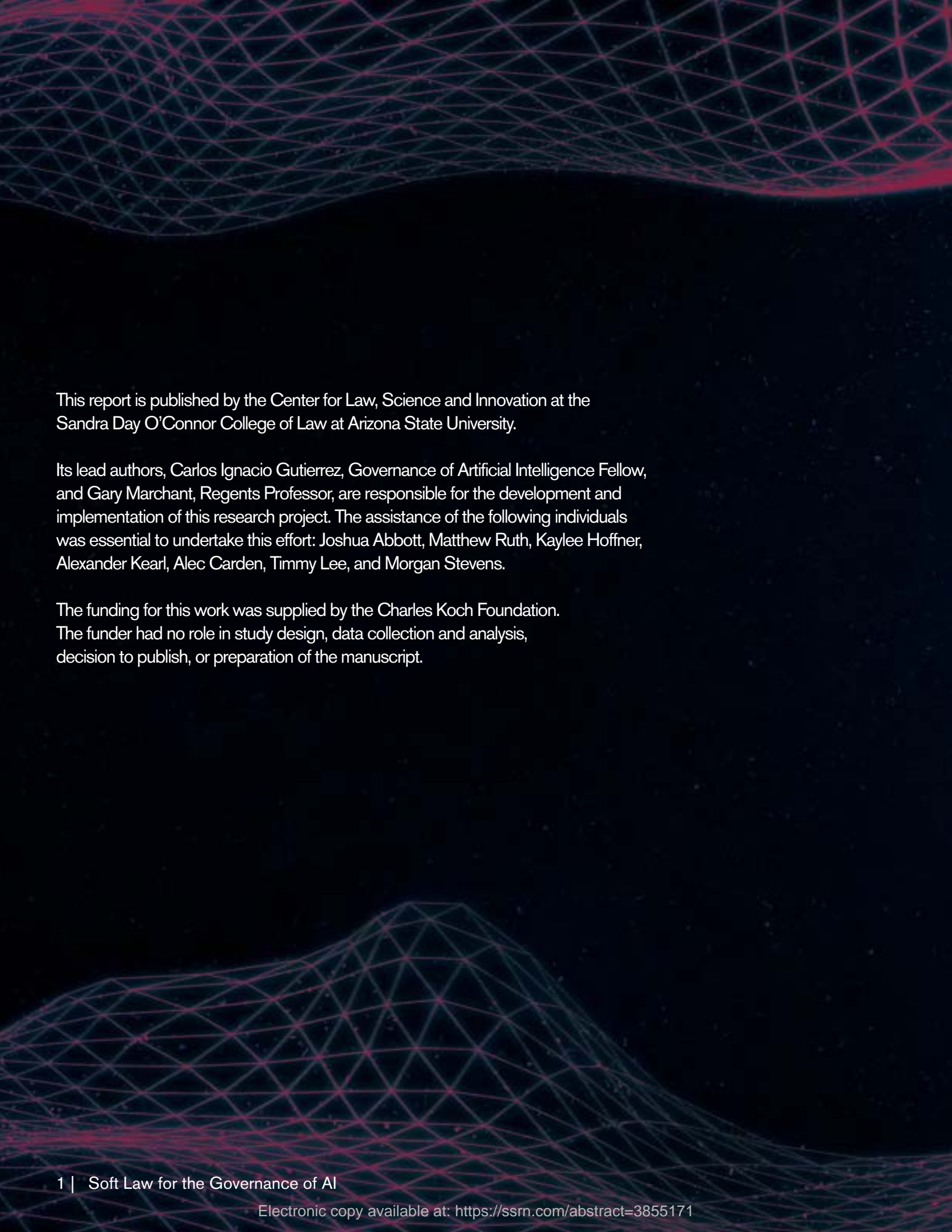


# A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence

**ASU** Sandra Day O'Connor  
College of Law  
Arizona State University

Carlos Ignacio Gutierrez, PhD  
Gary Marchant, PhD, JD

2021



This report is published by the Center for Law, Science and Innovation at the Sandra Day O'Connor College of Law at Arizona State University.

Its lead authors, Carlos Ignacio Gutierrez, Governance of Artificial Intelligence Fellow, and Gary Marchant, Regents Professor, are responsible for the development and implementation of this research project. The assistance of the following individuals was essential to undertake this effort: Joshua Abbott, Matthew Ruth, Kaylee Hoffner, Alexander Kearl, Alec Carden, Timmy Lee, and Morgan Stevens.

The funding for this work was supplied by the Charles Koch Foundation. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.



## Table of Contents

Executive summary.....	3
Introduction.....	4
1 Methodology.....	5
1.1 Identification.....	5
1.2 Screening and classification.....	6
1.3 Limitations.....	8
2 Results.....	9
2.1 Year of publication.....	9
2.2 Geography.....	10
2.3 Influence.....	11
2.4 Type of program.....	11
2.5 Role of stakeholders.....	13
2.6 Enforcement.....	16
2.7 Themes.....	23
Conclusion.....	35
References.....	36
Appendix 1 – Sources of information - linkhubs.....	49
Appendix 2 – Keyword search.....	52
Appendix 3 – Screening of soft law programs.....	53
Appendix 4 – Codebook for database.....	54



## Executive summary

Soft law is defined as a program that sets substantive expectations, but is not directly enforceable by government. Because soft law is not bound by a geographic jurisdiction and can be developed, amended, and adopted by any entity, it will be the dominant form of artificial intelligence (AI) governance for the foreseeable future. The objective of this document is to compile and analyze global trends on how this governance tool is used by government, non-profits, and the private sector to manage AI's methods and applications.

Inspired by similar efforts, this document contains a scoping review of AI soft law programs. Our process was divided into three steps: identification, screening, and classification. Our identification of programs began by establishing eligibility criteria. All programs had to: 1) conform to the definition of soft law, 2) emphasize the governance or management of a method or application of AI, and 3) were published by December 31st, 2019. These criteria made it possible to detect relevant programs through one of three methods that were implemented in a parallel manner. We found and mined over 80 linkhubs, resources that aggregate programs. We performed 370 keyword searches that combined our soft law program typology with a diverse list of themes, applications, and methods related to AI. Lastly, every screened-in program was vetted to search for references to other relevant programs, efforts such as these are denominated citation chaining.

In the screening process, we verified each program's compliance to the project's eligibility criteria. Out of the 1,599 programs initially identified, 965 were excluded because they were deemed to be articles or documents without a soft law component, unrelated to AI, or published after our 2019 cut-off. The final step of the process involved classifying the programs. Through several pilot exercises and by adopting best practices from relevant research, we developed 107 variables and themes to describe the programs. Variables provide information on how it is organized, functions, and its general characteristics, while themes communicate the subject matter discussed within a program's text.

Overall, we identified 634 soft law AI programs. Through our variables and themes, we were able to gather insights from this database. For one, the governance of this technology through soft law is a relatively new endeavor. Despite finding programs from the year 2001, over 90% of those in our sample were published between 2016-2019. Geographically, there appears to be limited diversity. The vast majority originate in countries classified as high income within Europe and North America. In the development of these tools, organizations appear to overwhelmingly prefer programs geared towards influencing the behavior of internal and external stakeholders, as opposed to those limited to internal stakeholders. At the same time, less than a third publicly mention enforcement or implementation tools meant to compel compliance with soft law program.

Despite having seven categories for classifying the type of soft law program, about 80% were labeled as principles or recommendations/strategy. This includes a list of 158 principles, one of the largest compilations dedicated to AI available in the literature. Our research also dispels the notion that soft law is the exclusive purview of industry self-regulation. Through the creation of governance triangles, we found that government entities led AI soft law development with a ~36% participation, followed by multi-stakeholder alliances with ~21%. One of the original contributions of this research is its classification of each program's text. By harnessing our 15 themes and 78 sub-themes, we uncovered that general mentions of transparency, general mentions of discrimination and bias, and AI literacy are the most represented issues in our database.

Soft law is not a panacea or silver bullet. By itself, it is unable to solve all of the governance issues experienced by society due to AI. Nevertheless, whether by choice or necessity, soft law is and will continue to play a central role in the governance of AI for some time. As such, it is important to build-upon the lessons that emanate from this research to make soft law as effective and credible as possible so it can address the governance challenges of AI systems, including safety, reliability, privacy, transparency, fairness, and accountability.

The ultimate goal of this research project is to inform decision-makers with evidence, practices, and recommendations that can be harnessed to enhance soft law programs with the objective of improving the management of applications and methods of AI under their responsibility.



## Introduction

Methods and applications of AI are pushing regulatory and social boundaries in every corner of the globe. To address these issues, the governance of AI can take two forms. There is hard law, represented by the creation or amendment of regulations. Generally, these processes can entail significant time delays and resources, which limits their responsiveness to emerging issues. On the other hand, there is soft law, which exists in the form of programs that set substantive expectations, but are not directly enforceable by government. Governance of this type can exist without jurisdictions and be developed, amended, and adopted by any entity. Throughout time, soft law has been treated as a preferred approach or delegated as a temporary alternative until hard law is promulgated. Regardless of its use, soft law's flexibility has made it the dominant form of AI governance.

This research project focuses on soft law for AI. Specifically, it compiles one of the most wide-ranging publicly-available databases on these programs ([the database can be accessed through this link](#)). It includes efforts from around the world, created by stakeholders of many varieties (e.g. private sector, non-profits, and government), and of different types (e.g. principles, standards, ethical codes, strategies, among others). This information is used to pinpoint global trends in how this technology is managed, where these programs originate, and whether they include enforcement mechanisms.

This document is divided into two sections. The first section concentrates on methodology. Here, readers will find our eligibility criteria for AI soft law programs, the process to identify and classify them, and a description of this project's limitations. The second section compiles our results. It contains a description of the most interesting findings divided into seven sub-sections that discuss everything from a program's organizational characteristics (e.g. year published or geographic jurisdiction) to a dissection of the issues discussed within its text.

Individuals and organizations interested in the governance of AI will benefit from the publication of this work in several ways. Those wanting to create AI soft law programs can examine how organizations have tackled issues of interest and learn from different approaches compiled in this project. Through this document and its database, researchers will enjoy access to a large sample of programs from which they can infer trends in the governance of this technology. Finally, this project can serve as an inspiration to regulators in the development or amendment of soft and hard law aimed at AI's methods and applications.

# 1 Methodology

This section details the steps taken by our team to identify, screen, and classify AI soft law programs. The methodology was heavily inspired by prior research aimed at characterizing the state of AI governance [1], [2]. It begins with a description of our multi-pronged strategy to identify a list of candidate programs. The second section details the screening of programs and our approach to classifying them with up to 107 variables and themes. The last section cautions readers on this project's limitations. Although precautions were taken to guarantee the database's reliability, we realize that areas of opportunity will be found within our methodology.

## 1.1 Identification

Inspired by and mirroring the work of Jobin et al.[2], this study performed a scoping review to identify the characteristics and trends of AI soft law programs at a global scale [3]–[5]. Scoping reviews are a particularly effective means of gathering information in areas where evidence on a subject is nascent or emerging [3]. One factor complicating this type of research is the heterogeneity in the creation and distribution of soft law programs. Any organization can create them without having to comply with paradigms as to how they are promulgated, published, or broadcast to the world. Therefore, there is no centralized repository or systematic means of classifying them. To overcome this challenge, we established a three-pronged eligibility criteria to identify AI programs dedicated to soft law:

1. **Comply with a definition of soft law:** programs in this database are ones who set forth substantive expectations that are not directly enforceable by government [6]. Governance tools that fall within this scope include: strategies, partnerships, codes of conduct, professional guidelines, among others [6]–[8]. This effort explicitly excludes initiatives or programs that are considered hard law, those where a public entity may force an individual or organization to perform an act via the threat of a penalty or punishment.
2. **Emphasis on the governance or management of AI:** Considerable debate exists on how to define AI [9], [10]. In determining the soft law programs to include in this database, the research team was instructed to add any program directed at an AI technology, regardless of how its author characterized or defined it. This includes soft law programs created to govern AI applications (e.g. facial recognition and autonomous vehicles) or methods (machine learning and neural networks).
3. **Temporal restriction:** Our research team limited the inclusion of programs to those in development or published by December 31st, 2019. Considering the constantly evolving AI governance space, our team wanted to focus on published initiatives.

Once the eligibility criteria were established, programs were detected through three parallel identification strategies: linkhubs, keyword searches, and citation chaining. The intention of this multi-pronged approach was to maximize the number of relevant AI soft law programs within this study.

### 1.1.1 Linkhubs

Linkhubs are resources that compile potentially relevant soft law programs. Our team began evaluating a small number of well-known sources related to the governance of AI (such as [1], [2], [11]–[13]). As our search progressed, the research team added more resources to the list. Eventually, 84 linkhubs were evaluated (see Appendix 1). They range from databases created by multilateral organizations, governments, non-profits, independent researchers that utilized crowdsourcing, popular media, and academic publications.



### 1.1.2 Keyword searches

Our team developed and implemented a keyword search protocol. It consisted of combing through the first ten pages of a private-browsing session of a search engine (Google) using three levels of terms [1]. Our intent in creating this segmentation was to broaden our reach into the grey literature in order to detect an assortment of programs. The first level of keywords depict types of soft law found in the literature [6]. The second level of keywords consists of themes, applications, and methods of AI generated through a brainstorm session performed by the research team. The last level indicates the technology of interest, AI. Overall, the team performed 370 searches (all of the keywords can be found in Appendix 2).

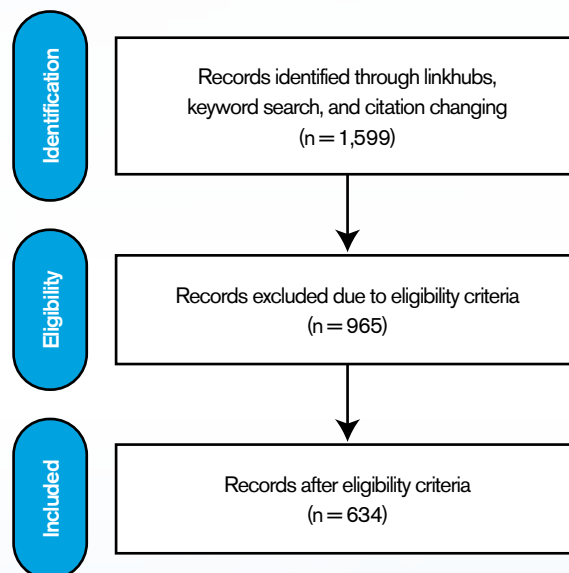
### 1.1.3 Citation chaining

The text and references within every screened-in program were vetted by at least one team member for mention of other programs. This was done in an effort to identify additional soft law programs that complied with the project's inclusion criteria.

## 1.2 Screening and classification

The output of the identification tasks led to the compilation of 1,599 candidate soft law programs. With this information in hand, the next step was to validate and classify them. The research team took a second look at each program to verify their compliance with the inclusion criteria. If they complied, the final step would be to extract information (see Figure 1 or a more detailed version of this figure is located in Appendix 3).

Figure 1 - Screening of Soft Law Programs



### 1.2.1 Screening

Once identified, candidate programs were subject to an additional layer of scrutiny. A team member, who was not originally charged with locating a program, verified its compliance to the project's inclusion criteria. If it did not meet the three requirements, it was excluded from the database. This led to the elimination of about 60% of programs.

Programs were eliminated from the database for one of six reasons (see Table 1). About 600 programs are articles in the popular or academic press that did not include a soft law element [14]–[16] or could not be placed under any soft law category [17], [18]. Another 121 programs passed the threshold of classifying as soft law, but were either not related to AI [19], [20] or published in 2020 [21], [22]. The final two categories are composed of documents classified

as recommendations, except that they did not create substantive expectations of action on any stakeholder. For instance, 171 documents were authored by non-profits that targeted recommendations to parties with whom there was no apparent relationship [23], [24]. Similarly, entities in the private sector published 80 documents selling their services (e.g. consulting firms) or directing recommendations to audiences of potential consumers and the public [25]–[27].

**Table 1 – Reasons for Excluding Programs**

<i>Justification for Exclusion</i>	# of programs
Article (popular press or academic) without a soft law component	367
Document without a soft law component	226
Non-profit providing unrequited recommendations	171
Private sector providing unrequited recommendations	80
Not related to artificial intelligence	68
Mechanism is potentially relevant, but was published in 2020	53
<b>Total</b>	<b>965</b>

### 1.2.2 Classification

One of the project's objectives is to detect trends in how organizations employ soft law in the management or governance of AI. To accomplish this, the core research team consulted similar efforts [1], [2] and brainstormed on what information could be extracted from these programs. This led to the differentiation between two types of data: variables and themes. Variables provide information on how a program is organized, functions, and its general characteristics. While themes communicate the subject matter discussed within its text. Overall, our project classified each program with up to 107 variables or themes (see Table 2).

**Table 2 – Classification of Programs**

Variables	14
Themes	15
Sub-themes	78
<b>Total</b>	<b>107</b>

The variables and themes in this analysis were developed via a pilot exercise with 35 randomly-selected screened-in programs. Each program was assigned to two individuals and their task was to examine the type of information that could be extracted. The pilot resulted in the creation of 14 variables and 15 themes.

Variables were generated through a consensus-driven process. They provide detailed information on how a program functions, its jurisdiction, and who participates in it. Themes were inspired by the work of Fjeld et al. [1]. In this sense, team members were asked to partake in a pile-sorting exercise [2]. Pile sorting is a process meant to elicit common attributes of information by bundling it into piles and discussing their shared characteristics until an agreement is reached. In this project, team members were asked to generate labels that described the text within a program. Subsequently, these labels were discussed and merged into groups until a consensus of a general taxonomy of



themes was reached. To operationalize this process, one individual was tasked with classifying the text of a program, while a second validated their work (see Appendix 4 for the inventory of variables and themes).

When the classification of the 634 programs concluded, a second pile-sorting exercise was performed to generate sub-themes. In this exercise, two individuals generated labels pertaining to a representative sample of 20% or at least 20 excerpts from each of the 15 themes. This information was used in a session where three team members met virtually to pile-sort the labels. The outcome was the creation of 78 sub-themes. Subsequently, one project team member was assigned to label a theme's text into its respective sub-themes.

