link, or searching for a final version of a document of which we encountered a publicly released draft. Through link chaining, manual Google searches, and by visiting individual government websites, we were able to identify nine additional sources. The resulting 109 documents were subsequently evaluated for eligibility.

3.3 Eligibility & Exclusion

Two of the paper's authors examined the title, abstract, table of contents, and source organisation of all 109 documents to assess their eligibility for this study. First, we excluded any documents which were not officially issued by a national government, for example academic papers or independent reports by consultancy firms (25 documents). Furthermore, we excluded sources which could not be considered as a national policy or strategy framework (*e.g.*, industry reports, ethical frameworks – 22 documents) or which did not have AI as their main focus (13 documents). Additionally, we excluded policy documents with a focus on only one dimension or a specific application area of AI, for example, autonomous driving, robotics (15 documents). This process resulted in the exclusion of 75 sources overall. The remaining 34 documents which met our eligibility criteria were passed to the last stage of our selection

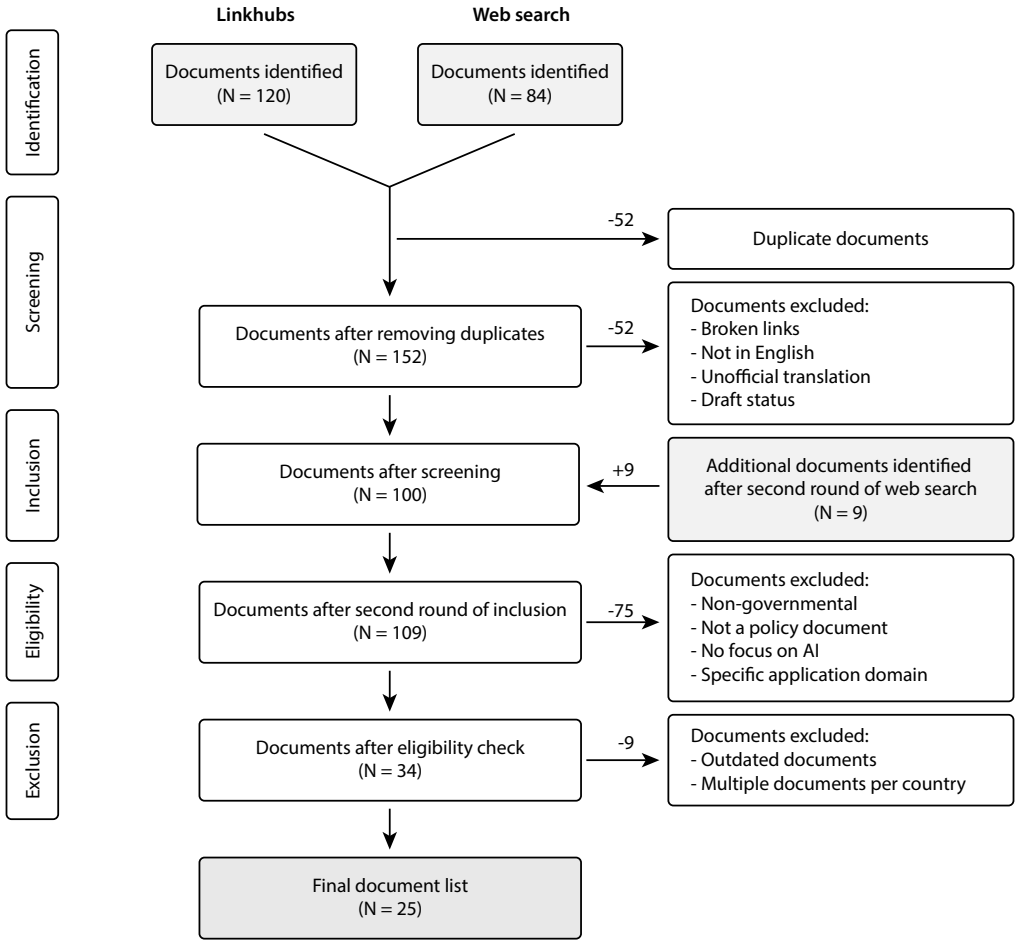


Figure 1: Identification, screening, inclusion, eligibility, and exclusion procedure following the PRISMA-ScR framework [42]. The overview highlights how we initially identified 204 documents, of which a total of 25 were finally selected.

process. Following our objective to select one primary document per country which best reflects the most updated and wide-ranging discourse about AI policy, we sorted our results by country and identified countries with more than one document. Nine more documents were removed during this stage, which resulted in the final set of 25 AI policy documents.

4 RESULTS

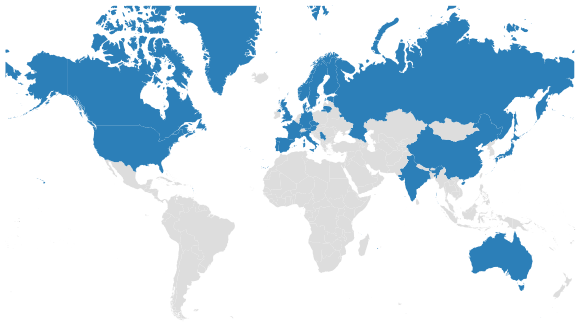
Our policy identification search revealed that a large number of countries do not have, or are currently in the process of developing, a national AI policy. Consequently, we believe that the time range of our search (5 years into the past) was suitable for our research purpose. We analyse a total of 25 identified documents. We highlight the countries included in our analysis in Figure 2 and include the full list of countries and respective abbreviation. In analysing the documents, we adhere to the following procedure;

- For each document we removed the cover page, the running headline, the (executive) abstract, and the table of contents when this was applicable.
- We follow established text-analysis methods and remove stop words (*e.g.*, have, the, about), numbers and punctuation marks, words that consist of less than 3 characters. To implement the pre-processing steps we use the *nltk*[2] Python library, one of the most common libraries in computational linguistics.

Following this procedure, the documents had an average length of 10426 words (median = 8983, SD = 7874). The documents differ considerably in their total length, with the shortest document totalling 1697 words (Sweden) and the longest document totalling 36969 words (France).