### 1.3 Limitations

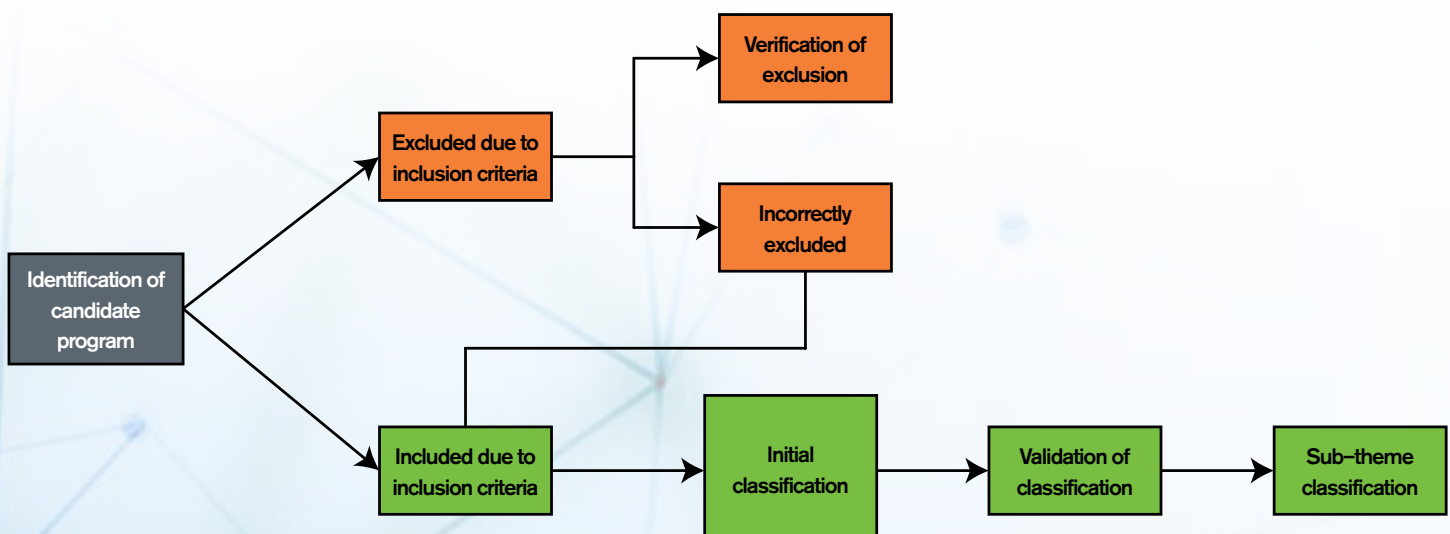
Our results have limitations and opportunities for improvement that need to be recognized. First, our search concentrated on media written in the English language. Thus, there is an explicit bias against programs that are not originally written in that language or have not been translated. Although we have undoubtedly missed important documents, our efforts to increase the diversity of programs included the use of free online translation tools (Google Translate) for non-English AI soft law programs found through our identification protocol.

Second, programs were identified, labeled, and reviewed by several individuals. Despite this, there is little doubt that important soft law programs or their characteristics could have been misunderstood, leading to mislabeling or erroneous exclusion from our analysis.

Third, the universe of soft law programs is rapidly expanding. As seen in this report, 42% of soft law examples in our sample were created in 2019. Despite the effects of Covid-19, we would not be surprised if 2020 represented an important year in the development of new soft law programs related to AI. Unfortunately, our research effort is a mere snapshot of AI Governance until 2019. In excluding programs created in 2020 and beyond, we have limited our perspective into the governance ecosystem of this technology.

In an effort to minimize the limitations mentioned above, this project validated each program's information several times. Figure 2 is an illustration of the steps taken to analyze our data, each box represents an individual pair of eyes. In the shortest case scenario, three individuals confirmed the exclusion of a program. Conversely, between five and six individuals evaluated the characteristics of every screened-in program. Although we know that this procedure is not infallible, our efforts were geared towards maximizing the reliability of information presented to our community of stakeholders.

Figure 2 - Procedure to Validate a Soft Law Program



## 2 Results

This project's methodology generated a database of 634 soft law programs coded with up to 107 variables or themes. The insights uncovered through the analysis of this data is divided into seven sections:

- 2.1. Year of publication:** despite finding a 20-year range of programs governing AI through soft law, its usage is a relatively recent phenomenon since over 90% of programs in our sample were published between 2016-2019.
- 2.2. Geography:** higher income countries dominate the generation of soft law, with a large concentration coming from the US and Europe.
- 2.3. Influence:** the vast majority of programs are created to influence the organization that created them and external stakeholders.
- 2.4. Type of program:** about 79% of programs are classified either as principles, recommendations, or strategies.
- 2.5. Stakeholders:** by organizing stakeholders into governance triangles we found that government entities led the highest number of programs (~36%).
- 2.6. Enforcement:** only 30% of programs publicly mention an enforcement or implementation mechanisms.
- 2.7. Themes:** we identified 15 themes and 78 sub-themes in the soft law governance of AI. The most prevalent theme was education – displacement of labor (identified in 815 excerpts) and the sub-theme was “general transparency” (present in ~43% of programs).

### 2.1 Year of publication

Publication dates for the programs in the database span 20-years between 2001 and 2019. The earliest example of AI soft law is a standard developed by a Chinese entity in 2001 titled “Artificial Intelligence--Basic Concepts and Expert System.” As shown in Table 3, this program did not immediately catalyze a barrage of similar efforts. Instead, the development of programs between 2001 and 2014 was limited to 20 specimens representing 3.5% of the database's total.

Table 3 - Year of Program Publication

Year	# of appearances	% of database
2019	269	42.43%
2018	192	30.28%
2017	91	14.35%
2016	26	4.10%
2015	22	3.47%
2014 and before	20	3.15%
Year not available	14	2.21%

Similar to trends uncovered in scholarly efforts devoted to analyzing the governance of AI, the creation of soft law for this technology is a recent phenomenon [2]. This is underscored by the finding that 91% of programs in our sample were published between 2016-2019. Interestingly, the publication of a program is a lagging indicator of the work performed by organizations. In other words, the impetus for managing AI methods and applications could be attributed to work that began several years prior to this timeline, such as between 2014 and 2017.



## 2.2 Geography

Programs in the database represent 64 unique geographic areas. In reviewing the World Bank's classification for the regions and wealth of countries in this list, a concentration is detected in terms of the type of nation(s) interested in dedicating resources to the creation of AI soft law programs [28]. Out of the 64 locations, 70% of these countries are either in Europe and Central Asia or East Asia and Pacific, and 80% are considered high-income (see Table 4).

**Table 4 - Location and Income of Countries in the Database**

Region	# of countries		# of countries
Europe and Central Asia	35	High Income	51
East Asia and Pacific	11	Upper Middle Income	6
Latin America and the Caribbean	5	Combination	4
Combination	4	Lower Middle Income	3
North America	3		
Middle East and North Africa	3	<b>Total</b>	<b>64</b>
Sub-Saharan Africa	2		
South Asia	1		
<b>Total</b>	<b>64</b>		

Note: Categories in this table are taken from the World Bank [28]

Further information was found in the two types of geographic relationships examined in the database (see Table 5). Country of origin describes the location where a soft law program was created, while jurisdiction represents its intended geographical area of application ("jurisdiction"). The U.S. is an important influence in this database as it is home to a large proportion of technology-based organizations that develop AI products and services, ~45% of which operate internationally. Compared to the US, countries such as Canada and China appear to have a mostly insular scope of geographic influence.

**Table 5 - Origin and Jurisdiction of Soft Law Programs**

Country of Origin	# of appearances	% of database	Jurisdiction	# of appearances	% of database
USA	163	25.71%	International	250	39.43%
International	105	16.56%	USA	88	13.88%
Europe	70	11.04%	Europe	68	10.73%
UK	52	8.20%	UK	35	5.52%
Germany	26	4.10%	Canada	17	2.68%
Canada	21	3.31%	China	16	2.52%
Japan	18	2.84%	Germany	14	2.21%
China	18	2.84%	Australia	12	1.89%
Netherlands	14	2.21%	Singapore	9	1.42%
France	13	2.05%	Netherlands	9	1.42%

## 2.3 Influence

In terms of organizational influence, programs are designed with one of two motivations. They are created solely to have an internal impact on the entity that adopted them or they are developed to affect both internal and external stakeholders. Based on our analysis, the vast majority (~82%) follow the latter pattern.

In these cases, parties develop their programs to inspire, encourage, and guide others through their work. This is the case of the AI principles developed by the OECD [29], all instruments developed by standards setting organizations, and government programs meant to encourage investment or participation in a particular sector [30], [31]. Conversely, ~18% of programs were conceived to be completely insular. This means that they are intended to be applicable only within the organization that created them. This is the case for Google's AI principles [32] and Deutsche Telekom's guidelines for implementing AI [33].

## 2.4 Type of program

The taxonomy of soft law programs was inspired by scholarship that examined its diversity [6]–[8]. Not all programs are created equally and, in this project, they were divided into seven categories (see Table 6). We find that the vast majority (~79%) are either principles or recommendations and strategies. In addition, we compiled one of the largest lists of principles related to the governance of AI applications and methods available in the contemporary literature.

**Table 6 - Type of Soft Law Mechanism**

	#	%
Recommendations and Strategies	344	54.26%
Principles	158	24.92%
Standards	60	9.46%
Professional Guidelines or Codes of Conduct	23	3.63%
Partnerships	21	3.31%
Certification or Voluntary Program	16	2.52%
Voluntary Moratorium or Ban	12	1.89%
<b>Total</b>	<b>634</b>	<b>100.00%</b>

### 2.4.1 Recommendations and strategies

This category contains two types of programs that comprise over half of the database (344 programs representing 54% of the sample). Strategies are roadmaps that highlight the direction an entity wishes to or should pursue. Meanwhile, recommendations were found in the form of suggestions, proposals, or evidence-based actions meant to improve an organization's status quo. Our team combined these programs because they tended to overlap.

The vast majority of recommendations and strategies are developed in-house by the entities that intend to adhere to them. However, the database incorporated a limited number of recommendations created by third-parties. To qualify, a substantive expectation of compliance needed to exist between organizations. For example, governments published AI specific strategies detailing plans to improve a country's AI competitiveness in research and development, transportation technologies, education, ethics, and other issue [34]–[36]. Many did so by forming multi-stakeholder alliances with non-profit institutions and/or the private sector (Gov-Np and Gov-Ps) [30], [37], [38]. Excluded from this category were programs generated by non-profits or private sector entities directed at institutions with whom they had no links to (pecuniary or non-pecuniary).



## 2.4.2 Principles

As broad statements that serve as high-level norms, principles have been the focus of important efforts aimed at implementing the soft law governance of AI [1], [2]. This database offers one of the largest known compilations of these programs containing 158 examples [39]–[41]. In contrast to the overall trend seen in Figure 1 (see section 2.5), most of the entities responsible for developing principles are in the private sector, followed by government with ~31% and ~28%, respectively. Interestingly, this research effort uncovered that, despite their high-level nature, a quarter of principles (38) incorporate or mention enforcement mechanisms (see section 2.6).

## 2.4.3 Standards

Any program developed by a standard-setting organization (SSOs) that addresses the technical needs of a field is within the scope of this category. As seen in Table 7, 12 organizations were responsible for the 60 standards in the database. As is customary with SSO's, virtually all programs are influenced by the multi-stakeholder input of governments, non-profits, and the private sector. In addition, their enforcement relies on employing an external entity whose job it is to verify compliance with the terms of a standard.

**Table 7 - Organizations Responsible for Publishing Standards**

<i>Institution</i>	<i># of standards</i>
International Organization for Standardization	27
Institute of Electrical and Electronics Engineers	16
European Telecommunications Standards Institute	4
German Institute for Standardization	4
World Forum for Harmonization of Vehicle Regulations	2
American Society for Testing and Materials International	1
British Standards Institute	1
International Electrotechnical Commission	1
International Telecommunication Union	1
Society of Automotive Engineers	1
Standards Council of Canada	1
Standards Press of China	1

## 2.4.4 Professional guidelines or codes of conduct

These programs describe behavior expectations applicable to individuals that work with AI applications or methods. They represent less than ~4% of the instruments in the database and are developed by a range of organizations including: professional associations who define the base level behavior expected from their members [42]–[44], industry associations who agglomerate private sector firms [45], and individual firms [46], [47].

## 2.4.5 Partnerships

A partnership is an initiative in which two or more entities collaborate to advance an agenda. Corresponding to ~3% of the database, these alliances are opportunities to advocate for an issue or generate synergies between stakeholders. The database bears witness to various flavors of these programs. Governments and the private sector may join to tackle a framework for responsible AI [48], governments cooperate to study alternatives for the technology's governance [49], and the private sector can work with a non-profit to advocate for ethical data governance [50].

## 2.4.6 Certification or voluntary programs

This soft law category includes two types of programs that represent ~2.5% of the sample. A certification is akin to a market signal, in the form of a “seal of approval” or statement, that indicates compliance to a set of pre-defined characteristics. Certifications exist for a number of issues including a commitment to minimizing the “abuse of facial analysis technology” [51] and the accountability of robotic products [52]. Excluded from this category are any programs related to educational certifications.

A voluntary program is a government initiative that invites non-government entities (private sector and non-profits) to comply with a non-binding set of actions or guidelines. Few of these were identified in this research effort. Among them, the government of Finland developed a program challenging local businesses to consider the ethical ramifications of their AI-based products [53].

## 2.4.7 Voluntary moratorium or ban

Moratoriums and bans are characterized by a call of action to avoid or cease the usage of an AI application or method. Generally, they target technologies that cause harm or negatively affect individuals. Only 12 of these programs were found in the sample. Ten focus on autonomous weapon systems [54]–[56] and the other two target AI-powered toys [57] and deepfake images [58].

## 2.5 Role of stakeholders

Our analysis of stakeholders is based on the work of Abbott and Snidal, who created a heuristic device illustrating the distribution of stakeholders in a particular sector called the “governance triangle” [59], [60]. Through it, we observe three types of actors that led the creation or implementation of soft law programs (government, non-profits, and the private sector). The corners of the triangle contain unilateral relationships indicating that one organization leads a program. These are connected to bilateral relationships whose main attribute is the inclusion of two actors who join forces in a program’s leadership. Lastly, the center of the triangle contains a category that combines all three actors.

**Figure 3 – Governance Triangle for Entities that Led Soft Law AI Programs**

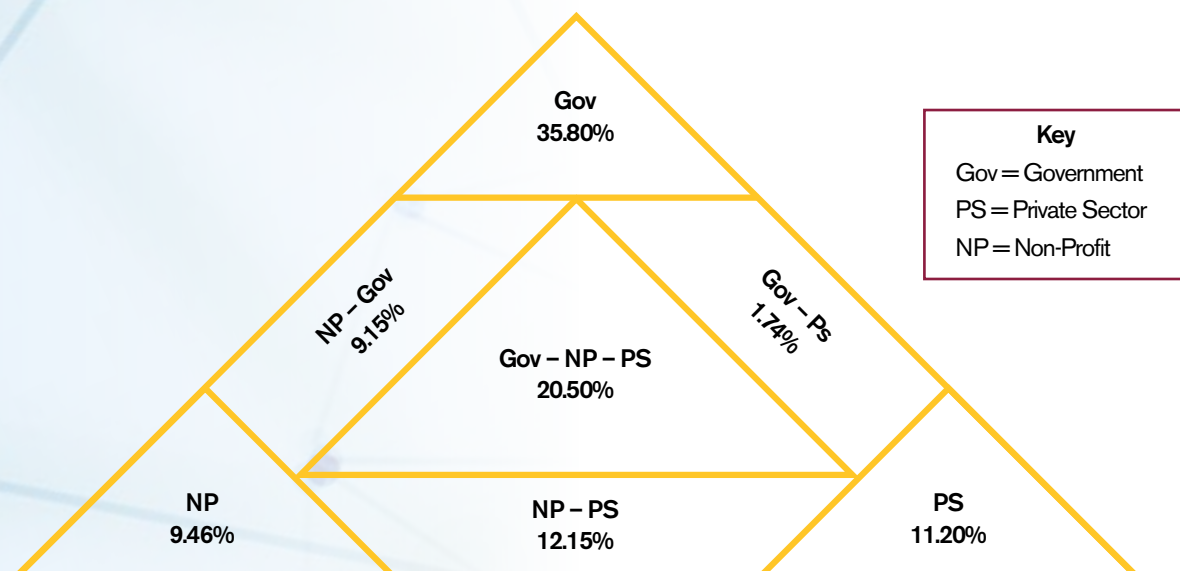


Figure 3 is a representation of the governance triangle for programs in this database. What stands out is the preponderance of government as the generator of AI soft law with ~36% of all programs being principally led by a public authority. This is followed by multi-stakeholder initiatives (~21%) and non-profits with private sector alliances (~12%). One of the interesting dynamics in the database are the differences in the distribution of stakeholders in the top three positions in the country of origin category: the U.S., international, and Europe (see Table 8 below).



**Table 8 – Stakeholders for Top Three Countries/Regions of Origin**

USA		International		Europe	
Gov	23.93%	Gov, Np, Ps	50.48%	Gov	51.43%
Ps	23.31%	Gov	24.76%	Gov, Np, Ps	17.14%
Np, Ps	21.47%	Np, Ps	10.48%	Gov, Np	12.86%
Np	12.27%	Np	7.62%	Np, Ps	11.43%
Gov, Np, Ps	9.20%	Gov, Ps	3.81%	Np	5.71%
Gov, Np	9.20%	Gov, Np	1.90%	Ps	1.43%
Gov, Ps	0.61%	Ps	0.95%	Gov, Ps	0%

In the U.S., stakeholders enacting AI soft law are distributed among many groups, with the private sector and non-profits sharing their participation in the leadership with government institutions. A hypothesis to explain this phenomenon is that many of the leading firms in the sector come from the U.S. and, because of their global influence, they have taken an active role in the development of soft law. A different perspective emphasizes the reluctance of government entities in this country from participating in the management of AI, hence its parity with other sectors.

Europe exhibits a picture unlike that of the U.S. Here, the government and its various alliances dominate the distribution of stakeholders with an 81% share, while the private sector by itself is marginally present with a 1.43% stake. Opposite to the U.S., European governments are known for adopting a proactive position in the governance of technologies, historically exhibiting a precautionary principle approach.

Furthermore, it can be argued that a smaller proportion of leading AI firms are based in Europe, which may explain the lower levels of penetration by this sector.

Finally, and far from surprising, the international front is led by multi-stakeholder efforts. Over half of programs are represented by all three sectors: government, the private sector and non-profits (e.g. standard-setting organizations, professional associations, among others). This is marginally followed by government, mainly in the form of multilateral initiatives.

### 2.5.1 Government

An entity representing one of the three branches (judicial, legislative, or executive) of a public authority at any level (local, state, or national) is denominated as government in this database. Despite a definition of soft law that excludes the direct enforcement of government power, this type of organization unilaterally generated the highest proportion of programs in the database (~36%) through the enactment of non-binding instruments. If alliances with other types of organizations are counted, government is present in ~67% of our sample.

To better understanding this dynamic, our team identified the level of government authority within all programs. As seen in Table 9, national and multilateral authorities are the most represented organizations in the leadership of soft law programs. In fact, over 78% of qualifying recommendations and strategies originate with government or one of its alliances, most of which were authored by national or multilateral authorities. This finding contradicts the popular narrative that soft law is simply industry self-regulation.