First, we estimate the semantic similarity among the documents using Universal Sentence Encoder (USE) [9]. USE is a language model that transforms sentences into high-dimensional vectors and is used in many Natural Language tasks such as semantic similarity and text classification. To calculate the similarities among the documents, we use the text of the documents as input and the

[2]https://www.nltk.org/



Australia	AUS	Lithuania	LTU
Austria	AUT	Luxembourg	LUX
Canada	CAN	Malta	MLT
China	CHN	Norway	NOR
Czech Republic	CZE	Portugal	PRT
Denmark	DNK	Russia	RUS
Estonia	EST	Serbia	SRB
Finland	FIN	Singapore	SGP
France	FRA	Spain	ESP
Germany	DEU	Sweden	SWE
India	IND	United States	USA
Italy	ITA	United Kingdom	GBR
Japan	JPN		

Figure 2: Overview of the 25 countries included in the analysis (marked in blue).

pre-trained Universal Sentence Encoder that is publicly available in Tensorflow-hub[3].

The outcome of this analysis was a 25x25 proximity matrix in which the overall semantic similarity between any pair of countries was indicated on a scale from zero (no similarity) to one (identical). We conducted Multidimensional Scaling (MDS) analysis on this matrix to create a two-dimensional representation of the results to increase interpretability of the data. MDS is a very common dimension reduction and visualisation technique that can be encountered both in the HCI and text analysis literature [8, 51]. Semantic similarity is transformed into Euclidean distance in a low-dimensional visualisation which allows easy identification of patterns by the examination of proximity of countries to each other. The MDS analysis we conducted resulted in the 2-dimensional map that can be seen in Figure 3. As is common in data reduction techniques, the two dimensions correspond to abstract coordinate axes which do not hold any distinctive meaning. Instead, it is the distance between objects which is of importance and should be interpreted.

This map reveals that to some extent policy documents were semantically similar among countries that are geographical close or belong to the same geographic cluster. Most northern European countries, for example, are relatively close to each other in the bottom right quadrant, southern European countries clustered around the vertical axis and the central European cluster is to the right of them. The western European countries France, Germany, Austria, with Luxembourg being the exception, are all positioned in the top left quadrant of the map. Asian and Anglosphere countries are not clearly grouped together as it is the case with other clusters but the grouping of countries in the bottom left quadrant all belong to the political association of commonwealth nations (*i.e.*, Great Britain, India, Australia, Singapore) and the top left quadrant includes North American countries. These results can be an indication that policy documents are influenced by countries with prior established collaborations and traditional cultural and political relationships.

4.1 Topic Modelling

To extract the topics from the documents, we apply Latent Dirichlet Allocation (LDA) topic modelling using Gibbs sampling [7]. LDA is a generative probabilistic model that treats each document as a mixture of latent topics and each topic as a distribution over words. To generate the topics, we split the documents into paragraphs and we remove the stopwords. We note that we initially attempted to extract the topics using the whole documents, but LDA extracted only very general topics with a low coherence score – a measure that indicates whether the words in the extracted topics are semantically related. Therefore, we decided to split the documents into paragraphs, a process that has been previously shown to outperform document-level analysis [13]. Additionally, we remove country names and subsequent adjectives (*e.g.*, Denmark, Danish) from the documents' text. To extract the topics, we use unigrams (*i.e.*, single terms), bigrams (*i.e.*, two consequent terms), and trigrams (*i.e.*, three consequent terms) and we keep the nouns, adjectives, and adverbs. Additionally, we apply lemmatisation, the process of converting a word to its lemma (*e.g.*, 'achieving' is converted to 'achieve'), to these words. To implement the topic model we use the *gensim*[4] Python library.

LDA takes as a parameter the number of topics, with the optimal number of topics dependent on the collection of documents. To learn the optimal number of topics, we calculate the coherence score for different number of topics ranging from 2 to 15. The coherence score measures the degree of semantic similarity between the top topic words of a single topic. The best coherence score was obtained for ten topics (0.572).

For each of the ten topics we obtain the top 20 words and their respective weight. Three of the authors sat together to determine the topics' titles as based on these words. Through a discussion process we reached consensus on all ten topic titles. Table 1 shows the topics that were extracted from the collection and provides a description of the topic.

We created a contingency table from document-topic matrix in which countries were represented in rows and topics probability distributions in columns. Each cell was representative of the proportion of a specific topic for a specific country on a scale from zero to 100. We performed Correspondence Analysis (CA) on this 25x10 table to create a low dimension representation that would allow the investigation of similarities between countries in regard to what topics dominated their policy documents. CA is a data dimensionality reduction technique typically used to visualise the association of strength between row and column entries in contingency tables [20].

[3]https://tfhub.dev/google/universal-sentence-encoder-large/3

[4]https://radimrehurek.com/gensim/

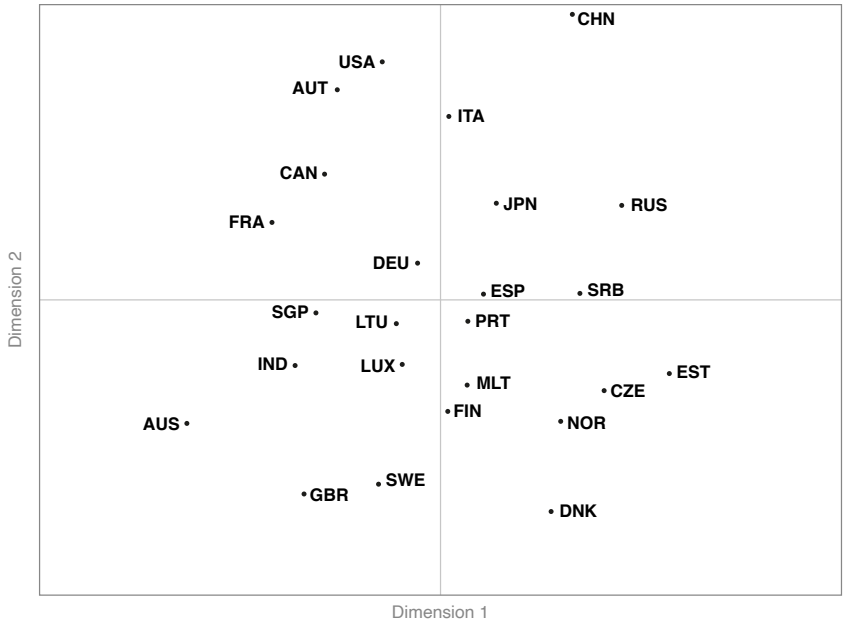


Figure 3: MDS based on semantic similarity of national policy documents.