**Table 9 - Level of Government Involvement**

	#	%
National	241	38.01%
No Government Involvement	219	34.54%
Multilateral	89	14.04%
Multilevel	60	9.46%
State	17	2.68%
Local	8	1.26%
<b>Total</b>	<b>634</b>	<b>100.00%</b>

### 2.5.2 Non-profit

Non-profits are organizations that do not distribute earnings amongst parties or officially represent a jurisdiction. Programs led solely by these entities correspond to ~9% of the database, a third originate from the U.S., and ~88% are created with the purpose of influencing external parties. Over two thirds of non-profit programs are either recommendations and strategies, which come in varieties such as institutional action plans [61], multi-stakeholder alliances [62], and principles [63], [64]. Interestingly, virtually all moratoriums or bans were promulgated by non-profits. Ten are related to autonomous weapon systems [54]–[56] and one to AI-powered toys [57].

### 2.5.3 Private sector

Approximately 11% of the database is composed of programs created solely by private sector organizations, half of which originate in the U.S. Firms had a strong preference for producing principles (~69%) and restricting their influence towards internal processes (63%) [65]–[68]. Creating these programs provide organizations, especially large technology conglomerates with unique advantages. The creation of broad statements for the governance of AI is a relatively straightforward step that allows executive teams to broadcast their intentions with the technology to internal and external stakeholders. In addition, they can do so without necessarily having to develop and implement enforcement mechanisms that alter the organization's operation.

Alternatively, private sector firms can choose to outsource the communication of their position with respect to AI by harnessing a special type of relationship classified in this research effort as alliances between non-profits and the private sector. Representing ~12% of the database, a large proportion of programs led by Np-Ps can be described as industry groups with significant ties to the private sector [69], [70] or professional associations that agglomerate and represent the views of industry professionals [71], [72].

### 2.5.4 All sector alliance

Alliances between all sectors are the second most prevalent type of relationship, found in ~21% of the database. A large proportion (~45%) are standard setting organizations, who rely on representatives from all sectors of the economy to create technical specifications for a field. The second most popular alliance is spearheaded by governments who, in an effort to create inclusive AI strategies, invite non-profit and private sector representatives to co-author regional, national, and sectoral strategies [73]–[75].



2.5.5 Participants

Along with identifying the entities that represent a program’s leadership, our research group distinguished organizations charged with a secondary role. Denominated as participants, they are characterized by their contributions in opining, discussing, or participating in a program’s development. Figure 4 contains the distribution of stakeholders, which is limited to a quarter of the database (27%) since only that proportion of programs publicly indicated such roles. The largest group is represented by multi-stakeholder initiatives (~12%) such as government strategies [76].

Figure 4 - Governance Triangle for Entities that Participated in Soft Law AI Programs

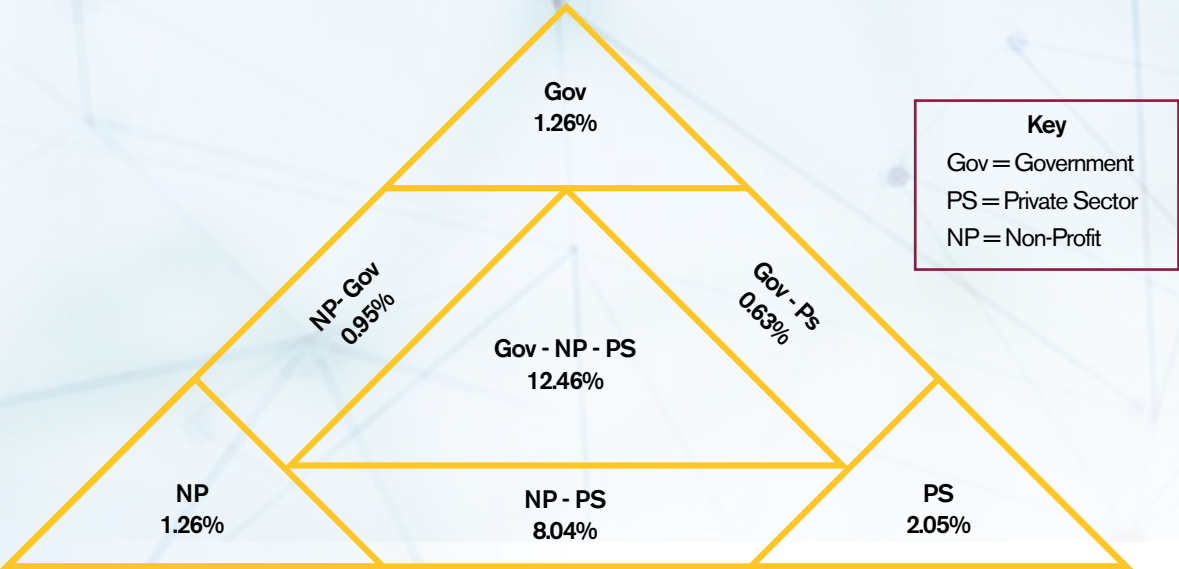


Figure 3 identifies stakeholders who had a secondary rather than primary role in developing AI soft law programs (when such information was publicly available).

2.6 Enforcement

In the pursuit of managing the consequences of AI, institutions throughout the world have created a range of soft law programs (e.g. principles, guidelines, strategies), many of which lack enforcement or implementation mechanisms. In fact, oft cited in discussions of soft law is its main weakness, its voluntary nature. In other words, all soft law programs rely on the alignment of incentives for implementation to take place, rather than the threat of direct enforcement. The literature on the subject offers insights into the menu of options available to facilitate this alignment [7], [77]. However, scant attention has been dedicated to documenting the trends in their overall existence or use in existing AI soft law programs.

This section addresses this research gap by identifying the organizational mechanisms, if present in a given soft law measure, that catalyze compliance, improve enforcement, or are used to plan for the implementation of soft law programs directed at AI. Further, it examines their characteristics and proposes a classification consisting of four categories: internal vs. external and levers vs. roles (see Table 10).

Table 10 - Characteristics of Enforcement Mechanisms in AI Soft Law Programs

	Internal	External
Levers	Training of employees Design of procedures Commitments Indicators Allocating a budget	Third-party verification Leverage
Roles	Champions Units Committees	Committees

Internal mechanisms are those whose operation depend on resources available within an organization. Naturally, external mechanisms are the opposite. They invite the participation of independent third parties to play a role in overseeing the implementation of a program. Within external and internal mechanisms, we identified levers. Levers represent the toolkit of actions or mechanisms that an organization can employ to implement or enforce a program. Its counterpart is characterized as roles. It describes how individuals, the most important resource of any organization, are arranged to execute the toolkit of levers.

This section is divided into four parts that analyze the mechanisms within the quadrants of Table 10. Readers will find that the voluntary nature of soft law need not, and often does not, represent an obstacle for a program's enforcement. Stakeholders within the database harness the diverse menu of levers and roles to transition their soft law program from an idea to actionable AI governance.

## **2.6.1 Internal levers**

The organizational toolkit available to implement or enforce soft law programs are denominated levers. One way to distinguish the five mechanisms listed in this quadrant is by their objective. The first three are designed to improve or guarantee the operation of a program through educational initiatives that inform or train employees, generate procedures to guide their actions, or allocate a budget to make a program's operation possible. The latter two are centered on goal-setting activities that motivate action such as making public commitments or creating indicators to measure the success of activities.

### *2.6.1.1 Educating the workforce*

Educational programs are a means to instill new norms or ideas to an organization's workforce. In the database, we found that these mechanisms were an opportunity to shape how employees interacted with a soft law program.

Specifically, some training initiatives were designed to introduce staff to an employer's approach for responsibly dealing with AI (explaining its principles or guidelines) [33], [78]. Others were directed at assisting employees in recognizing the intersection of AI and ethics, explain how to ethically develop products and applications using this technology, or build an understanding of the professional ethics expected in the field [33], [39], [78]–[80]. Mechanisms were also tailored to explain the negative effects of AI and instruct individuals on, for example, the actions they could take to “address potential human rights risks associated” with this technology [67].

Although most education-focused levers were generically described as trainings or workshops, one organization complemented these efforts by testing the creation of an “internal ethics certification” [81]. Its role within the firm was to generate an internal compliance tool that verified if staff with responsibilities directly associated to AI understood a firm's AI ethics guidelines.

### *2.6.1.2 Creation of procedures*

Once a soft law program is created, organizations may seek to enforce it by altering how their employees or representatives perform their duties through the creation and implementation of internal procedures. A wide scope of these type of levers was found in the database. For instance, firms created guidelines that allow employees to subjectively assess issues related to AI. This was the case in a firm who directed managers to evaluate the design of their AI products with an online questionnaire geared towards uncovering “potential human rights risks associated with the use of [AI]” [67].

Organizations also integrated soft law by developing procedures that employees are required to enforce (e.g. guidelines, review processes, daily meetings, testing procedures, among others) [39], [80], [82]. In one company, guidelines “must be followed by all officers and employees [...] when utilizing AI and/or conducting AI-related R&D” [66]. In another,



employees involved with the development and testing of AI applications were instructed to implement a review process, given the power to flag concerns and, if these were escalated, could ultimately halt the firm's work in testing an AI-based technology that was determined to be problematic [82].

#### *2.6.1.3 Allocating a budget*

Assigning resources to a soft law program is one of the strongest commitments an entity can make to guarantee its enforcement. The availability of a budget increases the likelihood that a set of tasks or actions will be implemented because the organization's management has deemed it enough of a priority to devote resources to it.

We found that such resources were allocated to two types of initiatives. First, governments assigned budgets for the implementation of their AI strategies. For one European government, this meant setting aside €20 million to “create a good basis for the development of automated driving” [83]. A North American counterpart bequeathed a fund of CA\$125 million to a research institution to “develop and lead” its AI strategy [84].

Secondly, we observed that resources were assigned to fund research initiatives and partnerships. These investments were made as one-off's, such as the apportionment for a multi-government research alliance [85], or on an annual basis, as was the case for the €2 million pledged by the Indian and French government for “basic and applied research projects, scholarships for training and research, exchange of experts and research projects, and awareness-raising measures” [86].

In terms of partnership, we uncovered programs that involve a variety of stakeholders contributing funding. For instance, the state of Karnataka in India partnered with a trade association to fund a Centre of Excellence for data science and AI, two German research institutes joined forces to create an AI center with the support of several levels of government, and Bank of America associated itself with Harvard University to support AI research [87]–[89].

#### *2.6.1.4 Commitments*

Publicly committing to a course of action is a signal to society that generates expectations about an organization's future actions. This section found a limited number of entities that make such promises to stakeholders. Although these commitments do not necessarily provide explicit detail on their enforcement, they can publicly bind the organization to act. In this regard, one firm asserted that their AI principles would “actively govern our research and product development and will impact our business decisions” [68]. Another stated that “all products and services are implemented and utilized by [...] employees, customers and partners” to “prevent and address human rights issues arising from AI utilization” [90].

General Electric signaled stakeholders about a variety of commitments related to their AI efforts [91]. The first commitment was to strive for diversity in their data science teams and to procure databases that reflected the population under analysis. The second commitment targeted the transparency of how their systems use data in their decision-making process. Finally, the last commitment is to work with stakeholders to ensure that the firm does “AI right – practically, methodically and for the benefit, safety and privacy of the patient” [91].

Signaling is not restricted to the private sector, as governments and multilateral organizations alike employ it. In one government strategy, we found that the agency implementing a soft law program committed to self-assess their goals based on advances in AI's state of the art [92]. Lastly, officials representing national governments in a multilateral pact agreed to annually review the “appropriate measures in order to adequately react to the emerging evolution of AI” [93].

### 2.6.1.5 Indicators

Indicators are akin to goal posts. Setting them allows an organization to evaluate its progress in achieving objectives and evaluate if taking further action is merited. Few examples of indicators as forcing mechanisms were found in the database. When identified, they were located within government strategies.

One government's AI strategy committed to create indicators in the near future, but in the meantime promised to track how stakeholders reacted to its strategy's activities [94]. A different government included AI as one of the technologies to be used in updating the administration of government. For this purpose, it developed indicators specific to the technology including: percentage of ministry families that use AI for service delivery or policy making with the goal of "all ministry families to have at least one AI project" [95]. Another government, whose strategy was completely devoted to AI, assigned performance indicators to each of its objectives [92]. It incentivized staff to achieve them by making their results a factor in determining future budget allocations [92].

## 2.6.2 Internal roles

The previously described levers were executed by individuals, groups of employees, or units within an organization. These individuals were assigned responsibilities with the purpose of steering, advising, or implementing soft law programs. It is self-evident that people are the most important institutional resource for the enforcement of soft law. Any program for that matter, including those directed at AI, will not function appropriately unless they are coordinated, championed, and pushed forward by individuals. This section examines the role of human resources in the execution of AI soft law programs. The analysis encountered three types of internal human resource roles: individual positions (e.g. champions), organizing groups of people through units, and the development of employee-led committees.

### 2.6.2.1 Champions

A champion is an individual who is bestowed the power within an organization to promote, educate, or assess issues related to an AI program. As expected, the power and responsibilities given to people in this role are wide-ranging. One government proposed the creation of an ombudsman to monitor the ethical practices of entities handling biometric technologies and how their associated data are utilized [96]. Another government created organizational archetypes (i.e. Chief Digital Strategy Officers and Chief Information Officers) to complement each other with the implementation of AI initiatives throughout the public sector [95].

In the private sector, AI champions have been assigned to assess how engineers understand ethics and transparency or identify if an AI-based product could produce a negative effect on the firm or society [67], [81]. Champions have also been given compliance responsibilities. One organization assigned two high level officers to evaluate high-risk products and determine if they satisfied an entity's AI privacy and decision-making standards [97].

### 2.6.2.2 Units

The creation of a unit within an organization represents a concerted effort to formalize the attainment of an objective through a permanent internal structure. We found one example of an organization with a unit dedicated to the enforcement of AI-related programs. Specifically, Microsoft's "Office of Responsible AI" is tasked with "setting the company-wide rules for enacting responsible AI," "defining roles and responsibilities for teams involved in this effort," and engaging with external efforts to shape soft law approaches to AI [40].

### 2.6.2.3 Committees

Internal committees or taskforces are entities whose composition includes representatives from within an organization (e.g. employees from executive management, engineering, human resources, legal, and product departments) [40], [98]–[100]. The role they perform varies by entity. In most cases, internal committees actively address the relationship between an organization's AI methods and applications with its consequences. This can be in the form of overseeing



AI-related processes, convening working groups, confirming that a firm's outcomes comply with AI commitments and soft law programs, or updating these programs when necessary [40], [78], [99], [101], [102]. In other models, committees are given a mandate to engage with relevant stakeholders from within and without the firm to generate and maintain appropriate feedback on the company's AI performance [100].

#### *2.6.2.4 Interaction between governance levels*

Few organizations publicly describe the relationship between different levels of their AI governance structure in detail. The ones that do, provide insights into the checks and balances in their implementation of soft law programs. In this database, the only organizations that provided this level of detail were Microsoft and Telefonica.

Microsoft built a structure in which its AI unit (Office of Responsible AI) and its internally-led committee (Aether) complement each other's role [40]. In their own words, these bodies "work closely with our responsible AI advocates and teams to uphold Microsoft responsible AI principles in their day-to-day work" [40].

Telefonica developed a multi-tiered model. In it, a product manager can signal that a method or application of AI may have a negative impact on the company to an AI champion [67]. These champions are tasked to solve the problem with a manager. If that is not possible, or if it represents a reputational risk for the company, "the matter is elevated to the Responsible Business Office which brings together all relevant department directors at [the] global level" [67].

### **2.6.3 External levers**

Within the organizational toolkits devoted to the enforcement of soft law, we found two levers that invited the participation of external entities: third-party verification and the use of leverage to compel the compliance of a target population.

#### *2.6.3.1 Third-party verification*

This mechanism entails the participation of an external entity in verifying the compliance to a set of guidelines. Standards, labels, and certifications are variations of programs that utilize third-party verification. In these cases, an independent entity is sought to reassure stakeholders on an entity's compliance with desirable guidelines. On the other hand, professional organizations depend on their members, individuals affiliated with them, but employed by other entities, to enforce their codes of conduct.

##### **2.6.3.1.1 Standards**

The purpose of standard setting organizations is to generate technical norms directed at the needs of the field's stakeholders. Entities interested in subjecting themselves to any of these standards have two choices on their level of commitment. In the first level, they can purchase the requirements of a standard and implement it by themselves. In choosing to do so, they are solely responsible for adhering to its rules. Which means that external parties have no means of verifying if the standard is adequately implemented.

Alternatively, entities can opt to hire a qualified third party to certify that the organization conforms to a standard's requirements. Doing so, represents a forcing mechanism that incentivizes the alignment to a soft law program. Successfully following a standard, grants an entity with an endorsement that can be communicated to stakeholders. This project found 60 standards related to AI that were in development or currently available. Examples of standards titles include: Guide for Verification of Autonomous Systems, Algorithmic Bias Considerations, and the Ethical Design and Use of Automated Decision Systems.

##### **2.6.3.1.2 Labels or certifications**

Similar to standards, labels or certifications are a signal to the market. However, these mechanisms validate a product or process according to a set of parameters developed by an institution with a distinct point of view. Organizations

voluntarily chose to subject themselves to these soft law programs to communicate the benefits of their products or services. The expectation is that this signal will generate a surplus of confidence in a target population (e.g. consumers, suppliers, government, among others). Many of these labels or certifications rely on third parties to verify an applicant's compliance to the requirements of a program. Successful applicants gain the ability to communicate that an independent entity has verified their claims, giving consumers more reasons to trust them.

The labels or certifications in our database are managed by non-profits, professional associations, and governments. In addition, they can be divided into two categories. The first category includes labels applicable to a wide scope of products and services. In one case, a foundation created a quality mark meant to implement the transparency and trust principles it believes should be present in all robots [52]. The organization relies on independent auditors to confirm that an applicant has adhered to its precepts on a yearly basis. Another is a test program by a professional association that offers a series of marks to certify all types of AI-based "products, systems, and services" [103]. Finally, a partnership between a non-profit and a government is creating a certification program meant to verify that AI products are "technically reliable and ethically acceptable" [104]. This instrument is an ambitious overarching tool that attempts to examine a product's "fairness, transparency, autonomy, control, data protection, safety, security and reliability" [104].

The second type of label or certification is specifically targeted at a particular application of AI. An example is the Safe Face Pledge mark meant for facial recognition technology. It requires participants to modify all internal procedures to comply with the values promoted by the creator of the label [51]. A different case is a collaboration between government, firms, and non-profits who are in the process of developing a certification for driving algorithms in Germany. This project aims to have firms upload their intellectual property into the system so it may analyze and "ensure their decisions are always favourable to the safety of the traffic around them" [105]. Lastly, a non-profit has created a seal tailored for the toy industry. It is designed to protect the rights of under-age individuals whose data and development is at risk from AI-enabled toys that could be used to exploit these vulnerable members of society [106]. The main target of this mark are parents who want their children to be safe from firms that could use the information they obtain from minors in a predatory manner.