Topic	Title	Description
1	Development strategy	Strategy for technological innovation, primarily from a government perspective.
2	Infrastructure	Technologies related to public infrastructure, such as energy and driving.
3	Private sector	Business projects, focus on development and investment opportunities.
4	Public sector	Administration and service, with a focus on citizens.
5	Data governance	Data access, data infrastructure, and datasets.
6	Ethical framework	Ethical consideration and related concepts such as regulation, privacy, and security.
7	Education	Training and obtaining of new skills, in particular digital skills (including programming).
8	Healthcare	Health services and the impact on both patients and economy.
9	Collaboration	Cooperation and alliances between public and private sector, as well as mentions of national, international and European.
10	Research	Research and development, both from an academic and industry perspective.

Table 1: Description of topics identified through topic modelling.

In the resulting CA plot, Figure 4, the geometric proximity between countries is an indication of similarity regarding how the ten topics where covered. The proximity to a specific topic label can be used to identify which topics were covered more extensively by individual or by a group of countries in their policy documents. The examination of this map shows that countries in the top left quadrant focused on development strategies while on the top right on collaborations. In the bottom quadrants, we can identify focus on research and the private sector while on the left corner documents did focus more on ethical consideration, healthcare, and the public sector.

However, while interpreting the results it has to be taken into consideration that countries positioned in the extreme edges of the plot were more distinct in the way they addressed the topics compared to those near the center of the axes. This means that countries close the origin of the axes were quite similar in how much they covered topics like infrastructure, education data, and data governance. Countries at the extreme were quite unique in the focus they gave to specific topics. For example, Japan in its policy document gave more relative focus on collaborations compared to the other countries. Similarly, Spain focused more on development strategies, Serbia and France on ethical frameworks, Australia on the Private sector, and Sweden on Research than the rest of the countries.

4.2 Discussion of Ethical Principles

The ethical aspects of artificial intelligence have become a widely discussed topic within both the academic and public domain. As

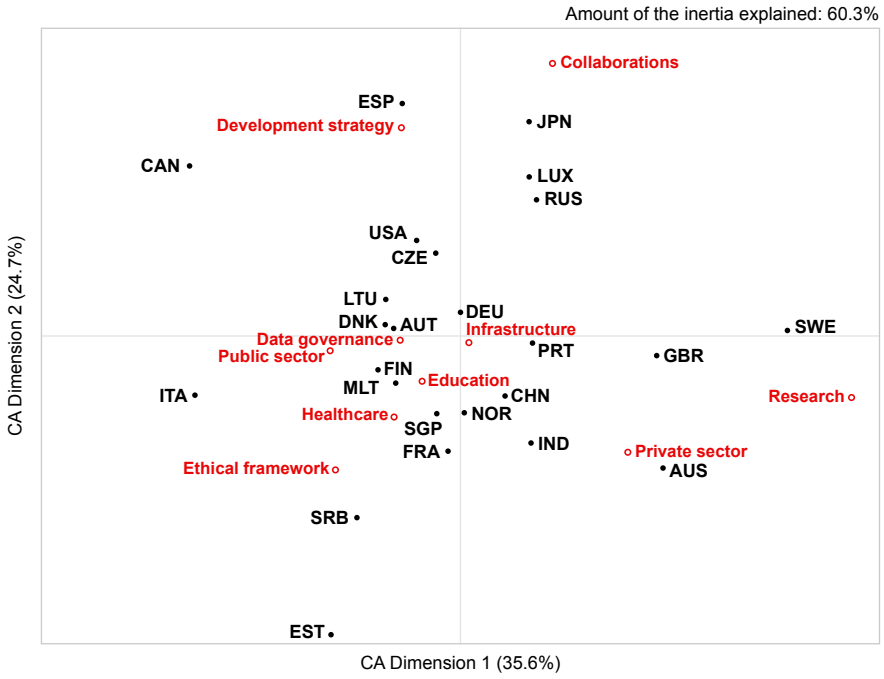


Figure 4: Correspondence analysis of the identified topics across 25 national policy documents.

ethics are such a critical aspect of future AI development, we analyse the frequency with which various ethical principles are discussed. We build on the prior work by Jobin et al. [25], and focus on the following eleven ethical principles; 'Transparency', 'Justice & Fairness', 'Non-maleficence', 'Responsibility', 'Privacy', 'Beneficence', 'Freedom & autonomy', 'Trust', 'Sustainability', 'Dignity', and 'Solidarity'. For each ethical principle, we follow the list of codes set by Jobin et al. [25] and calculate the frequency with which the codes occur in comparison to the total length of the document. For example, for the 'Freedom & autonomy' principle we analyse the text for the codes 'freedom', 'autonomy', 'consent', 'choice', 'self-determination', 'liberty', 'empowerment'. This analysis identifies not only whether an ethical principle is discussed in the text, but also to which extend the principle is discussed.

For all of the individual documents, we count the occurrence of terms related to the ethical principles as based on the pre-defined terms provided by Jobin et al. [25] and divide the total count by the number of total words in the document. Our results, shown in Figure 5, highlight a significant difference between the ethical principles and their occurrence in the policy documents (Kruskal-Wallis: $\chi^2(10) = 225$, $p < 0.001$). We run a post-hoc evaluations using a Dwass-Steel-Crichtlow-Fligner test (using p-value correction to account for type II error). We report the post-hoc results for the two most commonly observed principles. 'Justice & fairness' has a significantly higher occurrence compared to all other principles except for 'non-maleficence' and 'sustainability'. 'Non-maleficence' occurs more frequent than all other principles with the exception of 'Justice & fairness', 'Sustainability', 'Transparency', and 'Responsibility'.

Following the identification of the ethical principles across the corpus and prior work highlighting differences in cultural dimensions between countries [24], we compare the frequency with which these principles are discussed between existing geographical clusters of countries. We divide the corpus of 25 countries in the following clusters[5];

- **Anglosphere**, including Australia, Canada, the UK, and the USA.
- **Central and Eastern Europe**, including Chzech Republic, Russia, and Serbia.
- **Northern Europe**, including Denmark, Estonia, Finland, Lithuania, Norway, Sweden.
- **Western Europe**, including Austria, France, Germany, Luxembourg.
- **South/East Asia**, including China, India, Japan, and Singapore.
- **Southern Europe**, including Italy, Malta, Portugal, Spain.