#### 2.6.3.1.3 Professional associations

Professional associations agglomerate individuals from throughout the world to share experiences, debate major issues in their field, and set standards of conduct. In our search for soft law programs, individuals that form part of major global and regional associations related to AI require that their members report ethics code violations [107]–[109]. In some cases, not doing so constitutes itself a violation. In this sense, compliance to this particular brand of soft law depends on third-party cooperation for its success.

#### 2.6.3.2 Leverage

There are organizations who utilize their economic influence on others as leverage to compel adherence to soft law programs. In this analysis, we observed that two stakeholder groups were targeted: customers and suppliers.

One organization, a telecommunications firm, requested the compliance to its soft law program. Despite no mention of a binding mechanism, its documentation states that their AI ethical principles "must be observed by...business partners and suppliers" [110]. In contrast, firms can guarantee that their influence will affect the behavior of a target organization via written agreements. We found an organization that requires all of its customers and partners to adhere to its AI-related code of conduct [111]. In fact, if evidence of non-compliance is found, the firm reserves the right to terminate its business relationship. A similar commitment is required by a technology company with over 5,000 suppliers [97]. Each of them must agree to apply the firm's standards on cyber security and data processing.



## 2.6.4 External roles

The main conduit for inviting external human resources to participate in the enforcement of mechanisms is committees or taskforces. They represent an opportunity to include individuals who differ from their internal counterparts because they imbue a distinct perspective to the operation of an organization.

### 2.6.4.1 Committees

Organizations value the input of outsiders because they provide a viewpoint that is informed by different ideas and environments. Inviting these individuals to a board, committee, or taskforce offers a platform to contribute their knowledge and perspective. This section identifies how four types of organizations (private sector, government, non-profits, and professional associations) chose to harness these contributions to improve the implementation or enforcement of their AI-related soft law programs.

#### 2.6.4.1.1 Private sector

Boards or committees (whether they are legally constituted or act in a consultative manner) act in the interests of a public firm's shareholders by supervising the decision-making of its management. These bodies are composed of external members and are created specifically to deal with the risks and issues stemming from the integration or commercialization of AI.

They also differentiate themselves by their degrees of power. One firm's board has a role limited to counseling it on specific areas of concern such as "diversity and inclusion, algorithmic bias and data security and privacy" [112]. Another was motivated to assemble a body to "help guide and advise the company on ethical issues relating to its development and deployment" of AI-based products [113]. A significant difference between these two examples is that the latter was given the power to veto a firm's products and services. In fact, this board publicly declined to endorse the use of AI in a product, thus preventing its commercialization [114].

#### 2.6.4.1.2 Government

Similar to the private sector, membership in boards or advisory committees employed by public authorities include individuals from all segments of society (e.g. representatives from academia, industry, and non-governmental organizations). Our search discovered that governments seek boards with external members to oversee and implement their AI strategies. This can take the form of providing advice on delimited matters or assigning a multiplicity of tasks central to the implementation or enforcement of these soft law programs.

On one end of the spectrum, boards have a circumscribed role within the enforcement of AI strategies. They concentrate on tasks related to the "analysis and assessment of the ethical aspects of the use and implementation of AI", are charged with thinking about how to implement AI principles, "review [the] impact of technology on fundamental rights," or are asked to keep an eye on national and international AI trends [75], [92], [115]–[117].

On the other end, some countries assign these bodies with substantial responsibilities over the implementation or enforcement of their AI strategies. Finland's board has the remit of developing a diagnostic for improving the private sector's role in the AI ecosystem and drawing up an action plan for achieving it [118]. The Estonian government tasked its expert group with a similar directive as its Finnish counterpart, with the added responsibility of preparing draft laws related to AI and monitor the implementation of its strategy [119].

One of the most powerful boards is the one designated in Russia's AI strategy. Its responsibilities encompass the supervision and coordination of all efforts related to the implementation of its soft law AI program. This includes formulating the action plan, creating performance indicators, monitoring activities, and serving as the node between government and external stakeholders in all matters related to the government's AI efforts [120].

### 2.6.4.1.3 Research organizations and non-profits

A common theme for bodies with external members in research organizations and non-profits is a principal-agent relationship. In other words, these boards are constituted to ensure the accountability of the managing team in maintaining the entity's sustainability and reaching the AI-related goals of their soft law programs [41], [52], [61], [94], [121], [122]. In one case, an organization tasked its ethics committee with not only overseeing its ethical use of AI, but also with the creation of its soft law program [123].

### 2.6.4.1.4 Professional associations

The membership of professional associations is composed of individuals who practice in a designated field. The soft law programs within these organizations appear to be educational in scope. They recruit members, who are generally not employed by the association, and ask them to periodically inform other members on any AI-related developments in the field or assist in the creation of ad-hoc training programs related to this technology [124], [125].

## 2.7 Themes

Every program's text was classified into 15 themes and further subdivided into 78 sub-themes (see methodology section for details on how these divisions were created). Table 11 presents the top five results in both categories. It finds that education/displacement of labor is the theme with the highest number of excerpts in the database with 815. This means that text related to education/displacement of labor were found 815 times throughout the 634 soft law programs. Meanwhile, the sub-theme of general transparency appears in ~43% of programs. Readers of this section will find that each theme contains a description of the sub-theme, a table with the percentage of programs that contain each sub-theme, and representative excerpts. The database also contains the prevalence of sub-themes by type of soft law program.

**Table 11 - Top Five Themes and Sub-Themes**

Theme	# of labels	Sub-theme	% of database
1 Education – displacement of labor	815	1 General transparency	43.38%
2 Transparency and explainability	805	2 General mention of discrimination or bias	38.33%
3 Ethics	776	3 AI literacy	38.33%
4 Security	591	4 Acting in favor of AI ethics	29.34%
5 Bias	506	5 Human control and involvement in AI decision-making	27.92%

### 2.7.1 Accountability

Society is gradually bestowing AI-powered systems with autonomy to make decisions affecting individuals in lethal and non-lethal ways. In this theme, readers will find language highlighting the continuum of issues related to the bearing of responsibility for the unplanned actions and accidents caused by AI systems (see Table 12).

**Table 12 - Accountability Sub-Themes**

Sub-theme	%
Processes to ensure accountability	20.50%
General mentions of accountability	14.98%
What entity is ultimately responsible for the actions of AI: organizational responsibility	11.20%
What entity is ultimately responsible for the actions of AI: human responsibility	8.52%



About 15% of programs make general mention of accountability. They allude to the term, loosely define it, or state its importance to the program/society. A further ~21% recognizes the need for measures or mechanisms to ensure that accountability is considered. This is done by suggesting the creation of committees, the implementation of procedures, or anything in between. They range from general indications, such as what the Council of Europe's Commissioner for Human Rights describes: "member states must establish clear lines of responsibility for human rights violations that may arise at various phases of an AI system lifecycle" [126] or as specific as declared by the American Civil Liberties Union regarding accountability: "an entity must maintain a system which measures compliance with these principles including an audit trail memorializing the collection, use, and sharing of information in a facial recognition system" [127].

Some programs take a position as to who is primarily responsible for an AI system's actions. Around 11% single out organizations. They discuss the need to establish the type and extent of liability borne by firms or declare outright that legal persons should be the entities accountable for AI. This point of view is shared by the Association for Computing Machinery: "institutions should be held responsible for decisions made by the algorithms that they use, even if it is not feasible to explain in detail how the algorithms produce their results" [72].

With an opposing view, ~9% of programs affirm that humans, in the form of individual developers, operators, or decision-makers, are ultimately responsible for AI systems: "responsibility for these insights falls to humans, who must anticipate how rapidly changing AI models may perform incorrectly or be misused and protect against unethical outcomes, ideally before they occur" [128]. In between these positions, there is a 3% segment holding both parties accountable. They either differentiate the types of activities to which humans and non-humans are responsible for, assign responsibility to both, or are unsure as to which should bear the consequences:

- "Legal responsibility should be attributed to a person. The unintended nature of possible damages should not automatically exonerate manufacturers, programmers or operators from their liability and responsibility"[129]; and,
- "Institutions and decision makers that utilize AI technologies must be subject to accountability that goes beyond self-regulation"[51].

## 2.7.2 Artificial general intelligence

Defined as "highly autonomous systems that outperform humans at most economically valuable work", artificial general intelligence (AGI) is the next step in this technology's evolution [41]. Few programs spotlight AGI, which makes sense considering it is thought to be decades away from development. When discussed, ~1.4% of programs express traits desirable in such systems (see Table 13). The Chinese Academy of Sciences published a number of principles detailing the philosophy that should guide the creation of AI-based conscious beings, including: empathy, altruism, and have a sense of how to relate with current and future humans [130].

**Table 13 - AGI Sub-Themes**

Sub-theme	%
Development and governance of AGI	1.58%
Goals and traits of AGI	1.42%

About 1.6% of programs discuss how AGI should be developed and managed by decision-makers. This includes the research agenda to be prioritized (e.g. “autonomous decomposition of difficult tasks, as well as seeking and synthesizing solutions” [120]) or what governance mechanisms ought to be implemented (e.g. “urges the Commission to exclude from EU funding companies that are researching and developing artificial consciousness” [131]).

### 2.7.3 Bias

AI systems inevitably perpetuate the prejudices inherited in their design or emanating from the underlying data selected for their training. Over a third of the soft law in this database recognizes bias or discrimination in a general manner by stating the term or emphasizing the importance in avoiding its occurrence.

**Table 14 - Bias Sub-Themes**

Sub-theme	%
General mention of discrimination or bias	38.33%
Mechanisms for entities	15.62%
Diversity	15.46%
Inclusion	10.41%

In tackling this issue, programs take different approaches (see Table 14). In ~15% of cases, diversity is a term that represents the creation of a multidisciplinary workforce as a tool to combat the bias of AI systems: “we strive to use teams with people from diverse backgrounds to design solutions using artificial intelligence” [132] and “unless we build AI using diverse teams, data sets and design, we are at risk of repeating the inequality of previous revolutions” [133].

Meanwhile, there are programs that highlight the relevance of including populations that are generally excluded due to demographic or health characteristics (~10%): “AI should facilitate the diversity and inclusion of individuals with disabilities in the workplace” [134]. Lastly, ~16% of programs address bias by suggesting actionable mechanisms to decrease its impact: “a board should be created at EU level to monitor risks of discrimination, bias and exclusion in the use of AI systems by any organisation” [135].

### 2.7.4 Displacement of labor and education

The impetus for this theme was to unearth the relationship between the labor market and AI. Closely linked to it are the educational and research initiatives highlighting alternatives to ameliorate the overarching effects of this technology on population dynamics or to improve its contributions to society. Considering this, the text herein was distributed into three groups: labor, education, and research (see Table 15).

**Table 15 - Displacement of Labor and Education Sub-Themes**

Sub-theme	%
Education: AI literacy	38.33%
Research: research projects generally	22.24%
Research: link between society and research	16.25%
Labor markets changes: job loss and avoid job loss	15.93%
Education: skills / retraining	11.67%
Education: AI talent	9.78%
Labor markets changes: AI helping people get jobs or improve their current job	6.94%
Labor markets changes: solutions to job loss	5.52%
Education: AI aiding education	1.89%



The first group clusters the perceived consequences of AI on labor. It begins with ~16% programs that mention the possibility of job loss and the importance of avoiding it: “we have a responsibility to ensure that vulnerable workers in our supply chain are not facing significant negative impacts of AI and automation” [136] and “observe principles of fair employment and labor practices” [52]. A second group, ~6% of the sample, proposes a variety of alternatives to fight job loss, such as incentivizing communication-based activities: “all stakeholders should engage in an ongoing dialogue to determine the strategies needed to seize upon artificial intelligence’s vast socio-economic opportunities for all, while mitigating its potential negative impacts” [137]. The last label within this group, ~7% of the database, stresses the opposite of the first two, the labor efficiencies possible through AI such as: “simplifying processes and eliminating redundant work increases productivity” [138] and “accessible AI promotes growth and increased employment, and benefits society as a whole” [139].

Education is inextricably linked to preparing future generations for the demographic shifts caused by this technology. One of the most popular sub-themes in this database, appearing in ~38% of programs, remarks on the importance of providing the pedagogical and andragogical tools to facilitate AI literacy:

- “the IBM company will work to help students, workers and citizens acquire the skills and knowledge to engage safely, securely and effectively in a relationship with cognitive systems, and to perform the new kinds of work and jobs that will emerge in a cognitive economy” [78]; and,
- “update the education curriculum to refocus skills sets on AI under the umbrella of media and information literacy in preparation of the next generation of workers for AI adoption” [140].

Another section details the skills or retraining (~12%), not necessarily related to AI, that individuals whose livelihood is directly affected by this technological shift will face in order to continue earning a living: “low-skilled workers are more likely to suffer job losses...improving skills and competences is thus important to enable wider participation in the opportunities offered by new forms of work and for promoting an inclusive labour market” [141]. To complement both of these efforts, a small percentage of programs (~2%) referred to the ability of AI to aide in the provision of education: “conversational agents have huge potential to educate students...AI enhances our ability to understand the meaning of content at scale and serve it in meaningful and customized ways” [142].

While organizations await the influx of a new wave of AI literate workers, there are active efforts to recruit experts and specialists from around the world (~10%): “it is widely acknowledged that there is a skill gap in the agritech space and companies do not have time to wait for New Zealand to develop talent entirely on its own. Immigration policy should be continually monitored to allow rapid importing of the skills across the continuum to meet expected growing demand” [30].

The third grouping in this theme centers on research. All types of organizations (e.g. universities, firms, and governments) are incenting basic and applied AI research to improve their competitiveness. There are programs that describe research projects currently in progress or ideas that should be undertaken (~22%): “research Councils could support new studies investigating the consequences of deepfakes for the UK population, as well as fund research into new detection methods” [143]. The last sub-theme links research with society (~16%). Here, readers will find text on technology transfer opportunities, commercialization of AI discoveries, or partnerships with academia to bring research to the public: “DoD should advance the science and practice of VVT&E of AI systems, working in close partnership with industry and academia” [144].

## 2.75 Environment

The impact of AI on the environment is not covered extensively in the database (see Table 16). Relating the technology to its planetary impact through general statements occurred in ~9% of programs. One professional association asks its members to “promote environmental sustainability both locally and globally” [43].

**Table 16 - Environment Sub-Themes**

Sub-theme	%
Environmental protection generally	9.31%
Efficient use/conservation of resources	3.00%
Disaster management	0.47%

Specific mention of AI's aptitudes to improve the conservation of resources through efficiencies (3%) or in disaster management scenarios (~0.5%) was even more rarely discussed: “AI can highly improve the energy sector in Mauritius namely by...using IoT and neural algorithms to increase energy efficiency” [145] and “AI can be used in many aspects of preparation for and response to natural disasters and extreme events, such as hurricane winds and storm-related flooding” [146].

## 2.76 Ethics

This theme exhibits the moral compass or ideals that guide how organizations employ AI (see Table 17). At the surface level, the term ethics is mentioned without offering much detail as to its meaning (~19%). A similar phenomenon occurs with values (~25%) and culture (~5%) where, in many cases, they are used broadly: “enable a kind of a ‘passport of values’ whereby systems can learn one’s personal value preferences, an important part of prosocial behavior” [147] and “the development of AI technologies and their effects must always be in accordance with current legislation and respect local cultural and social norms” [148].

**Table 17 - Ethics Sub-Theme**

Sub-theme	%
Mechanisms to ensure AI ethics	29.34%
Values	25.39%
General mentions of AI ethics	18.77%
AI must be for a social benefit	18.61%
Protection of rights	17.19%
Ethical consequences	8.68%
Culture	4.57%

Many programs expressed hopeful thoughts or commitments about the need to ensure that the technology has a positive impact on society (~19%): “data and AI should enhance societies, strengthen communities, and ameliorate the lives of vulnerable groups” [61]. Further, rights, in particular human rights, were extolled as a vital requirement to be respected by the technology (~17%): “A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights” [149].

Conversely, there are programs that emphasize AI's negative ethical consequences (~9%): “calls on the Commission to propose a framework that penalises perception manipulation practices when personalised content or news feeds lead to negative feelings and distortion of the perception of reality that might lead to negative consequences” [131]. Almost a



third of programs (~29%) mention or suggest actions to ensure AI remains ethical. These include measures such as: “ban AI-enabled mass scale scoring of individuals as defined in our Ethics Guidelines” [150] and “establish a charter of ethics for Intelligent IT to minimize any potential abuse or misuse of advanced technology by presenting a clear ethical guide for developers and users alike” [151].

## 2.7.7 Health

The convergence of health technologies and AI promises to deliver significant value-added to the provision of medical services (see Table 18). About 6% of the database stresses the variety of benefits possible for this application of AI. These descriptions range from general statements such as “certain medical treatments or diagnoses might be carried out better with a robot” [152] to specific advantages: “carry out large-scale genome recognition, proteomics, metabolomics, and other research and development of new drugs based on AI, promote intelligent pharmaceutical regulation” [34]. Furthermore, some programs (~3%) spotlight the health and well-being of patients as a central node in the field: “a guiding principle for both humans and health technology is that, whatever the intervention or procedure, the patient’s well-being is the primary consideration” [153].

**Table 18 - Health Sub-Themes**

Sub-theme	%
Standards and oversight	6.62%
Benefits of AI in healthcare writ large (to hospital, people, and field)	5.99%
Patient Health - looking out for the health of people	3.31%
Collaboration of clinicians and developers	2.37%
Access	1.58%

To ensure the enduring nature of this technology’s advantages, programs stress its development, governance, and ability of individuals to access it. In terms of development, having manufacturers and clinicians work together can help ensure that AI is safely created and implemented effectively (~2%): “clinicians can and must be part of the change that will accompany the development and use of AI” [154]. Text that delves in the governance of healthcare AI attempts to verify that any device that assists in making life and death decisions does so in a manner that follows agreed upon practices or industrial standards (~7%). One standard created specifically for this purpose is aimed at helping manufacturers “through the key decisions and steps to be taken to perform a detailed risk management and usability engineering processes for medical electrical equipment or a medical electrical system, hereafter referred to as mee or mes, employing a degree of autonomy” [155]. Finally, if access to this technology is out of reach for large swaths of the population, its ability to positively contribute to society will be hampered. Statements discussing the need to make this technology available are represented in ~2% of the sample: “fair distribution of the benefits associated with robotics and affordability of homecare and healthcare robots in particular” [156].