For each cluster we calculate the average frequency with which each ethical principle is discussed within the overall corpus. Figure 6 visualises the differences between the six geographical clusters, highlighting both consistencies between clusters (*e.g.*, a limited focus on 'Solidarity' and 'Dignity') as well as differences between the geographical clusters (*e.g.*, a relatively large focus on sustainability in South/East Asia as compared to Central and Eastern Europe).

4.3 Qualitative assessment of ethical principles

Building on the frequency analysis of ethical principles presented in Section 4.2, we manually assess the ethical discussion in the

[5]European countries clustered according to the EU's EuroVoc [32].

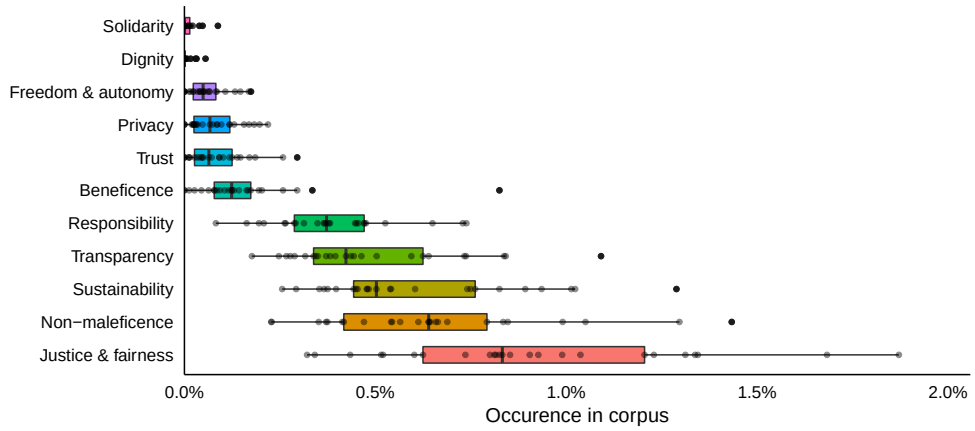


Figure 5: Occurrence of eleven ethical principles across the corpus of national AI policy documents.

corpus in order to obtain a better understanding of the different ways in which the ethical principles are interpreted. In particular, in this initial qualitative analysis we explore the topics of 'justice & fairness' and 'non-maleficence' due to their frequent occurrence within our corpus.

4.3.1 Justice & fairness. When discussing the justice & fairness implications of (future) AI technologies, countries within the Anglosphere cluster refer to justice & fairness within the context of technological innovation. For example, the USA's strategy document states that; "*Scientists must also study to what extent justice and fairness considerations can be designed into the system, and how to accomplish this within the bounds of current engineering techniques.*" [43]. In contrast, the Nordic documents often refer to pre-existing cultural conditions as a base for expanding their AI efforts, *e.g.*, "*Finland is known for its high level of citizen trust [...] It is a privilege to step into the age of artificial intelligence from such an exceptional setting. At the same time, it practically obliges us to an active approach, understanding of the prerequisites of trust in the age of artificial intelligence [...].*" [16]. These pre-existing conditions are even cited as a potential competitive advantage in the development of AI technology; "*Norwegian society is characterised by trust and respect for fundamental values such as human rights and privacy. This is something we perhaps take for granted in Norway, but leading the way in developing human-friendly and trustworthy artificial intelligence may prove a key advantage in today's global competition.*" [30]. In line with the frequency analysis (Figure 6), we find that countries in the South/East Asia cluster typically have a more narrow focus on the topic of justice & fairness and primarily address this principle by referring to the establishment of regulatory frameworks.

The introduction of additional regulation is discussed in all clusters. Countries from the Anglosphere and the South/East Asia cluster propose to expand existing national legislative power structures, *e.g.* the Canadian policy document recommends to "*Establish a Department of Digital Policy and an Office of the Chief Algorithmic Intelligence Auditor to manage government responses to AI.*" [10]. Countries in the European clusters were more likely to refer to the need need for European or international regulations in addition to their own national rule set; "*Lithuania needs to work for Lithuanian, European and international standards and regulations that promote the use of AI and prevent risks.*" [27] As these countries are either a member of the EU or uphold close relationships with the EU this is perhaps not surprising. Several policies highlight not only their favour towards EU-wide regulations, but simultaneously contrast such regulations to other countries; "*Europe and Denmark should not copy the US or China. Both countries are investing heavily in artificial intelligence, but with little regard for responsibility, ethical principles and privacy.*" [12].

4.3.2 Non-maleficence. The topic of non-maleficence deals with the potential for AI systems to cause harm, as well as ways in which societies can protect themselves from these negative outcomes of AI. All clusters point to the concerns around AI for various demographic groups, *e.g.* the USA; "*Many concerns have been voiced about the susceptibility of data-intensive AI algorithms to error and misuse, and the possible ramifications for gender, age, racial, or economic classes.*" [43] and Denmark "*Artificial intelligence should not reproduce prejudices that marginalise specific population groups. There will be active work to prevent unwanted bias and promote designs that avoid classification discriminating on ethnicity, sexuality and gender, for example.*" [12].

In their policies, Western European countries stress the potential harmful consequences for societies and professionals if AI systems 'dictate' the optimal outcome of a decision and limiting professional manoeuvring space; "*the increased use of these technical solutions will lead to an increased pressure to standardize the decisions made by institutions: it is far easier for a judge to follow the recommendations of an algorithm which presents a prisoner as a danger to society than to look at the details of the prisoner's record himself and ultimately decide to free him.*" [45].