## 2.7.8 Meaningful human control

AI systems are capable of decision-making at speeds that are beyond human capabilities. This theme discloses the desire to rein-in the technology through diverse means (see Table 19). At its most basic level, it remarks that humans need to be involved in the operation of AI systems (~28%), be it through governance (“we can make sure that robot actions are designed to obey the laws humans have made” [157]) or mechanically (“we are able to deactivate and stop AI systems at any time (kill switch)” [33]).

**Table 19 - Meaningful Human Control Sub-Themes**

Sub-theme	%
Human control and involvement in AI decision-making	27.92%
Development and feedback	15.93%
Right to redress or review	14.04%
Human autonomy	7.41%
User consent	7.41%
Right to refuse	3.94%

The meaningful-human control sub-themes also describe a continuum of human participation in AI decision-making. For instance, at any time, individuals should be given the ability to opt-out of these systems (~4%), “to establish a right to be let alone, that is to say a right to refuse to be subjected to profiling” [158], or be free to make their own decisions without being nudged in a particular direction (~7%), “algorithms and automated decision-making may raise concerns over loss of self-determination and human control” [159].

Prior to the engagement of these systems, about 16% of programs discuss the need to involve stakeholders (e.g. the public and affected entities) in their development: “no jurisdiction should adopt face recognition technology without going through open, transparent, democratic processes, with adequate opportunity for genuinely representative public input and objection” [95]. While ~7% stipulate that consent of any kind should be requested from users before participating in processes that involves an AI system: “advocate for general adoption of revised forms of consent...for appropriately safeguarded secondary use of data” [128].

Subsequent to being subjected to a decision enacted by this technology, a proportion of programs (~14%) advocate for the right of individuals to seek an explanation for decisions, have these overturned, or dispute them after the fact: “make available externally visible avenues of redress for adverse individual or societal effects of an algorithmic decision system, and designate an internal role for the person who is responsible for the timely remedy of such issues” [160].

## 2.7.9 Privacy

Freedom from surveillance or wholesale analysis of an individual's data exhaust is a timely subject (Table 20). This is especially the case in an era where AI applications can intrude into the public's life in ways that no humans were ever capable of doing in the past. About 22% of programs mention the word privacy in a general manner or stress the importance of its protection: “any system, including AI systems, must ensure people's private data is protected and kept confidential” [161].

**Table 20 - Privacy Sub-Theme**

Sub-theme	%
Individual privacy	22.08%
Privacy protection mechanisms	21.45%
Compliance with laws	17.51%
AI privacy systems	2.52%



In second place, at ~21%, programs mention systems or mechanisms that may protect user's information: "restricting third party access unless disclosed and necessary to the original purpose or application as stated in the Purpose Specification or in response to a legal order" [162]. To complement these mechanisms, ~18% of programs discuss their compliance with regulations whose purpose is primarily to ensure privacy: "while there is no single approach to privacy, IBM complies with the data privacy laws in all countries and territories in which we operate" [163]. Lastly, a small proportion of programs (~3%) discusses harnessing AI to improve privacy practices: "technologies for cyber security and privacy protection must be advanced" [164].

### 2.7.10 Private sector development

Private firms are the spearhead behind the research, development, and commercialization of most AI innovations (see Table 21). This sub-theme compiles the programs attempting to catalyze the development of the private sector, most of which (~79%) have government involvement. In fact, ~15% of programs describe, in general terms, the role of government in supporting the AI industry: "uphold open market competition to prevent monopolization of AI" [165]. Furthermore, we found programs that specifically backed efforts related to promoting the sector's competitiveness (~9%) and entrepreneurship via small and medium businesses (~11%):

- "Sweden's greatest opportunities for competitiveness within AI lies within a mutual interaction between innovative AI application in business and innovative organization of society" [166], and;
- "Assist SMEs to develop AI applications through AI Pilot projects, data platforms, test fields and regulatory co-creation processes" [167].

**Table 21 - Private Sector Development Sub-Themes**

Sub-theme	%
General mention of support for private sector development	15.46%
Entrepreneurship	10.57%
Competitiveness	8.83%
Private sector and non-government action	5.52%
Social responsibility	0.63%

Non-government parties can also act to improve the conditions and progress of the AI sector. In ~6% of programs, firms created mechanisms such as internal governance structures, performance indicators, or strategies that recognize the potential of AI. Meanwhile, ~1% of the sample discusses attempts by private and non-government entities to align themselves with corporate social responsibility goals (e.g. creating sustainable development goals).

### 2.7.11 Role of government/governance

This theme contains text with governance efforts related to the general management of AI (see Table 22). Without specifying a particular sector, a quarter of programs reference public entities as a key promoter or arbiter of AI for example: "avoid excessive legal constraints on artificial intelligence research" [168]. Organizations outside of government, mainly private sector and non-profits, also comment on their role in working and supervising the technology (~13%): "we need both governance and technical solutions for the responsible development and use of AI" [169].

**Table 22 - Role of Government/Governance Sub-Themes**

Sub-theme	%
Role of government as a promoter of AI	24.76%
Cooperation/engagement	21.45%
Non-government governance efforts (private sector and non-profits)	12.78%

Any text that highlights public-private partnerships, creation of alliances, or participation in multilateral fora related to AI systems was classified in the cooperation between parties to govern AI sub-theme (~21%): “we encourage states to promote the worldwide application of the eleven guiding principles as affirmed by the GGE and as attached to this declaration and to work on their further elaboration and expansion” [170].

### 2.7.12 Safety

One of the most important debates regarding AI systems relates to their ability to cause bodily harm and how to minimize it (see Table 23). Whether it is purposefully as an autonomous weapon or as an unplanned event in the form of an accident, this sub-theme delves into how programs contend with safety issues.

The first part of this sub-theme relates to the overarching safety of AI. In this sense, around 19% of programs include normative statements on the need for the technology to be safe and avoid or minimize physical harm to people: “there is a need for a public discussion about the safety society expects from automated cars” [171]. This is followed by a discussion on the mechanisms that ought to be implemented to ensure the technology’s safety (~15%), including instituting procedures or processes, as well as standards and regulations: “all the stakeholders including industry, government agencies and civil society should deliberate to evolve guidelines for safety features for the applications in various domains” [172].

**Table 23 - Safety Sub-Themes**

Sub-theme	%
General safety	18.93%
Safety mechanisms	15.30%
Restriction on AWS	5.99%
Military use of AI	5.84%
National security	3.31%
Law enforcement	3.31%

The second part of the sub-theme focuses on the weaponization of AI. Discussion of the military uses of the technology and the imposition of restrictions on autonomous weapon systems both appear in about 6% of programs:

- “Considering the increasing proliferation of autonomous systems, including among adversaries, the RNLA should continue to experiment with systems that may enhance its portfolio” [173]; and,
- “We deny that AI should be employed for safety and security applications in ways that seek to dehumanize, depersonalize, or harm our fellow human beings” [174].



The third, and last part, deliberates on AI as an information gathering technology at the national security (~3%) and the local level through law enforcement (~3%):

- “Understanding the need to protect privacy and national security, AI systems should be deployed in the most transparent manner possible” [175]; and,
- “Law enforcement needs for AI and robotics should be identified, structured, categorized and shared to facilitate development of future projects” [176].

### 2.7.13 Security/reliability

This theme is divided into two areas relative to protecting the integrity of AI systems (security) and ensuring their optimal operation (reliability) (see Table 24). On the security side, it encompasses risks to system integrity and the mechanisms to prevent adversarial attacks of the cyber variety (~26%): “manufacturers providing vehicles and other organisations supplying parts for testing will need to ensure that all prototype automated controllers and other vehicle systems have appropriate levels of security built into them to manage any risk of unauthorised access” [177]. Within the context of security, our team added a sub-theme that targets text discussing the protection of data from third-parties (~14%): “the development of AIS must preempt the risks of user data misuse and protect the integrity and confidentiality of personal data” [178]. The last part of security entails any text discussing working with AI to thwart cyber-attacks (~5%): “by using different algorithms to parse and analyze data, machine learning empowers AI to become capable of learning and detecting patterns that would help in identifying and preventing malicious acts within the cybersecurity space” [140].

**Table 24 - Security/Reliability Sub-Themes**

Sub-theme	%
Security: Cyber security	25.71%
Reliability: Reliability generally	17.82%
Reliability: Verifiability	15.14%
Security: Data protection	13.88%
Reliability: Data quality	8.52%
Reliability: Controlling for failure	7.41%
Security: Role of AI in cyber security	4.73%

The second section of this theme concerns reliability. About 18% of programs include normative statements on reliability, interoperability, or trustworthiness of AI systems: “utilize emerging frameworks that will help ensure AI technologies are safe and reliable” [179]. In case of a system outage, ~7% of programs highlight the need for procedures to offset the failure of the technology: “organizations should ensure that reliable contingencies are in place for when AI systems fail, or to provide services to those unable to access these systems” [180]. The last two sub-themes labeled text describing factors that affect data quality (~9%) and mechanisms to confirm the functionality of an AI system (~15%):

- “Users and data providers should pay attention to the quality of data used for learning or other methods of AI systems” [181]; and,
- “Solutions should be rigorously tested for vulnerabilities and must be verified safe and protected from security threats” [182].

### 2.7.14 Transparency and explainability

This theme focuses on conveying information to stakeholders on AI systems in a manner that is understandable and clear (see Table 25). Two sub-themes labeled data on the general use of transparency (~43%) and explainability (~24%) throughout programs, the former being the most popular sub-theme of the database.

A group of sub-themes deals with the information relationship between AI systems and individuals. For instance, ~6% programs suggest that individuals should be informed about any interaction with AI: “individuals should always be aware when they are interacting with an AI system rather than a human” [183]. Another sub-theme focuses on how individuals are subjected to consequential decisions by this technology, how they ought to know of them, and receive an explanation (~12%): “data subjects... have a right to obtain information on the reasoning underlying AI data processing operations applied to them” [184].

**Table 25 - Transparency and Explainability Sub-Themes**

Sub-theme	%
General transparency	43.38%
General explainability	24.29%
Sharing information/open access	15.93%
Open lines of communication with the public	13.88%
Data provenance	11.99%
Consequential decisions should be transparent and explainable	11.83%
Individuals should be informed about their interaction with an AI system	5.68%

To counter information asymmetry, one of the sub-themes highlights efforts to increase the awareness surrounding AI systems to the public or generally creating open lines of communication amongst stakeholders (~14%): “law enforcement should endeavor to completely engage in public dialogue regarding purpose-driven facial recognition use” [185]. A complementary label is applied to efforts looking to share AI-relevant databases amongst institutions (~16%): “develop shared public datasets and environments for AI training and testing” [186]. The last sub-theme in this section indicates where the data used by AI systems originated or how it is used in the training of systems (~12%): “identification of the type of biometric that is captured/stored and its relevance to the purpose for which it is being captured/store” [162].

### 2.7.15 Transportation/urban planning

This theme guides readers through programs interested on AI applications related to transportation (land and air) and their interaction to the urban environment. Starting with the user, programs discuss how an individual controls and communicates with AI transportation systems (2%): “when the vehicle is driven by vehicle systems that do not require the driver to perform the driving task, the driver can engage in activities other than driving” [187]. The theme scales up to one application of this technology, aircraft vehicles (~1%): “develop standards and guidelines for the safety, performance, and interoperability of fully autonomous flights” [188].



**Table 26 - Transportation/Urban Planning Sub-Themes**

Sub-theme	%
Governance of AI in transportation and urban areas	10.41%
Traffic safety and management	5.68%
Infrastructure	4.26%
Urban planning efficiency	3.00%
Role of the operator	2.37%
Autonomous aircraft	1.26%

The next set of sub-themes focus on the physical and non-physical support systems for AI-based transportation. Many programs discuss the infrastructure requirements needed for these applications to operate (~4%): “AI industry to work with telecommunications providers on specific needs for AI-supportive telecommunications infrastructure” [189]. Others center on the array of rules, guidelines, and regulations meant to govern their utilization (~10%): “this document establishes minimum functionality requirements that the driver can expect of the system, such as the detection of suitable parking spaces” [190]. Meanwhile, there are a number of proposals for managing traffic (~6%): “we can make mobility safer assisting human abilities and greener through platooning heavy goods vehicles to lower emissions and promoting public transport” [191]. Finally, there is an urban planning efficiency sub-theme dealing with sustainability efforts and resource management related to AI, but unrelated to traffic (3%): “AI-enabled solutions in the mobility and transportation sectors could go a long way in making cities more sustainable” [159].

## Conclusion

This research project identified 634 soft law programs directed at methods and applications of AI. In compiling this data, we learned a great deal about the state of AI soft law. For instance, this type of governance is a relatively recent phenomenon with over 90% of programs being created between 2017 and 2019. We dispelled the notion that these instruments are uniquely suited for private sector self-regulation since the largest group, approximately 36%, were generated by public sector entities. We found that most programs were generated in a cluster of high-income countries, dominated by the US, UK, Europe, or are global in nature. We confirmed that soft law's main characteristics, its voluntary nature, continues to be a leading disadvantage as 69% of programs do not publicly list enforcement or implementation mechanisms. Lastly, we created a library of over 6,000 excerpts that catalog the text of programs using 15 themes and 78 sub-themes.

Soft law is not a panacea or silver bullet. By itself, it is unable to solve all of the governance issues experienced by society due to AI. Nevertheless, whether by choice or necessity, soft law is and will continue to play a central role in the governance of AI for some time. As such, it is important to build-upon the lessons that emanate from this research to make soft law as effective and credible as possible so it can address the governance challenges of AI systems, including safety, reliability, privacy, transparency, fairness, and accountability.

The ultimate goal with this research project is to inform decision-makers with evidence, practices, and recommendations that can be harnessed to enhance soft law programs. Through this information, our hope is that all parties can improve how they manage applications and methods of AI under their responsibility.



## References

- [1] J. Fjeld, N. Achten, H. Hilligoss, A. Nagy, and M. Srikumar, "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," *SSRN Electron. J.*, Feb. 2020, doi: 10.2139/ssrn.3518482.
- [2] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nat. Mach. Intell.*, vol. 1, no. 9, pp. 389–399, 2019, doi: 10.1038/s42256-019-0088-2.
- [3] Z. Munn, M. D. J. Peters, C. Stern, C. Tufanaru, A. McArthur, and E. Aromataris, "Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach," *BMC Med. Res. Methodol.*, vol. 18, no. 1, p. 143, Nov. 2018, doi: 10.1186/s12874-018-0611-x.
- [4] H. Arksey and L. O'Malley, "Scoping studies: Towards a methodological framework," *Int. J. Soc. Res. Methodol. Theory Pract.*, vol. 8, no. 1, pp. 19–32, Feb. 2005, doi: 10.1080/1364557032000119616.
- [5] S. Anderson, P. Allen, S. Peckham, and N. Goodwin, "Asking the right questions: Scoping studies in the commissioning of research on the organisation and delivery of health services," *Health Research Policy and Systems*, vol. 6, no. 1. BioMed Central, p. 7, Jul. 09, 2008, doi: 10.1186/1478-4505-6-7.
- [6] G. Marchant, "'Soft Law' Governance Of Artificial Intelligence." *AI Pulse*, 2019, [Online]. Available: <https://ai-pulse.org/soft-law-governance-of-artificial-intelligence/>.
- [7] G. Marchant, L. Tournas, and C. I. Gutierrez, "Governing Emerging Technologies Through Soft Law: Lessons for Artificial Intelligence," *Jurimetrics*, vol. 61, no. 1, 2021.
- [8] G. E. Marchant and B. Allenby, "Soft law: New tools for governing emerging technologies," *Bull. At. Sci.*, vol. 73, no. 2, pp. 108–114, Mar. 2017, doi: 10.1080/00963402.2017.1288447.
- [9] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2014.
- [10] S. Legg and M. Hutter, "A Collection of Definitions of Intelligence," Jun. 2007, Accessed: Jan. 09, 2021. [Online]. Available: <http://arxiv.org/abs/0706.3639>.
- [11] OECD.AI, "The OECD Artificial Intelligence Policy Observatory," 2020, 2020. <https://oecd.ai/> (accessed Dec. 14, 2020).
- [12] Algorithm Watch, "AI Ethics Guidelines Global Inventory," *Algorithm Watch*, vol. 2019, pp. 1–9, 2019, Accessed: Dec. 14, 2020. [Online]. Available: <https://inventory.algorithmwatch.org/>.
- [13] AI Ethicist, "Ethical AI," 2021. <https://www.aiethicist.org/> (accessed Jan. 10, 2021).
- [14] T. Lysaght, H. Y. Lim, V. Xafis, and K. Y. Ngiam, "AI-Assisted Decision-making in Healthcare: The Application of an Ethics Framework for Big Data in Health and Research," *Asian Bioeth. Rev.*, vol. 11, no. 3, pp. 299–314, Sep. 2019, doi: 10.1007/s41649-019-00096-0.

- [15] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, "Key challenges for delivering clinical impact with artificial intelligence," *BMC Medicine*, vol. 17, no. 1. BioMed Central Ltd., p. 195, Oct. 29, 2019, doi: 10.1186/s12916-019-1426-2.
- [16] M. Perc, M. Ozer, and J. Hojnik, "Social and Juristic Challenges of Artificial Intelligence," *Palgrave Communications*, vol. 5, no. 1. Palgrave Macmillan Ltd., pp. 1–7, Dec. 01, 2019, doi: 10.1057/s41599-019-0278-x.
- [17] INDIGENOUS AI, "Indigenous Protocol and Artificial Intelligence Working Group." <http://www.indigenous-ai.net/> (accessed Jan. 11, 2021).
- [18] Finn AI, "3 Ethical Considerations When Implementing AI For Your Bank - Finn AI." <https://www.finn.ai/ethics-of-ai-in-banking/> (accessed Jan. 11, 2021).
- [19] Force 11, "Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0." <https://www.force11.org/fairprinciples> (accessed Jan. 11, 2021).
- [20] National Police Foundation, "National Police Foundation Publishes Best Practices Guide for Police Open Data." <https://www.policefoundation.org/national-police-foundation-publishes-best-practices-guide-for-police-open-data/> (accessed Jan. 11, 2021).
- [21] Philips, "Five guiding principles for responsible use of AI in healthcare and healthy living - Blog." <https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/2020/20200121-five-guiding-principles-for-responsible-use-of-ai-in-healthcare-and-healthy-living.html> (accessed Jan. 11, 2021).
- [22] Government of Norway, "National Strategy for Artificial Intelligence," 2020. Accessed: Jan. 11, 2021. [Online]. Available: [www.regjeringen.no/ki-strategi](http://www.regjeringen.no/ki-strategi).
- [23] Sipri, "The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk," 2019. Accessed: Jan. 11, 2021. [Online]. Available: [www.sipri.org](http://www.sipri.org).
- [24] Women Leading in AI, "10 Principles of Responsible AI." <http://womenleadinginai.org/wp-content/uploads/2019/02/WLiAI-Report-2019.pdf> (accessed Jan. 11, 2021).
- [25] M. Gelles, A. Mirkow, and J. Mariani, "The future of law enforcement Policing strategies to meet the challenges of evolving technology and a changing world."
- [26] Booz Allen Hamilton, "4 Ways to Preserve Privacy in Artificial Intelligence." <https://www.boozallen.com/s/solution/four-ways-to-preserve-privacy-in-ai.html> (accessed Jan. 11, 2021).
- [27] Ernst & Young, "How AI can be a force for good in government | EY - Global." [https://www.ey.com/en\\_gl/consulting/how-ai-can-be-a-force-for-good-in-government](https://www.ey.com/en_gl/consulting/how-ai-can-be-a-force-for-good-in-government) (accessed Jan. 11, 2021).
- [28] World Bank, "World Bank Country and Lending Groups," The World Bank. 2020, Accessed: Dec. 28, 2020. [Online]. Available: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>.