Both the Southern and Northern European countries strongly refer to interdisciplinary education of AI concepts to establish a better understanding and thus protection against maleficent AI system among its citizens – as stated by the Norwegian policy document; "*Knowledge of artificial intelligence, and related fields such*

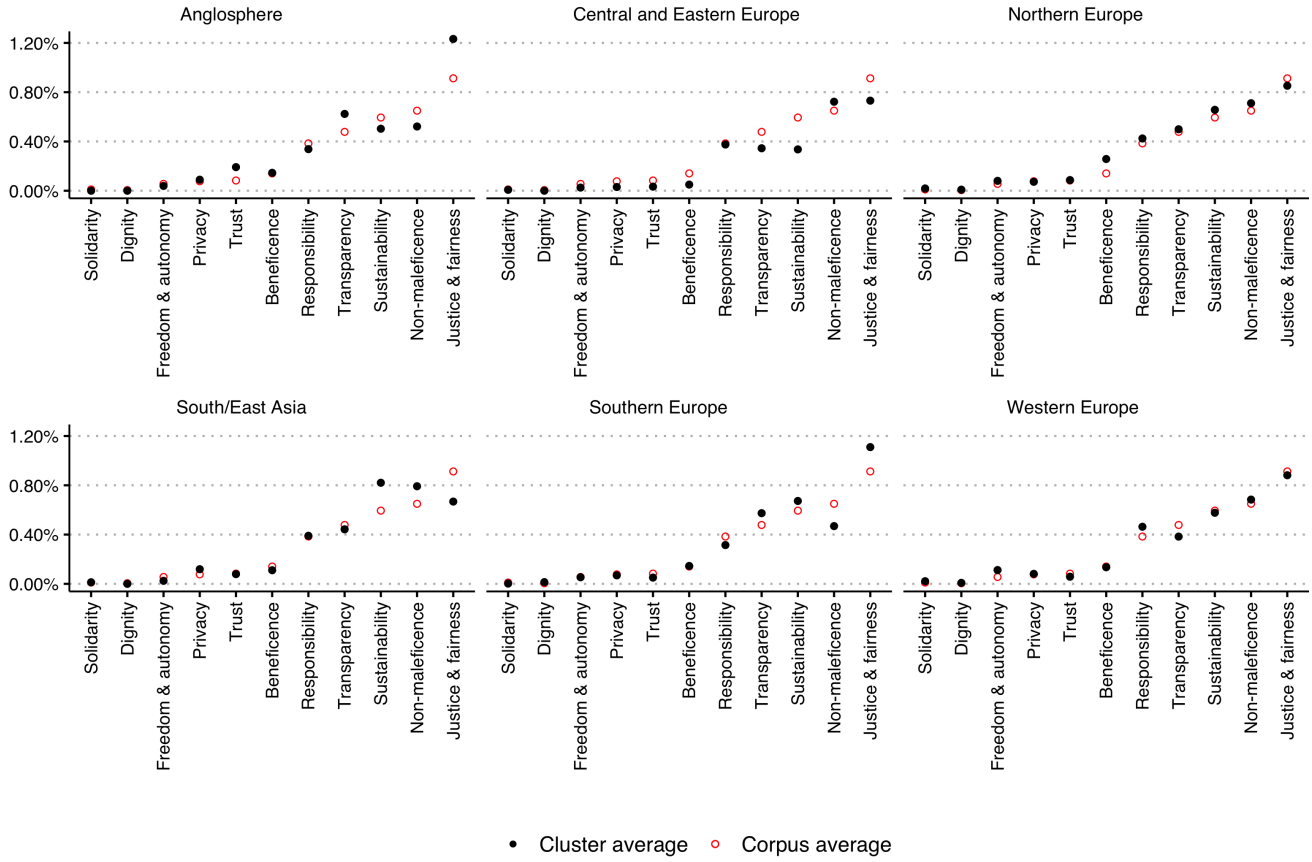


Figure 6: Average occurrence frequency of eleven ethical principles across six geographical clusters.

as ethics and data protection associated with AI applications, will be important in study programmes oriented towards the educational sector, health, crime prevention, law and several other fields." [30]. Discussions around AI education focus primarily on the embedding of AI knowledge in the existing curriculum and are focused especially on children and young adults, as seen in *e.g.* the Portuguese strategy document; "*young students should understand the risks and threats that they face in the same way as the rest of the community, aggravated with the fact that they spend most of their time immersed in the cyberspace, with the false feeling that being 'digital natives', so at ease with technology and devices, their – frail and superficial - expertise protects them from hazards and attacks.*" [18].

5 DISCUSSION

Through a systematic identification of national policy documents, following the screening, inclusion, eligibility, and exclusion criteria outlined in Figure 1, we identified a total of 25 AI-focused national policy and strategy documents. These documents, spanning a total of four continents, highlight the global interest among governments in AI and its impact on their citizens. Our findings suggest that significant differences exist between the national AI policy documents of the analysed countries. Through a semantic similarity assessment, assessing the similarity in meaning of the texts, and subsequent mapping of the respective distances through MDS (Figure 3), we find that geographical clusters (*e.g.*, Northern European) emerge. This is in line with previous work from *e.g.* Hofstede, whose analysis of cultural dimensions of countries often highlights geographical/cultural groupings [24]. For example, on the dimension of 'power distance', Hofstede reports "*Power Distance Index scores are listed for 76 countries; they tend to be higher for East European, Latin, Asian and African countries and lower for Germanic and English-speaking Western countries*" [24].

To better understand the various policy documents, we performed a topic modelling analysis on the corpus and identified ten distinct topics. This analysis allows us to explore two distinct aspects. First, Figure 4 provides an indication of the relative importance of each topic to the respective countries. Countries positioned near the origin of the axes can be considered as average in their coverage of the topics, whereas countries positioned on the periphery are distinct in regard to a specific topic (*e.g.*, Spain on 'Development strategy'). Second, Figure 4 highlights the topic similarity. This data visualisation allows us to make additional inferences on the topics discussed. 'Research', for instance, is more closely related to the private than the public sector – highlighting the perspectives of policymakers on the significant role of commercially-driven innovation in the AI domain. Identifying these differences and similarities between countries allows researchers, policymakers,

and AI developers to align their focus area with the respective country.

Our analysis also highlights a number of similarities between countries. During the data collection phase, we found that almost all policy documents stress the need for the respective country to become a leader in AI research – both for the purpose of economical gain and to allow the country to align the technology more closely to their respective ethical values. Jobin et al., in analysing ethical principles across their corpus of AI guidelines, performed a manual identification of the occurrence of different ethical principles [25]. As such, their analysis identifies a total of eleven ethical principles and a count on the number of documents in which these principles are discussed. We build on this framework of ethical principles to analyse not only the occurrence of these principles across the corpus, but to assess the frequency with which they are discussed. This highlights a number of interesting differences and overlaps. Our analysis identifies 'justice & fairness' and 'non-maleficence' as the two most frequently discussed ethical principles. In contrast, Jobin et al. identify sustainability as the second-least occurring theme [25]. This stark difference around the discussion of the sustainability dimension might be due to the fact that Jobin et al. focused on the analysis of 'AI ethic guideline' documents, whereas our work considered AI policies. Both our analysis and the analysis by Jobin et al. identify the ethical principles of justice & fairness and non-maleficence as dominant themes.

5.1 Localised Perspectives on Artificial Intelligence

Our analysis reveals critical differences between national AI policies in terms of their semantic similarity, topic prioritisation, and their discussion of ethical principles. Perhaps most critical for the development of future AI standards and applications is the observation that a number of countries form clusters with other countries that share cultural and geographical similarities. These clusters can form the basis of cooperation between countries in prioritising AI research as well as defining what is an acceptable use of AI. Although our analysis is restricted to national policy documents, the emerging cooperation on supra-national levels showcases how a number of governments have acknowledge the need for cooperation, see *e.g.* the discussion on Nordic-Baltic collaboration on AI [29].

In their *Nature* article 'Most people are not WEIRD' (Western, Educated, Industrialised, Rich, and Democratic), Henrich et al. argue that the majority of the research on human behaviour and psychology incorrectly assumes that different populations share the same cognitive and affective processes [23]. As such, sampling participants primarily from 'WEIRD' societies – in particular US undergraduates – limits the broader understanding of human behaviour. Henrich et al. suggest that this obstacle can be overcome by, *inter alia*, prioritising cross-cultural research and evaluating studies with judiciously chosen populations [23]. The same implications apply to Human-AI research, where an *e.g.* 'Silicon Valley-based' mindset to product development does not necessarily align with end-user needs across societies.

The respective topics and clusters identified in our work can be taken as a starting point to study Human-AI interaction cross-culturally. We argue that in order for AI to develop into a human-centred technology, it is key for researchers and developers to understand the values and principles of societies in which their technologies are used.

5.1.1 Capturing Local Values. The idea of geographical and/or cultural collaboration in the pursuit of a shared set of guidelines is an already widely established idea. These larger cooperation frameworks highlight how local values can lead to product differentiation within existing products and services. For example, the Nordic Swan Ecolabel – introduced by the Nordic Council in 1989 – has developed itself as an ecolabel for sustainability for sixty product groups[6]. These product groups cover not only physical products such as textiles or batteries, but also apply to services such as investment funds – prohibiting investments in, *inter alia*, companies that use fossil fuels to generate power. Lange et al. analysed the position of the Nordic Swan Ecolabel, as well as various other local ecolabels (*e.g.*, France, Germany), in relation to the EU Ecolabel [26]. Their analysis identifies a number of strategies for local ecolabel strategies; differentiation (uphold a strong national profile), assimilation (phase out national criteria when EU policy is introduced), or prioritising the focus area of local policies on (local) growth areas.

The examples provided in the ecolabel literature highlight both the possibility and opportunity for (clusters of) countries to differentiate themselves even under the umbrella of an overlapping organisation [26]. Translating cultural values, as *e.g.* quantified by Hofstede [24], to concrete recommendations for AI applications is, however, no straightforward endeavour. Initiatives by *e.g.* Google to reduce bias in image search show that a diverse set of training data is essential to the development of inclusive technology [5]. As such, we argue against 'global' AI applications which behave identical to all users and in accordance with the preferences and values of only a small and uniform subset of the global population. Instead, we promote the collection of localised perspectives on AI through methods such as localised crowdsourcing [4] or the targeting of (marginalised) sub-groups through interviews and focus groups [49]. Friedman et al., in a comparative study in the USA and Sweden, demonstrate how survey and interview data can be used to infer cultural differences in the perception of privacy [19]. These localised perspectives can subsequently be used to drive the development of future AI applications.

5.2 Implications for HCI Researchers

Our results highlight regional differences in national (governmental) perspectives on AI policy. The implications of these results are relevant to HCI research when designing AI applications and studying the use and non-use of such applications. Based on our results, we outline three implications for HCI research.

First, we need to develop methods for capturing values and principles. Our results highlighted differences between geographical clusters in the frequency with which ethical principles are discussed (Figure 5). Our initial qualitative analysis also indicates dissimilarities in the perception of these principles. Developing suitable methods for capturing and quantifying the values and principles on

[6] https://www.nordic-ecolabel.org/

both community and individual level is essential to better inform the decisions made by an AI system. As highlighted by Smith et al., AI systems which do not align with the values of the communities in which they operate will ultimately fail to solve the problems for which they were designed [38].

Second, we must be aware of regional and cultural differences between different user groups. In understanding the user needs and expectations of Human-AI interaction, HCI researchers must consider the impact of cultural differences between user groups. Prior work has shown that these differences can have a large effect on technology usage. For example, Oliveira et al. show that users from individualist countries are more likely to contribute answers and comments on question and answer websites in comparison to users from collectivist countries [33].

Third, we must analyse AI policies to inform about near future opportunities. Our findings indicate that policy documents can provide a useful and public resource for HCI researchers to identity topics within AI that warrant more work due to their regional relevance. Although methods such as design fiction have enabled HCI researchers to imagine the implications of future technologies, see *e.g.* [35], they largely rely on the considerations of the author(s) and are often focused on the distant future. By drawing on the issues identified in national policies, researchers can increase the likelihood that research outcomes are applicable in the near term and relevant to the surrounding society at large.

5.3 Limitations

As is evident from Figure 2, our analysis is largely limited to countries in the Global North. As observed on the linkhubs used in our policy identification search, many countries in the Global South have not (yet) published a national policy on AI. Similar to the analysis from Jobin et al. [25], our policy document discovery process did not yield any documents originating from Africa or South-America. We acknowledge that an analysis of documents from these regions, if such documents exist, likely would give rise to a distinct perspective. We leave this as future work as these documents do currently not exist. Our analysis reveals that even within the same continent (Europe) well defined clusters emerge (Figure 3). In addition to the omission of strategy documents from the Global South, we note that we encountered several countries which published a national AI policy in the country's native language (*e.g.*, the Netherlands, Poland), without an English translation available.