- [29] OECD, "Recommendation of the Council on Artificial Intelligence." 2019, doi: 10.1017/ilm.2020.5.
- [30] AI Forum - New Zealand, "Artificial Intelligence for Agriculture in New Zealand," 2019. <https://aiforum.org.nz/reports/ai-for-agriculture-in-new-zealand/> (accessed Jan. 29, 2021).
- [31] Government of Luxembourg, "Artificial Intelligence: opportunities, risks and recommendations for the financial sector," 2018.
- [32] S. Pichai, "AI at Google: our principles," 2018-07-07, p. 1, 2018, Accessed: Dec. 07, 2020. [Online]. Available: <https://www.blog.google/technology/ai/ai-principles/>.
- [33] Deutsche Telekom, "Digital Ethics Guidelines on AI," 2018. [Online]. Available: <https://www.telekom.com/resource/blob/544510/ca70d6697d35ba60fbc29aeef4529e8/dl-181008-digitale-ethik-en-data.pdf>.
- [34] Graham Webster, Rogier Creemers, Paul Triolo, and Elsa Kania, "Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)," 2017. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/> (accessed Jan. 29, 2021).
- [35] Government of Japan, "New Robot Strategy: Japan's Robot Strategy," 2015. Accessed: Jan. 29, 2021. [Online]. Available: [https://www.meti.go.jp/english/press/2015/pdf/0123\\_01b.pdf](https://www.meti.go.jp/english/press/2015/pdf/0123_01b.pdf).
- [36] European Commission, "The Age of Artificial Intelligence: Towards a European Strategy for Human-Centric Machines | JRC Science Hub Communities," 2018. Accessed: Jan. 29, 2021. [Online]. Available: <https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/age-artificial-intelligence-towards-european-strategy-human-centric>.
- [37] International Association of Privacy Professionals, "White Paper – Building Ethics into Privacy Frameworks for Big Data and AI," 2018. <https://iapp.org/resources/article/building-ethics-into-privacy-frameworks-for-big-data-and-ai/> (accessed Jan. 29, 2021).
- [38] European Commission, "COMMITTEE AND THE COMMITTEE OF THE REGIONS: A European strategy on Cooperative Intelligent Transport Systems, a milestone towards cooperative, connected and automated mobility," 2016. Accessed: Jan. 29, 2021. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX-%3A52016DC0766&from=EN>.
- [39] Samsung Research, "Principles for AI Ethics." <https://research.samsung.com/artificial-intelligence> (accessed Dec. 15, 2020).
- [40] Microsoft, "Responsible AI Principles from Microsoft," 2020. Accessed: Dec. 15, 2020. [Online]. Available: <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3Aprimaryr6>.
- [41] OpenAI, "About OpenAI," 2019, [Online]. Available: <https://openai.com/about/>.
- [42] Association of Computing Machinery, "Code of Ethics and Professional Conduct," 2018. <https://www.acm.org/code-of-ethics> (accessed Jan. 29, 2021).

- [43] Association for the Advancement of Artificial Intelligence, "CODE OF PROFESSIONAL ETHICS AND CONDUCT," 2018. <https://www.aaai.org/Conferences/code-of-ethics-and-conduct.php> (accessed Jan. 29, 2021).
- [44] Society of Actuaries, "Ethical Use of Artificial Intelligence for Actuaries 2 Ethical Use of Artificial Intelligence for Actuaries," 2019.
- [45] Digital Signage Federation, "Digital Signage Privacy Standards," 2011. Accessed: Jan. 29, 2021. [Online]. Available: <http://www.ftc.gov/os/2009/02/P085400behavadreport.pdf>.
- [46] Appen, "Appen Announces Crowd Code of Ethics to Build Better AI | Appen," 2019. <https://appen.com/blog/appen-announces-crowd-code-of-ethics-to-build-better-ai/> (accessed Jan. 29, 2021).
- [47] Kogniz, "Code of conduct." <https://web.archive.org/web/20200614133201/https://www.kogniz.com/code-of-conduct> (accessed Jan. 29, 2021).
- [48] AI Singapore and PwC Singapore, "AI Singapore and PwC Singapore collaborate to enhance digital trust competencies through responsible AI | AI Singapore," 2018. <https://www.aisingapore.org/2018/11/ai-singapore-and-pwc-singapore-collaborate-to-enhance-digital-trust-competencies-through-responsible-ai/?highlight=governance> (accessed Jan. 29, 2021).
- [49] Government of Canada, "Canada-France Statement on Artificial Intelligence," 2018. [https://www.international.gc.ca/world-monde/international\\_relations-relations\\_internationales/europe/2018-06-07-france\\_ai-ia\\_france.aspx?lang=eng](https://www.international.gc.ca/world-monde/international_relations-relations_internationales/europe/2018-06-07-france_ai-ia_france.aspx?lang=eng) (accessed Jan. 29, 2021).
- [50] Element AI, "Element AI and Mozilla Foundation Partner to Build Data Trusts and Advocate for the Ethical Data Governance of AI | Element AI," 2019. <https://www.elementai.com/press-room/element-ai-and-mozilla-foundation-partner-to-build-data-trusts-and-advocate-for-the-ethical-data-governance-of-ai> (accessed Jan. 29, 2021).
- [51] Algorithmic Justice League and Center on Technology & Privacy at Georgetown Law, "Safe Face Pledge," 2018. <https://www.safefacepledge.org/> (accessed Dec. 07, 2020).
- [52] Foundation for Responsible Robotics, "FRR Quality Mark for (AI Based) Robotics." <https://responsiblerobotics.org/quality-mark/> (accessed Dec. 15, 2020).
- [53] Government of Finland, "Enterprises to join in the ethical use of artificial intelligence," 2018. <https://www.tekoalyaika.fi/en/background/ethics/> (accessed Jan. 29, 2021).
- [54] Future of Life Institute, "An Open Letter to the United Nations Convention on Certain Conventional Weapons - Future of Life Institute," 2017. <https://futureoflife.org/autonomous-weapons-open-letter-2017/> (accessed Jan. 29, 2021).
- [55] University of Ottawa, "Call for an International Ban on the Weaponization of Artificial Intelligence | Centre for Law, Technology and Society | University of Ottawa," 2017. <https://techlaw.uottawa.ca/bankilleraai> (accessed Jan. 29, 2021).
- [56] NGO Coalition, "The Campaign To Stop Killer Robots," website, 2020. <https://www.stopkillerrobots.org/> (accessed Dec. 07, 2020).



- [57] Campaign for a commercial-free childhood, “Nov. 9. 2015. Advocates Say ‘Hell No Barbie’ to Stop Mattel from Spying on Kids - Campaign for Commercial Free Childhood,” 2015. <https://commercialfreechildhood.org/advocates-say-hell-no-barbie-to-stop-mattel-from-spying-on-kids/> (accessed Jan. 29, 2021).
- [58] Twitter, “Synthetic and manipulated media policy,” 2019. <https://help.twitter.com/en/rules-and-policies/manipulated-media> (accessed Jan. 29, 2021).
- [59] K. W. Abbott and D. Snidal, “The governance triangle: Regulatory standards institutions and the shadow of the state,” *Polit. Glob. Regul.*, vol. 44, pp. 44–88, 2009.
- [60] K. W. Abbott, G. E. Marchant, and E. A. Corley, “Soft law oversight mechanisms for nanotechnology,” *Jurimetrics*, pp. 279–312, 2012.
- [61] Ada Lovelace Institute, “Our Strategy: 2019-2020,” 2019.
- [62] T20, *The Future of Work and Education for the Digital Age*. 2018.
- [63] Amnesty International and Access Now, “The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems,” 2018. Accessed: Feb. 01, 2021. [Online]. Available: <http://www.ohchr.org/EN/ProfessionalInterest/Pages/Vienna.aspx>.
- [64] AlgorithmWatch, “The ADM Manifesto,” 2019. <https://algorithmwatch.org/en/the-adm-manifesto/> (accessed Feb. 01, 2021).
- [65] Samsung, “Artificial Intelligence | Samsung Research,” 2019. <https://research.samsung.com/artificial-intelligence> (accessed Jan. 29, 2021).
- [66] Sony Corporation, “AI Engagement within Sony Group Sony Group AI Ethics Guidelines,” pp. 1–3, Accessed: Dec. 15, 2020. [Online]. Available: [https://www.sony.net/SonyInfo/csr\\_report/humanrights/AI\\_Engagement\\_within\\_Sony\\_Group.pdf](https://www.sony.net/SonyInfo/csr_report/humanrights/AI_Engagement_within_Sony_Group.pdf).
- [67] Telefónica, “AI Principles | Our commitments | Responsible Business | Telefónica,” 2018. <https://www.telefonica.com/en/web/responsible-business/our-commitments/ai-principles> (accessed Dec. 16, 2020).
- [68] S. Pichai, “AI at Google: our principles,” 2018-07-07, p. 1, 2018, Accessed: Dec. 15, 2020. [Online]. Available: <https://www.blog.google/technology/ai/ai-principles/>.
- [69] International Chamber of Commerce, “ICC Advertising and Marketing Communications Code,” 2018.
- [70] Interactive Advertising Bureau, “Artificial Intelligence in Marketing Where Brands and Consumers Meet Through Data,” 2019.
- [71] American Medical Association, “H-480.939 Augmented Intelligence in Health Care,” 2019. [https://policysearch.ama-assn.org/policyfinder/detail/augmented intelligence?uri=%2FAMADoc%2FHOD.xml-H-480.939.xml](https://policysearch.ama-assn.org/policyfinder/detail/augmented%20intelligence?uri=%2FAMADoc%2FHOD.xml-H-480.939.xml) (accessed Feb. 01, 2021).
- [72] Association for Computing Machinery, “Principles for Algorithmic Transparency and Accountability,” 2017.

- [73] Government of Kenya, "Emerging Digital Technologies for Kenya: Exploration & Analysis," 2019. [Online]. Available: <https://www.ict.go.ke/wp-content/uploads/2019/07/blockchain.pdf>.
- [74] Government of India, "Artificial Intelligence Task Force," 2018. <https://www.aitf.org.in/> (accessed Feb. 01, 2021).
- [75] Government of Spain, "Spanish RDI Strategy in Artificial Intelligence," 2019. Accessed: Dec. 16, 2020. [Online]. Available: <https://cpage.mpr.gob.es>.
- [76] Government of Australia - National Transport Commission (NTC), "Changing driving laws to support automated vehicles," 2018. <https://www.ntc.gov.au/transport-reform/ntc-projects/changing-driving-laws-support-AVs>.
- [77] C. I. Gutierrez, G. E. Marchant, and L. Tournas, "Lessons for Artificial Intelligence from Historical Uses of Soft Law Governance - A Conclusion," *Jurimetrics*, vol. 61, no. 1, 2021.
- [78] IBM, "IBM Cognitive Principles," 2017. <https://www.ibm.org/responsibility/policies#cognitivePrinciples> (accessed Dec. 16, 2020).
- [79] Tom Moore, "Our Guiding Principles of Artificial Intelligence at AT&T," 2019. [https://about.att.com/innovation-blog/2019/05/our\\_guiding\\_principles.html](https://about.att.com/innovation-blog/2019/05/our_guiding_principles.html) (accessed Dec. 16, 2020).
- [80] Daimler AG, "Daimler Principles for Artificial Intelligence | Daimler > Sustainability > Data," 2019. <https://www.daimler.com/sustainability/data/ki-guidelines.html> (accessed Dec. 16, 2020).
- [81] Tieto, "Tieto strengthens commitment to ethical use of AI," 2018. <https://www.tietoevry.com/en/newsroom/all-news-and-releases/press-releases/2018/10/tieto-strengthens-commitment-to-ethical-use-of-ai/>.
- [82] Apple, "Our Approach to Automated Driving System Safety," 2019. Accessed: Dec. 16, 2020. [Online]. Available: <https://www.apple.com/privacy/>.
- [83] Government of Austria, "Automated - Connected - Mobile." 2016.
- [84] CIFAR, "Pan-Canadian AI Strategy," 2017. Accessed: Dec. 15, 2020. [Online]. Available: <https://cifar.ca/ai/>.
- [85] Government of Germany, Government of France, and Government of Japan, "DFG, German Research Foundation - Trilateral French-Japanese-German Research Projects on Artificial Intelligence," 2019. [https://www.dfg.de/en/research\\_funding/announcements\\_proposals/2019/info\\_wissenschaft\\_19\\_58/index.html](https://www.dfg.de/en/research_funding/announcements_proposals/2019/info_wissenschaft_19_58/index.html) (accessed Dec. 16, 2020).
- [86] Government of India, "Indo-French Roadmap on Cybersecurity and Digital Technology (August 22, 2019)," 2019. <https://mea.gov.in/bilateral-documents.htm?dtl/31757/IndoFrench+Roadmap+on+Cybersecurity+and+Digital+Technology+August+22+2019> (accessed Dec. 16, 2020).
- [87] "Karnataka partners with Nasscom for CoE in data science and artificial intelligence - The Hindu BusinessLine," 2018. <https://www.thehindubusinessline.com/info-tech/karnataka-partners-with-nasscom-for-coe-in-data-science-and-artificial-intelligence/article9963988.ece> (accessed Dec. 16, 2020).
- [88] Center for Explainable and Efficient AI Technologies, "CEE AI," 2019. <https://cee-ai.org/cee-ai/> (accessed Dec. 16, 2020).



- [89] D. Schrag, "Ensuring Responsible Use of Artificial Intelligence | Belfer Center for Science and International Affairs," 2018. <https://www.belfercenter.org/publication/ensuring-responsible-use-artificial-intelligence> (accessed Dec. 16, 2020).
- [90] NEC Group, "NEC Group AI and Human Rights Principles," NEC Corp., no. 3, pp. 4–5, 2019, Accessed: Dec. 15, 2020. [Online]. Available: <https://www.nec.com/en/press/201904/images/0201-01-01.pdf>.
- [91] K. Bigelow, "GE Healthcare," 2018, Accessed: Dec. 15, 2020. [Online]. Available: <https://www.gehealthcare.com/article/ethics-in-healthcare-arent-new-but-their-application-has-never-been-more-important/>.
- [92] Government of Belgium - Flanders, "Flemish Policy Plan for Artificial Intelligence," 2019. <https://www.ewi-vlaanderen.be/nieuws/vlaams-actieplan-artificiele-intelligentie-gelanceerd> (accessed Dec. 16, 2020).
- [93] Government of the European Union, "Declaration of Cooperation on Artificial Intelligence." 2018.
- [94] Government of the UK, "Centre for Data Ethics (CDEI) 2 Year Strategy - GOV.UK," 2019. Accessed: Dec. 16, 2020. [Online]. Available: <https://www.gov.uk/government/publications/the-centre-for-data-ethics-and-innovation-cdei-2-year-strategy/centre-for-data-ethics-cdei-2-year-strategy>.
- [95] Government of Singapore, "Digital Government Blueprint," 2018.
- [96] The Scottish Government, "Independent Advisory Group on the Use of Biometric Data in Scotland," 2018.
- [97] Huawei, "Thinking Ahead About Security and Privacy Protection," no. September, 2019, Accessed: Dec. 15, 2020. [Online]. Available: <https://e.huawei.com/ae/material/event/HC/87b18e61b7864b5cb186e3dcd6205e7f>.
- [98] Genesys, "It's Time to Talk About AI Ethics," 2018. <https://www.genesys.com/blog/post/its-time-to-talk-about-ai-ethics> (accessed Dec. 16, 2020).
- [99] SAP, "SAP's guiding principles for artificial intelligence (AI)," 2018. <https://news.sap.com/2018/09/sap-guiding-principles-for-artificial-intelligence/> (accessed Dec. 15, 2020).
- [100] Sidewalk Labs, "Master Innovation & Development Plan - Digital Innovation Appendix," Toronto, 2019.
- [101] ELISA, "Elisa Ethical Principles for Artificial Intelligence," 2019. Accessed: Dec. 15, 2020. [Online]. Available: <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities->.
- [102] Defense Innovation Board, "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense," 2019.
- [103] IEEE, "The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)," 2019. <https://standards.ieee.org/industry-connections/ecpais.html> (accessed Dec. 16, 2020).
- [104] K. Berkler, "White paper: Building trust in artificial intelligence - Fraunhofer IAIS," <https://www.iais.fraunhofer.de/en/press/press-release-190702.html> (accessed Dec. 16, 2020).