Our analysis of the national policy documents is mostly limited to a structured, quantitative approach. While this allows us to quantify the topics discussed in the work, it does not allow for a nuanced understanding of the differences in perspective when discussing these topics. Therefore, our analysis is based on the premise that topics which are more frequently discussed are also of greater importance to the respective country. Furthermore, our analysis does not consider the publication date of the analysed documents. Although all of the documents in the corpus are published between 2015 and early 2020, it is not unlikely that topic interests have shifted over these years. We aim to explore the corpus through a qualitative approach in future work and invite other researchers to build on our work. To this end, we publish the source code of our analysis and the list of analysed documents[7].

6 CONCLUSION

We presented the first empirical assessment of national AI policy documents, finding that a total of ten distinct topics among a corpus of 25 national AI policy documents. Our results highlight semantic and topical similarities between the AI policies of existing cultural and geographical clusters. Understanding and quantifying these differences in perspectives between countries is a first step towards tailoring AI applications to align with the values of end-users cross-culturally. Furthermore, we find a strong overlap between the analysed AI policy documents in terms of the frequency with which they discuss ethical principles. Future work can explore the analysed corpus qualitatively to obtain an understanding of the motivations of the different countries. We encourage further exploration of differences between the identified clusters through experimental studies in order to support the development of AI applications that are aligned to local cultural and ethical values.

REFERENCES

[1] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y. Lim, and Mohan Kankanhalli. 2018. Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article 582, 18 pages. https://doi.org/10.1145/3173574.3174156

[2] ACM U.S. Public Policy Council and ACM Europe Policy Committee. 2017. Statement on Algorithmic Transparency and Accountability. *Commun. ACM* (2017).

[3] Richard J. Adams, Palie Smart, and Anne Sigismund Huff. 2017. Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. *International Journal of Management Reviews* 19, 4 (2017), 432–454. https://doi.org/10.1111/ijmr.12102

[4] Andy Alorwu, Niels van Berkel, Jorge Goncalves, Jonas Oppenlaender, Miguel Bordallo López, Mahalakshmy Seetharaman, and Simo Hosio. 2020. Crowdsourcing sensitive data using public displays—opportunities, challenges, and considerations. *Personal and Ubiquitous Computing* (2020). https://doi.org/10.1007/s00779-020-01375-6

[5] James Atwood, Yoni Halpern, Pallavi Baljekar, Eric Breck, D. Sculley, Pavel Ostyakov, Sergey I. Nikolenko, Igor Ivanov, Roman Solovyev, Weimin Wang, and Miha Skalic. 2020. The Inclusive Images Competition. In *The NeurIPS '18 Competition*, Sergio Escalera and Ralf Herbrich (Eds.). Springer International Publishing, Cham, 155–186.

[6] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2018. The Moral Machine experiment. *Nature* 563, 7729 (2018), 59. https://doi.org/10.1038/s41586-018-0637-6

[7] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.

[8] E. Cambria, Y. Song, H. Wang, and N. Howard. 2014. Semantic Multidimensional Scaling for Open-Domain Sentiment Analysis. *IEEE Intelligent Systems* 29, 2 (2014), 44–51. https://doi.org/10.1109/MIS.2012.118

[9] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Brian Strope, and Ray Kurzweil. 2018. Universal Sentence Encoder for English. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP '18)*. 169–174. https://doi.org/10.18653/v1/D18-2029

[10] CIFAR and Brookfield Institute for Innovation + Entrepreneurship. 2019. Rebooting Regulation: Exploring the Future of AI Policy in Canada. https://brookfieldinstitute.ca/report/rebooting-regulation-exploring-the-future-of-ai-policy-in-canada/.

[11] European Commission. 2020. National Strategies - Knowledge for policy. https://ec.europa.eu/knowledge4policy/ai-watch/national-strategies_en.

[12] Denmarks Ministry of Finance and Ministry of Industry, Business and Financial Affairs. 2019. National Strategy for Artificial Intelligence. https://en.digst.dk/policy-and-strategy/denmark-s-national-strategy-for-artificial-intelligence/.

[7] https://github.com/agiahanou/Text-Analysis-of-National-Artificial-Intelligence-Policies.

[13] Lan Du, Wray Buntine, and Huidong Jin. 2010. A segmented topic model based on the two-parameter Poisson-Dirichlet process. *Machine Learning* 81, 1 (2010), 5–19. https://doi.org/10.1007/s10994-010-5197-4

[14] European Commission. 2018. Communication from the European Commission – Artificial Intelligence for Europe no 2018/137. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.

[15] European Commission. 2020. *On Artificial Intelligence - A European approach to excellence and trust.* Technical Report. 27 pages. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

[16] Finland's Ministry of Economic Affairs and Employment. 2019. Leading the way into the era of artificial intelligence : Final report of Finland's Artificial Intelligence Programme 2019. http://urn.fi/URN:ISBN:978-952-327-437-2.

[17] Organisation for Economic Co-operation and Development. Year unknown. AI initiatives worldwide. http://www.oecd.org/going-digital/ai/initiatives-worldwide/.

[18] Foundation for Science and Technology. 2019. AI Portugal 2030. https://www.incode2030.gov.pt/en/ai-portugal-2030.

[19] Batya Friedman, Kristina Hook, Brian Gill, Lina Eidmar, Catherine Sallmander Prien, and Rachel Severson. 2008. Personlig Integritet: A Comparative Study of Perceptions of Privacy in Public Places in Sweden and the United States. In *Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges (NordiCHI '08)*. 142–151. https://doi.org/10.1145/1463160.1463176

[20] Michael Greenacre. 2017. *Correspondence analysis in practice.* CRC press.

[21] Nina Grgić-Hlača, Elissa M. Redmiles, Krishna P. Gummadi, and Adrian Weller. 2018. Human Perceptions of Fairness in Algorithmic Decision Making: A Case Study of Criminal Risk Prediction. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 903–912. https://doi.org/10.1145/3178876.3186138

[22] Shathel Haddad, Joanna McGrenere, and Claudia Jacova. 2014. Interface Design for Older Adults with Varying Cultural Attitudes toward Uncertainty. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1913–1922. https://doi.org/10.1145/2556288.2557124

[23] Joseph Henrich, Steven J Heine, and Ara Norenzayan. 2010. Most people are not WEIRD. *Nature* 466, 7302 (2010), 29–29.