- [105] German Research Center for Artificial Intelligence and TÜV SÜD, "TÜV for Artificial Intelligence," 2018. <https://www.tuvsud.com/en/e-ssentials-newsletter/automotive-essentials/e-ssentials-1-2018/tuv-sud-and-dfki-develop-tuv-for-artificial-intelligence> (accessed Dec. 16, 2020).
- [106] For Humanity, "SAFEAI - A Tool For Concerned Parents," 2017. <https://www.forhumanity.center/blog-posts/2017/7/18/safeai-a-tool-for-concerned-parents> (accessed Dec. 16, 2020).
- [107] AAAI, "AAAI CODE OF PROFESSIONAL ETHICS AND CONDUCT," 2019. <https://aaai.org/Conferences/code-of-ethics-and-conduct.php> (accessed Dec. 15, 2020).
- [108] R. E. Anderson, "ACM Code of Ethics and Professional Conduct," *Commun. ACM*, vol. 35, no. 5, pp. 94–99, 1992, doi: 10.1145/129875.129885.
- [109] P. Grindrod and J. Bernabé Moreno, "Code of conduct - Oxford Munich Code of Conduct," 2018. <http://www.code-of-ethics.org/code-of-conduct/> (accessed Dec. 16, 2020).
- [110] Telefónica Deutschland, "Focused on People: Telefónica Deutschland declares six ethical principles regarding artificial intelligence | Telefónica Deutschland," 2019. <https://www.telefonica.de/news/press-releases-telefonica-germany/2019/02/focused-on-people-telefonica-deutschland-declares-six-ethical-principles-regarding-artificial-intelligence.html> (accessed Dec. 16, 2020).
- [111] Kogniz, "Code of conduct." <https://www.kogniz.com/code-of-conduct>.
- [112] HireVue, "Bias, AI ethics and the HireVue approach," 2019. <https://www.hirevue.com/why-hirevue/ai-ethics>.
- [113] Axon, "Second Report of the Axon AI & Policing Technology Ethics Board: Automated License Plate Readers," 2019.
- [114] Axon, "First report of the axon ai & policing technology ethics board," 2019. Accessed: Dec. 07, 2020. [Online]. Available: [https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/5d13d7e1990c4f00014c0aeb/1561581540954/Axon\\_Ethics\\_Board\\_First\\_Report.pdf](https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/5d13d7e1990c4f00014c0aeb/1561581540954/Axon_Ethics_Board_First_Report.pdf)<https://www.policingproject.org/axon>.
- [115] Government of Denmark, "National Strategy for Artificial Intelligence," 2019. Accessed: Dec. 16, 2020. [Online]. Available: [https://niti.gov.in/writereaddata/files/document\\_publication/NationalStrategy-for-AI-Discussion-Paper.pdf](https://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf).
- [116] Government of Lithuania, "Lithuanian Artificial Intelligence Strategy: A Vision of the Future," 2019.
- [117] Government of Japan, "Social Principles of Human-Centric AI," 2019.
- [118] Government of Finland, "Finland's Age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence Objective and recommendations for measures," 2017.
- [119] Estonian Ministry of Economic Affairs and Communications, "Eesti tehisintellekti kasutuselevõtu eksperdirühma aruanne (Estonian Artificial Intelligence: Report of the Expert Group on Deployment)," 2019. Accessed: Dec. 15, 2020. [Online]. Available: [https://www.riigikantselei.ee/sites/default/files/riigikantselei/strateegiaburoo/eesti\\_tehisintellekti\\_kasutuselevotu\\_eksperdiruhma\\_aruanne.pdf](https://www.riigikantselei.ee/sites/default/files/riigikantselei/strateegiaburoo/eesti_tehisintellekti_kasutuselevotu_eksperdiruhma_aruanne.pdf).

- [120] Government of Russia, "Decree of the President of the Russian Federation on the Development," 2019.
- [121] The Partnership on AI, "About - The Partnership on AI," <https://www.partnershiponai.org/about/> (accessed Dec. 15, 2020).
- [122] BDVA and euRobotics, "Joint Vision Paper for an Artificial Intelligence Public Private Partnership (AI PPP)," 2019. Accessed: Dec. 16, 2020. [Online]. Available: [https://www.eu-robotics.net/sparc/upload/Monitoring-report-2017-final-SPARC-2018\\_5v0-with-annexes.pdf](https://www.eu-robotics.net/sparc/upload/Monitoring-report-2017-final-SPARC-2018_5v0-with-annexes.pdf).
- [123] Digital Catapult and Machine Intelligence Garage, "Ethics Framework," 2018. <https://www.migarage.ai/ethics/ethics-framework/> (accessed Dec. 16, 2020).
- [124] American Academy of Sleep Medicine, "Artificial Intelligence in Sleep Medicine Committee," 2019. <https://aasm.org/membership/volunteer/committees-current/> (accessed Dec. 16, 2020).
- [125] American Society of Neuroradiology, "Artificial Intelligence Task Force," 2019. <https://www.asnr.org/research/asnr-study-groups/artificial-intelligence/artificial-intelligence-task-force/> (accessed Dec. 16, 2020).
- [126] Council of Europe, "Unboxing artificial intelligence: 10 steps to protect human rights," 2019. <https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights> (accessed Feb. 02, 2021).
- [127] American Civil Liberties Union, "An Ethical Framework for Facial Recognition," 2014.
- [128] J. R. Geis et al., "Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement," *Radiology*, vol. 293, no. 2, pp. 436–440, Nov. 2019, doi: 10.1148/radiol.2019191586.
- [129] Green Digital Working Group, "Position on Robotics and Artificial Intelligence," 2016.
- [130] Chinese Academy of Sciences and Institute of Automation, "Harmonious Artificial Intelligence Principles," 2018. <http://bii.ia.ac.cn/hai/> (accessed Feb. 02, 2021).
- [131] European Parliament, "A comprehensive European industrial policy on artificial intelligence and robotics - Tuesday, 12 February 2019," 2019. [https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_EN.html) (accessed Feb. 02, 2021).
- [132] Solita, "How to navigate the ethical challenges of using AI in business and society The Impact of AI," 2019.
- [133] Sage, "The Ethics of Code: Developing AI for Business with Five Core Principles," 2017. <https://www.sage.com/~media/group/files/business-builders/business-builders-ethics-of-code.pdf> (accessed Feb. 03, 2021).
- [134] Employer Assistance and Resource Network on Disability Inclusion, "Use of Artificial Intelligence to Facilitate Employment Opportunities for People with Disabilities," 2019. Accessed: Feb. 03, 2021. [Online]. Available: <https://www2.deloitte.com/insights/us/en/focus/human-capital-trends/2017/predictive-hiring-talent-acquisition.html>.
- [135] European Commission, "The future of work? Work of the future!," 2019. <https://ec.europa.eu/digital-single-market/en/news/future-work-work-future> (accessed Feb. 02, 2021).



- [136] Cisco, "Human Rights Position Statements," 2018. <https://webcache.googleusercontent.com/search?q=cache:j6yTEbYdcMKJ:https://www.cisco.com/c/dam/assets/csr/pdf/Human-Rights-Position-Statements-2018.pdf+&c-d=1&hl=en&ct=clnk&gl=us> (accessed Feb. 03, 2021).
- [137] Internet Society, "Artificial Intelligence and Machine Learning: Policy Paper," 2017. Accessed: Feb. 03, 2021. [Online]. Available: [https://en.wikipedia.org/wiki/Artificial\\_intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence).
- [138] Department of Health Abu Dhabi, "The AI Lab." [Online]. Available: <https://www.doh.gov.ae/ai-assets/doc/ai-lab-playbook-v13-print.pdf>.
- [139] Salesforce, "Ethics - Salesforce Research." <https://einstein.ai/ethics> (accessed Feb. 03, 2021).
- [140] Artificial Intelligence and the Internet of Things in Smart Sustainable Cities in Latin America, "The Buenos Aires Declaration," 2018.
- [141] Council of the European Union, "Thematic Reviews on 'Social Aspects of Digitalisation' (SPC) and 'Digitalisation and Robotisation of Work' (EMCO) - Endorsement of the joint key messages," 2018. <https://data.consilium.europa.eu/doc/document/ST-14411-2018-INIT/en/pdf> (accessed Feb. 03, 2021).
- [142] Starbutter, "Ethics and Principles." <https://web.archive.org/web/20201124121357/https://www.starbutter.com/ethics-and-principles> (accessed Feb. 03, 2021).
- [143] Government of the UK and Centre for Data Ethics and Innovation, "Deepfakes and Audiovisual Disinformation," 2019. <https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation> (accessed Feb. 03, 2021).
- [144] D. Tarraf et al., The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations. RAND Corporation, 2019.
- [145] Government of Mauritius, "Mauritius Artificial Intelligence Strategy," 2018. [https://mitci.govmu.org/Documents/Strategies/Mauritius\\_AI\\_Strategy.pdf#search=Ai strategy](https://mitci.govmu.org/Documents/Strategies/Mauritius_AI_Strategy.pdf#search=Ai%20strategy) (accessed Feb. 03, 2021).
- [146] Computing Community Consortium, "AI Roadmap - 5. Recommendations," 2019. <https://cra.org/ccc/ai-roadmap-recommendations/> (accessed Feb. 03, 2021).
- [147] Ethicsnet, "Dataset on Ethics." [Online]. Available: <https://www.ethicsnet.org/about>.
- [148] IA Latam, "Statement of Ethics for the Design, Development, and Use of Artificial Intelligence." [Online]. Available: <https://ia-latam.com/etica-ia-latam/>.
- [149] IEEE, "Ethics In Action in Autonomous and Intelligent Systems," 2019. <https://ethicsinaction.ieee.org/> (accessed Feb. 03, 2021).
- [150] Government of the European Union, "Policy and Investment Recommendations for Trustworthy Artificial Intelligence," 2019. [Online]. Available: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60343](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60343).

- [151] Government of the Republic of Korea, "Mid- to Long-Term Master Plan in Preparation for the Intelligent Information Society," 2016. [Online]. Available: <http://www.msip.go.kr/dynamic/file/afieldfile/msse56/1352869/2017/07/20/Master Plan for the intelligent information society.pdf>.
- [152] British Standards Institution, "Robots and robotic devices: Guide to the ethical design and application of robots and robotic systems," 2016. <https://drive.google.com/file/d/0B0McUk5vAotHcDYzbzh2Y0c4aUE/view> (accessed Feb. 03, 2021).
- [153] PATH, "Develops Ethical Guidelines on the Use of AI in Healthcare," 2019. <https://www.pnewswire.com/news-releases/path-develops-ethical-guidelines-on-the-use-of-ai-in-healthcare-300882699.html> (accessed Feb. 03, 2021).
- [154] Academy of Medical Royal Colleges, "Artificial Intelligence in Healthcare," 2019.
- [155] International Organization for Standardization, "ISO - IEC/TR 60601-4-1:2017 - Medical electrical equipment – Part 4-1: Guidance and interpretation – Medical electrical equipment and medical electrical systems employing a degree of autonomy," 2017. <https://www.iso.org/standard/70755.html?browse=tc> (accessed Feb. 03, 2021).
- [156] European Parliament, "Texts adopted - Civil Law Rules on Robotics - Thursday, 16 February 2017," 2017. [https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_EN.html?redirect](https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html?redirect) (accessed Feb. 04, 2021).
- [157] EPSRC, "Principles of robotics - EPSRC," 2010. <https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/> (accessed Feb. 04, 2021).
- [158] Rathenau Instituut, "Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality Report," 2017.
- [159] International Chamber of Commerce, "ICC Policy Statement on Artificial Intelligence," 2018. <http://www.iccin-diaonline.org/ps2019/1.pdf> (accessed Feb. 04, 2021).
- [160] FAT/ML, "Principles for Accountable Algorithms and a Social Impact Statement for Algorithms," 2018. <https://www.fatml.org/resources/principles-for-accountable-algorithms> (accessed Feb. 04, 2021).
- [161] Government of Australia, "Artificial Intelligence: Australia's Ethics Framework," 2019. Accessed: Feb. 04, 2021. [Online]. Available: <https://consult.industry.gov.au/>.
- [162] International Biometrics & Identification Association, "IBIA Privacy Best Practice Recommendations for Commercial Biometric Use," 2014. Accessed: Feb. 04, 2021. [Online]. Available: [www.ibia.org](http://www.ibia.org).
- [163] IBM, "Data Responsibility at IBM," 2017. <https://www.ibm.com/blogs/policy/dataresponsibility-at-ibm/> (accessed Feb. 04, 2021).
- [164] Government of Japan, "Report on Artificial Intelligence and Human Society Unofficial translation Advisory Board on Artificial Intelligence and Human Society," 2017.
- [165] UNESCO, "Steering AI and advanced ICTs for knowledge societies: a Rights, Openness, Access, and Multi-stakeholder Perspective," 2019. <https://unesdoc.unesco.org/ark:/48223/pf0000372132> (accessed Feb. 04, 2021).

- [166] Government of Sweden, *Artificial Intelligence in Swedish Business and Society: Analysis of Development and Potential*. 2018.
- [167] Government of Taiwan, "AI Taiwan," 2018. <https://ai.taiwan.gov.tw/> (accessed Feb. 04, 2021).
- [168] U. and D. A. I. Toward a Controlled, "Toward a Controlled, Useful and Demystified Artificial Intelligence," 2017. Accessed: Feb. 04, 2021. [Online]. Available: [www.senat.fr](http://www.senat.fr).
- [169] Business for Social Responsibility, "Artificial Intelligence: A Rights-Based Blueprint for Business," 2018. <https://www.bsr.org/en/our-insights/report-view/artificial-intelligence-a-rights-based-blueprint-for-business> (accessed Feb. 04, 2021).
- [170] Alliance for Multilateralism, "Declaration by the Alliance for Multilateralism on Lethal Autonomous Weapons Systems (LAWS)," 2019.
- [171] E. Palmerini et al., "RoboLaw Grant Agreement number: 289092 Project acronym: RoboLaw Project title: Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics Funding scheme: Collaborative project (CP), FP7-SiS-Challenge 1-3: Regulating emerging scientific and technological developments D6.2 Guidelines on Regulating Robotics," 2014. Accessed: Feb. 04, 2021. [Online]. Available: [www.robolaw.eu](http://www.robolaw.eu).
- [172] Government of India, "Report of Committee - D: On Cyber Security, Safety, and Ethical Issues," 2019.
- [173] E. Chavannes and A. Arkhipov-Goyal, "The Ethics of Robotic and Autonomous Systems in a Military Context," 2019.
- [174] Ethics and Religious Liberty Commission, "Artificial Intelligence: An Evangelical Statement of Principles." <https://erlc.com/resource-library/statements/artificial-intelligence-an-evangelical-statement-of-principles>.
- [175] Government of Canada, "Responsible AI in the Government of Canada." <https://docs.google.com/document/d/1Sn-qBZUXEUG4dVk909eSg5qvfbpNIRhzlefWPtBwbxY/edit#> (accessed Dec. 07, 2020).
- [176] UNICRI, "United Nations Interregional Crime and Justice Research Institute," 2019. [http://www.unicri.it/in\\_focus/on/interpol\\_unicri\\_report\\_ai](http://www.unicri.it/in_focus/on/interpol_unicri_report_ai) (accessed Feb. 05, 2021).
- [177] Isle of Man Government, "Code of Practice for the testing of autonomous vehicles," 2017.
- [178] Université de Montréal, "Declaration of Montréal for a responsible development of AI," 2018. <https://www.montrealdeclaration-responsibleai.com/> (accessed Feb. 05, 2021).
- [179] Microsoft Corporation, "Healthcare, Artificial Intelligence, Data and Ethics - A 2030 Vision," 2018.
- [180] Government of Canada, "Responsible AI in the Government of Canada," 2018. <https://docs.google.com/document/d/1Sn-qBZUXEUG4dVk909eSg5qvfbpNIRhzlefWPtBwbxY/edit> (accessed Feb. 05, 2021).
- [181] Government of Japan, "Draft AI Utilization Principles," 2018.



- [182] Foundation for Responsible Robotics, “FRR Quality Mark for (AI Based) Robotics - Responsible Robotics,” 2019. <https://responsiblerobotics.org/quality-mark/> (accessed Jan. 29, 2021).
- [183] Government of Malta, “Malta: Towards Trustworthy AI,” 2019. [https://malta.ai/wp-content/uploads/2019/10/Malta\\_Towards\\_Ethical\\_and\\_Trustworthy\\_AI\\_vFINAL.pdf](https://malta.ai/wp-content/uploads/2019/10/Malta_Towards_Ethical_and_Trustworthy_AI_vFINAL.pdf) (accessed Feb. 05, 2021).
- [184] Council of Europe, “Guidelines on Artificial Intelligence and Data Protection,” 2019. Accessed: Feb. 05, 2021. [Online]. Available: <https://www.coe.int/en/web/human-rights-rule-of-law/artificial-intelligence/glossary>:
- [185] Integrated Justice Information Systems Institute International Association of Chiefs of Police, “Law Enforcement Facial Recognition Use Case Catalog,” 2019.
- [186] Select Committee on Artificial Intelligence of the National Science & Technology Council, “The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update,” 2019.
- [187] United Nations, “Report of the Working Party on Road Traffic Safety on its seventy-fourth session Contents,” 2017.
- [188] ANSI UASSC, “Standardization Roadmap for Unmanned Aircraft Systems,” 2018.
- [189] Government of the UK, “AI Sector Deal,” 2019. <https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal> (accessed Feb. 05, 2021).
- [190] International Organization for Standardization, “Intelligent transport systems – Assisted parking system (APS) – Performance requirements and test procedures,” 2017. <https://www.iso.org/standard/73768.html?browse=tc> (accessed Feb. 05, 2021).
- [191] BusinessEurope, “Artificial Intelligence recommendations,” 2019. <https://www.buinesseurope.eu/publications/artificial-intelligence-recommendations> (accessed Feb. 05, 2021).