[24] Geert Hofstede. 2011. Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture* 2, 1 (2011), 8. https://doi.org/10.9707/2307-0919.1014

[25] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 9 (2019), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[26] Peter Lange, Ulrik Boe Kjeldsen, Maja Tofteng, Anja Krag, and Kasper Lindgaard. 2014. *The coexistence of two Ecolabels: The Nordic Ecolabel and the EU Ecolabel in the Nordic Countries.* Nordic Council of Ministers.

[27] Lithuanian Ministry of Economy and Innovation. 2019. Lithuanian Artificial Intelligence Strategy: A vision of the future. http://kurklt.lt/wp-content/uploads/2018/09/StrategyIndesignpdf.pdf.

[28] Quenby Mahood, Dwayne Van Eerd, and Emma Irvin. 2013. Searching for grey literature for systematic reviews: challenges and benefits. *Research Synthesis Methods* 5, 3 (2013), 221–234. https://doi.org/10.1002/jrsm.1106

[29] Nordic Council of Ministers for Digitalisation. 2018. AI in the Nordic-Baltic region. https://www.norden.org/en/declaration/ai-nordic-baltic-region.

[30] Norwegian Ministry of Local Government and Modernisation. 2019. National Strategy for Artificial Intelligence. https://www.regjeringen.no/en/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594.

[31] Future of Life Institute. 2020. National and International AI Strategies. https://futureoflife.org/national-international-ai-strategies/.

[32] Publications Office of the European Union. 2020. Concept scheme - 7206 Europe. https://op.europa.eu/s/n3ru.

[33] Nigini Oliveira, Nazareno Andrade, and Katharina Reinecke. 2016. Participation Differences in Q&A Sites Across Countries: Opportunities for Cultural Adaptation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. 10. https://doi.org/10.1145/2971485.2971520

[34] Organisation for Economic Co-operation and Development. 2019. Recommendation of the Council on Artificial Intelligence – OECD/Legal/0449. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[35] Daniel Pargman, Elina Eriksson, Rob Comber, Ben Kirman, and Oliver Bates. 2018. The Futures of Computing and Wisdom. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. Association for Computing Machinery, 960–963. https://doi.org/10.1145/3240167.3240265

[36] Joachim Schöpfel. 2010. Towards a Prague definition of grey literature. In *Twelfth International Conference on Grey Literature*. 11–26.

[37] Farid Shirazi, Adnan Seddighi, and Amna Iqbal. 2017. Cloud Computing Security and Privacy: An Empirical Study. In *Human-Computer Interaction. Interaction Contexts*, Masaaki Kurosu (Ed.). Springer International Publishing, 534–549.

[38] C. Estelle Smith, Bowen Yu, Anjali Srivastava, Aaron Halfaker, Loren Terveen, and Haiyi Zhu. 2020. Keeping Community in the Loop: Understanding Wikipedia Stakeholder Values for Machine Learning-Based Systems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. 1–14. https://doi.org/10.1145/3313831.3376783

[39] Jacopo Soldani. 2019. Grey Literature: A Safe Bridge Between Academy and Industry? *SIGSOFT Softw. Eng. Notes* 44, 3 (Nov. 2019), 11–12. https://doi.org/10.1145/3356773.3356776

[40] Jacopo Soldani, Damian Andrew Tamburri, and Willem-Jan Van Den Heuvel. 2018. The pains and gains of microservices: A Systematic grey literature review. *Journal of Systems and Software* 146 (2018), 215 – 232. https://doi.org/10.1016/j.jss.2018.09.082

[41] The United States Government. 2020. Artificial Intelligence for the American People. https://www.whitehouse.gov/ai/ai-american-values/.

[42] A. C. Tricco, E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. J. Peters, T. Horsley, L. Weeks, S. Hempel, E. A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M. G. Wilson, C. Garritty, S. Lewin, C. M. Godfrey, M. T. Macdonald, E. V. Langlois, K. Soares-Weiser, J. Moriarty, T. Clifford, Ö. Tunçalp, and S. E. Straus. 2018. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* 169, 7 (10 2018), 467–473.

[43] USA's Select Committee on Artificial Intelligence of the National Science & Technology Council. 2019. The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update. https://www.whitehouse.gov/ai/ai-american-innovation/.

[44] Niels van Berkel, Jorge Goncalves, Danula Hettiachchi, Senuri Wijenayake, Ryan M. Kelly, and Vassilis Kostakos. 2019. Crowdsourcing Perceptions of Fair Predictors for Machine Learning: A Recidivism Case Study. *Proceedings of the ACM on Human-Computer Interaction - CSCW* 3 (2019), 28:1–28:21. https://doi.org/10.1145/3359130

[45] Cédric Villani, Yann Bonnet, Bertrand Rondepierre, et al. 2018. For a Meaningful Artificial Intelligence: Towards a French and European Strategy.

[46] Sarah Theres Völkel, Christina Schneegass, Malin Eiband, and Daniel Buschek. 2020. What is 'Intelligent' in Intelligent User Interfaces? A Meta-Analysis of 25 Years of IUI. In *Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI '20)*. 477–487. https://doi.org/10.1145/3377325.3377500

[47] Hao-Chuan Wang, Susan F. Fussell, and Leslie D. Setlock. 2009. Cultural Difference and Adaptation of Communication Styles in Computer-Mediated Group Brainstorming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 669–678. https://doi.org/10.1145/1518701.1518806

[48] Ruotong Wang, F. Maxwell Harper, and Haiyi Zhu. 2020. Factors Influencing Perceived Fairness in Algorithmic Decision-Making: Algorithm Outcomes, Development Procedures, and Individual Differences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. 1–14. https://doi.org/10.1145/3313831.3376813

[49] Allison Woodruff, Sarah E. Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article 656, 14 pages. https://doi.org/10.1145/3173574.3174230

[50] Mike Woolridge, Peter Millican, and Paula Boddington. 2020. Ethics for Artificial Intelligence. https://www.cs.ox.ac.uk/efai/resources/alphabetical-list-of-resources/.

[51] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article 130, 11 pages. https://doi.org/10.1145/3173574.3173704

[52] Chen Zhao, Pamela Hinds, and Ge Gao. 2012. How and to Whom People Share: The Role of Culture in Self-Disclosure in Online Communities. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 67–76. https://doi.org/10.1145/2145204.2145219

[53] Chen Zhao and Gonglue Jiang. 2011. Cultural Differences on Visual Self-Presentation through Social Networking Site Profile Images. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1129–1132. https://doi.org/10.1145/1978942.1979110