## Appendix 1 – Sources of information - linkhubs

#	Title	Authors	Site
1	The global landscape of AI ethics guidelines	Anna Jobin, Marcello Ienca and Effy Vayena	<a href="https://www.nature.com/articles/s42256-019-0088-2">https://www.nature.com/articles/s42256-019-0088-2</a>
2	Algorithm Watch	-	<a href="https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/">https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/</a>
3	Principled Artificial Intelligence Project	Berkman Klein Center	<a href="https://clinic.cyber.harvard.edu/2019/06/07/introducing-the-principled-artificial-intelligence-project/">https://clinic.cyber.harvard.edu/2019/06/07/introducing-the-principled-artificial-intelligence-project/</a>
4	AI Ethics - Too Principled to Fail?	Brent Mittelstadt	
5	Artificial Intelligence in Society	OECD	<a href="https://www.oecd-ilibrary.org/sites/cf3f3be0-en/index.html?itemId=/content/component/cf3f3be0-en">https://www.oecd-ilibrary.org/sites/cf3f3be0-en/index.html?itemId=/content/component/cf3f3be0-en</a>
6	An Overview of National AI Strategies	Medium	<a href="https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd">medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd</a>
7	Best Practices for Common Uses of Facial Recognition Technologies	FTC	<a href="https://www.ftc.gov/sites/default/files/documents/reports/facing-facts-best-practices-common-uses-facial-recognition-technologies/121022facialtechrpt.pdf">https://www.ftc.gov/sites/default/files/documents/reports/facing-facts-best-practices-common-uses-facial-recognition-technologies/121022facialtechrpt.pdf</a>
8	Privacy Multistakeholder Process	NTIA	<a href="https://www.ntia.doc.gov/files/ntia/publications/background_resources_facial_recognition_4_25_2014.pdf">https://www.ntia.doc.gov/files/ntia/publications/background_resources_facial_recognition_4_25_2014.pdf</a>
9	Global AI Policy	Future of Life Institute	<a href="https://futureoflife.org/ai-policy/">https://futureoflife.org/ai-policy/</a>
10	Artificial Intelligence Governance and Ethics: Global Perspectives	-	<a href="https://arxiv.org/ftp/arxiv/papers/1907/1907.03848.pdf">https://arxiv.org/ftp/arxiv/papers/1907/1907.03848.pdf</a>
11	AI Repository	ITU	<a href="https://www.itu.int/en/ITU-T/AI/Pages/ai-repository.aspx">https://www.itu.int/en/ITU-T/AI/Pages/ai-repository.aspx</a>
12	Reports	All-Party Parliamentary Group on Artificial Intelligence	<a href="http://www.appg-ai.org/reports/">http://www.appg-ai.org/reports/</a>
13	Policy Papers	All-Party Parliamentary Group on Artificial Intelligence	<a href="http://www.appg-ai.org/policy-papers/">http://www.appg-ai.org/policy-papers/</a>
14	Evidence	All-Party Parliamentary Group on Artificial Intelligence	<a href="https://www.appg-ai.org/evidence/">https://www.appg-ai.org/evidence/</a>
15	An Updated Round Up of Ethical Principles of Robotics and AI	Alan Winfield's Web Log	<a href="http://alanwinfield.blogspot.com/2019/04/an-updated-round-up-of-ethical.html">http://alanwinfield.blogspot.com/2019/04/an-updated-round-up-of-ethical.html</a>
16	Artificial Intelligence: the global landscape of ethics guidelines	Anna Jobin, Marcello Ienca, Effy Vayena	<a href="https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf">https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf</a>
17	Awesome AI Guidelines	-	<a href="https://github.com/EthicalML/awesome-artificial-intelligence-guidelines">https://github.com/EthicalML/awesome-artificial-intelligence-guidelines</a>
18	Business, industry and NGO standards on AI and machine learning	AI Law Consultancy	<a href="https://ai-lawhub.com/business/">https://ai-lawhub.com/business/</a>
19	The Ethics of AI Ethics	Dr. Thilo Hagendorff	<a href="https://arxiv.org/ftp/arxiv/papers/1903/1903.03425.pdf">https://arxiv.org/ftp/arxiv/papers/1903/1903.03425.pdf</a>
20	Building Ethics into Artificial Intelligence	Han Yu, Zhiqi Shen, Chunyan Miao, Cyril Leung, Victor R. Lesser, Qiang Yang	<a href="http://www.ntu.edu.sg/wp-content/uploads/conference/Building_Ethics_into_Artificial_Intelligence_accepted.pdf">http://www.ntu.edu.sg/wp-content/uploads/conference/Building_Ethics_into_Artificial_Intelligence_accepted.pdf</a>
21	OECD		<a href="https://www.oecd.org/going-digital/ai/initiatives-worldwide/">https://www.oecd.org/going-digital/ai/initiatives-worldwide/</a>
22	The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions	Jess Whittlestone, Rune Nyrop, Anna Alexandrova and Stephen Cave	<a href="http://lcfi.ac.uk/media/uploads/files/AIES-19_paper_188_Whittlestone_Nyrop_Alexandrova_Cave_OcF7jnp.pdf">http://lcfi.ac.uk/media/uploads/files/AIES-19_paper_188_Whittlestone_Nyrop_Alexandrova_Cave_OcF7jnp.pdf</a>
23	Linking Artificial Intelligence Principles	-	<a href="http://www.linking-ai-principles.org/principles">http://www.linking-ai-principles.org/principles</a>
24	New perspectives on ethics and the laws of artificial intelligence	-	<a href="https://policyreview.info/articles/analysis/new-perspectives-ethics-and-laws-artificial-intelligence">https://policyreview.info/articles/analysis/new-perspectives-ethics-and-laws-artificial-intelligence</a>
25	Report on Artificial Intelligence: Part I – the existing regulatory landscape	-	<a href="https://www.howtoregulate.org/artificial_intelligence/">https://www.howtoregulate.org/artificial_intelligence/</a>
26	Regulation of Artificial Intelligence	Library of Congress	<a href="https://www.loc.gov/law/help/artificial-intelligence/index.php">https://www.loc.gov/law/help/artificial-intelligence/index.php</a>
27	Alan Winfield's Web Log	-	<a href="http://alanwinfield.blogspot.com/2017/12/a-round-up-of-robotics-and-ai-ethics.html">http://alanwinfield.blogspot.com/2017/12/a-round-up-of-robotics-and-ai-ethics.html</a>
28	Towards a Code of Ethics for Artificial Intelligence	Paula Boddington	<a href="https://link.springer.com/content/pdf/10.1007/978-3-319-60648-4.pdf">https://link.springer.com/content/pdf/10.1007/978-3-319-60648-4.pdf</a>
29	Overview of ethics codes and principles for AI	Rethenau Instituut	<a href="https://www.rathenau.nl/en/digital-society/overview-ethics-codes-and-principles-ai">https://www.rathenau.nl/en/digital-society/overview-ethics-codes-and-principles-ai</a>
30	Resources	Leverhulme Centre for the Future of Intelligence	<a href="http://lcfi.ac.uk/resources/">http://lcfi.ac.uk/resources/</a>
31	Ethical frameworks, tool kits, principles, and oaths - Oh my!	Kathy Baxter	<a href="https://blog.einstein.ai/frameworks-tool-kits-principles-and-oaths-oh-my/">https://blog.einstein.ai/frameworks-tool-kits-principles-and-oaths-oh-my/</a>
32	The Oxford Handbook of Ethics of AI: Online Supplement	Oxford	<a href="https://c4ejournal.net/the-oxford-handbook-of-ethics-of-ai-online-companion/">https://c4ejournal.net/the-oxford-handbook-of-ethics-of-ai-online-companion/</a>

33	The Regulation of Artificial Intelligence – A Case Study of the Partnership on AI	Gijs Leenders	<a href="https://becominghuman.ai/the-regulation-of-artificial-intelligence-a-case-study-of-the-partnership-on-ai-c1c22526c19f">https://becominghuman.ai/the-regulation-of-artificial-intelligence-a-case-study-of-the-partnership-on-ai-c1c22526c19f</a>
34	Background research on policy initiatives in the area of artificial intelligence	EU Agency for Fundamental Rights (FRA)	-
35	ASU Database on AI Governance	ASU	-
36	BSI White Paper – Overview of standardization landscape in artificial intelligence	BSI	<a href="https://www.bsigroup.com/en-GB/industries-and-sectors/artificial-intelligence/download-the-artificial-intelligence-whitepaper/">https://www.bsigroup.com/en-GB/industries-and-sectors/artificial-intelligence/download-the-artificial-intelligence-whitepaper/</a>
37	What is the State of Artificial Intelligence Governance Globally?	James Butcher & Irakli Beridze	<a href="https://www.tandfonline.com/doi/full/10.1080/03071847.2019.1694260">https://www.tandfonline.com/doi/full/10.1080/03071847.2019.1694260</a>
38	Artificial Intelligence and Autonomous Systems Legal Update (2Q19)	Gibson Dunn	<a href="https://www.gibsondunn.com/artificial-intelligence-and-autonomous-systems-legal-update-2q19/">https://www.gibsondunn.com/artificial-intelligence-and-autonomous-systems-legal-update-2q19/</a>
39	Establishing AI Ethics – Public and Private Sector Initiatives	Emerj	<a href="https://emerj.com/ethics-and-regulatory/establishing-ai-ethics-public-and-private-sector/">https://emerj.com/ethics-and-regulatory/establishing-ai-ethics-public-and-private-sector/</a>
40	Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development	Future of Humanity Institute	<a href="https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_FHI-Technical-Report.pdf">https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_FHI-Technical-Report.pdf</a>
41	Repository	Oceanis	<a href="https://ethicsstandards.org/repository/">https://ethicsstandards.org/repository/</a>
42	AI Strategy in EU Towards 2020	Medium	<a href="https://towardsdatascience.com/ai-strategy-in-eu-2018-2019-44393b7eaf28">https://towardsdatascience.com/ai-strategy-in-eu-2018-2019-44393b7eaf28</a>
43	Artificial intelligence across industries	IEC	<a href="https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en-jp">https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en-jp</a>
44	Administering Artificial Intelligence	Alicia Solow-Niederman	<a href="https://robots.law.miami.edu/2019/wp-content/uploads/2019/03/Administering-Artificial-Intelligence-We-Robot.pdf">https://robots.law.miami.edu/2019/wp-content/uploads/2019/03/Administering-Artificial-Intelligence-We-Robot.pdf</a>
45	Artificial Intelligence and Machine Learning: Ethics, Governance and Compliance Resources	Future of Privacy Forum	<a href="https://fpf.org/artificial-intelligence-and-machine-learning-ethics-governance-and-compliance-resources/">https://fpf.org/artificial-intelligence-and-machine-learning-ethics-governance-and-compliance-resources/</a>
46	AI Strategies & Public Sector Components	Observatory of Public Sector Innovation	<a href="https://oecd-opsi.org/projects/ai/strategies/">https://oecd-opsi.org/projects/ai/strategies/</a>
47	Linking Artificial Intelligence Principles	Institute of Automation, Chinese Academy of Sciences	<a href="http://ceur-ws.org/Vol-2301/paper_15.pdf">http://ceur-ws.org/Vol-2301/paper_15.pdf</a>
48	Professional Codes of Ethics	Ethics for Artificial Intelligence	<a href="https://www.cs.ox.ac.uk/efai/is-ai-ethics-special/professional-codes-of-ethics/">https://www.cs.ox.ac.uk/efai/is-ai-ethics-special/professional-codes-of-ethics/</a>
49	From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices	Jessica Morley, Luciano Floridi, Libby Kinsey & Anat Elhalal	<a href="https://link.springer.com/article/10.1007/s11948-019-00165-5">https://link.springer.com/article/10.1007/s11948-019-00165-5</a>
50	Mapping Regulatory Proposals for Artificial Intelligence in Europe	Access Now	<a href="https://www.accessnow.org/cms/assets/uploads/2018/11/mapping_regulatory_proposals_for_AI_in_EU.pdf">https://www.accessnow.org/cms/assets/uploads/2018/11/mapping_regulatory_proposals_for_AI_in_EU.pdf</a>
51	Artificial Intelligence Index Report 2019	AI Index Steering Committee, Human-Centered AI Institute, Stanford University	<a href="https://hai.stanford.edu/sites/g/files/sbiybj10986/t/ai_index_2019_report.pdf">https://hai.stanford.edu/sites/g/files/sbiybj10986/t/ai_index_2019_report.pdf</a>
52	In Pursuit of Autonomy: Examining Twelve National A.I. Strategies	Observer Research Foundation	<a href="https://www.orfonline.org/wp-content/uploads/2018/11/Ai_Book.pdf">https://www.orfonline.org/wp-content/uploads/2018/11/Ai_Book.pdf</a>
53	Artificial Intelligence: An Overview of State Initiatives	FutureGrasp	<a href="http://www.unicri.it/in_focus/files/Report_AI-An_Overview_of_State_Initiatives_FutureGrasp_7-23-19.pdf">http://www.unicri.it/in_focus/files/Report_AI-An_Overview_of_State_Initiatives_FutureGrasp_7-23-19.pdf</a>
54	Comparison of National Strategies to Promote Artificial Intelligence - Part 2	Cambrian	<a href="https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+2.pdf/4c6f3a0d-beaa-09f3-1db4-c66467739653?version=1.1&amp;t=1560500520623">https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+2.pdf/4c6f3a0d-beaa-09f3-1db4-c66467739653?version=1.1&amp;t=1560500520623</a>
55	Comparison of National Strategies to Promote Artificial Intelligence - Part 1	Cambrian	<a href="https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&amp;t=1560500570070">https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&amp;t=1560500570070</a>
56	A round up of robotics and AI ethics: part 1 principles	Robohub	<a href="https://robohub.org/a-round-up-of-robotics-and-ai-ethics-part-1-principles/">https://robohub.org/a-round-up-of-robotics-and-ai-ethics-part-1-principles/</a>
57	A Review of Future and Ethical Perspectives of Robotics and AI	Robotics and Intelligent Systems Group, Department of Informatics, University of Oslo	<a href="https://www.frontiersin.org/articles/10.3389/frobt.2017.00075/full">https://www.frontiersin.org/articles/10.3389/frobt.2017.00075/full</a>
58	Artificial Intelligence (AI) R&D - Interagency Working Group	NITRD	<a href="https://www.nitrd.gov/nitrdgroups/index.php?title=AI">https://www.nitrd.gov/nitrdgroups/index.php?title=AI</a>



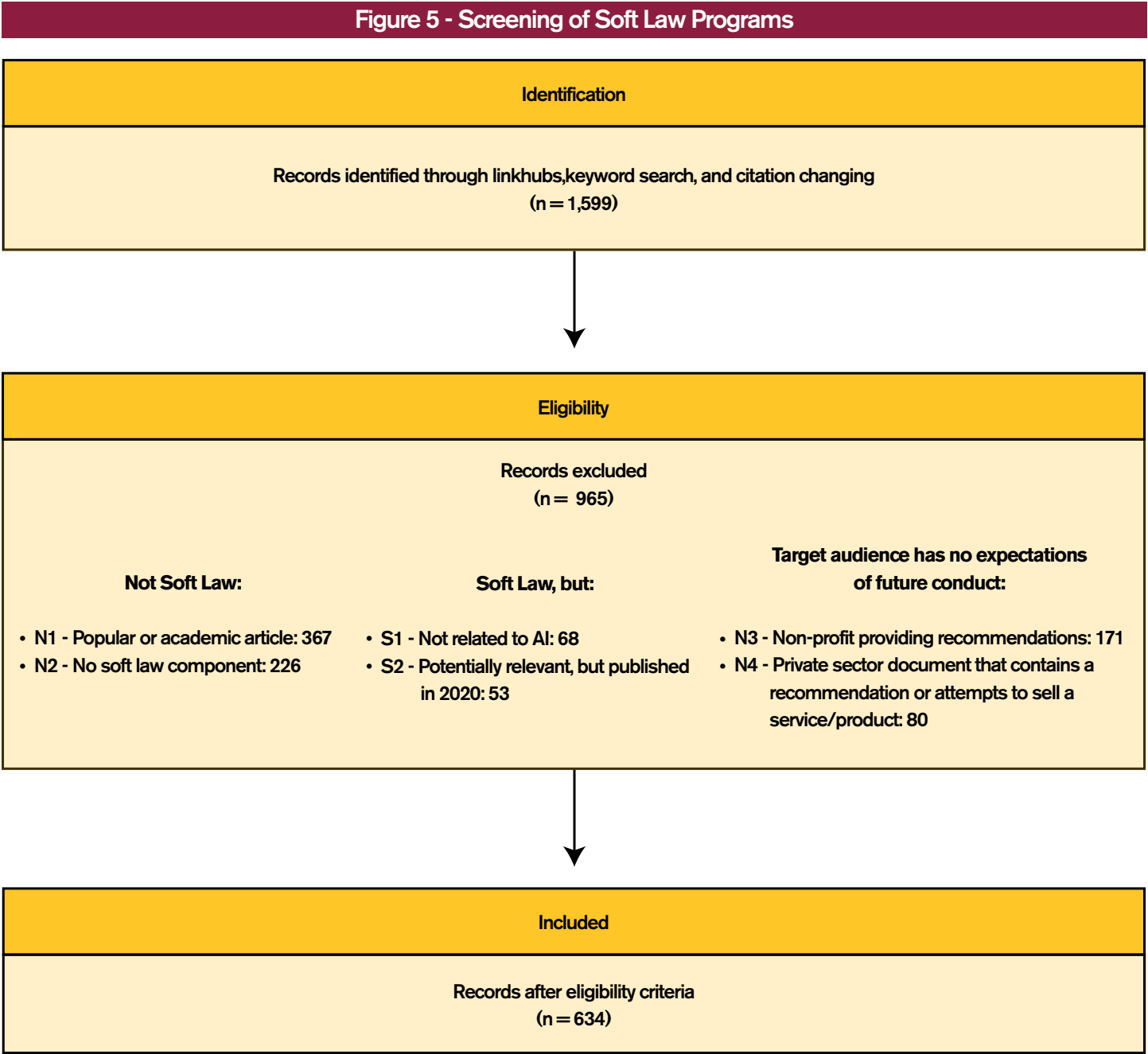
59	OECD AI Policy Observatory	OECD	<a href="https://oecd.ai/">https://oecd.ai/</a>
60	Artificial Intelligence Across Industries	International Electrotechnical Commission	<a href="https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/#">https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/#</a>
61	BUILDING AN AI WORLD Report on National and Regional AI Strategies	CIFAR	<a href="https://www.cifar.ca/docs/default-source/ai-society/buildinganaeworld_eng.pdf?sfvrsn=fb18d129_4">https://www.cifar.ca/docs/default-source/ai-society/buildinganaeworld_eng.pdf?sfvrsn=fb18d129_4</a>
62	Toward ethical, transparent and fair AI/ML: a critical reading list	Eirini Malliaraki	<a href="https://medium.com/@eirinimalliaraki/toward-ethical-transparent-and-fair-ai-ml-a-critical-reading-list-d950e70a70ea">https://medium.com/@eirinimalliaraki/toward-ethical-transparent-and-fair-ai-ml-a-critical-reading-list-d950e70a70ea</a>
63	Can AI Fairness Be Regulated?	Ann R. Thryft	<a href="https://www.eetimes.com/can-ai-fairness-be-regulated/">https://www.eetimes.com/can-ai-fairness-be-regulated/</a>
64	U.S. AI Regulation Guide: Legislative Overview and Practical Considerations	Yoon Chae; The Journal of Robotics, Artificial Intelligence & Law	<a href="https://www.bakermckenzie.com/-/media/files/people/chae-yoon/rail-us-ai-regulation-guide.pdf">https://www.bakermckenzie.com/-/media/files/people/chae-yoon/rail-us-ai-regulation-guide.pdf</a>
65	Ethics, Governance, and Compliance Resources	Future of Privacy Forum	<a href="https://sites.google.com/fpf.org/futureofprivacyforumresources/ethics-governance-and-compliance-resources?-authuser=1">https://sites.google.com/fpf.org/futureofprivacyforumresources/ethics-governance-and-compliance-resources?-authuser=1</a>
66	Universal Guidelines for Artificial Intelligence	The Public Voice	<a href="https://thepublicvoice.org/ai-universal-guidelines/memo/">https://thepublicvoice.org/ai-universal-guidelines/memo/</a>
67	AI Ethics Toolkits	Intel	<a href="https://www.intel.ai/ai-ethics-toolkits/#gs.x3ye7t">https://www.intel.ai/ai-ethics-toolkits/#gs.x3ye7t</a>
68	Key requirements for AI fairness	Ann R. Thryft	<a href="https://www.embedded.com/key-requirements-for-ai-fairness/">https://www.embedded.com/key-requirements-for-ai-fairness/</a>
69	Want to get serious about artificial intelligence? You'll need an AI strategy	Hannah Miller, André Petheram and Emma Martinho-Truswell	<a href="https://www.oxfordinsights.com/insights/2018/1/23/aistrategies">https://www.oxfordinsights.com/insights/2018/1/23/aistrategies</a>
70	Accountability in AI: Promoting Greater Societal Trust	Dr. Jason Millar, Brent Barron, Dr. Koichi Hori, Rebecca Finlay, Kentaro Kotsuki, Dr. Ian Kerr	<a href="https://www.ic.gc.ca/eic/site/133.nsf/vwapj/3_Discussion_Paper_-_Accountability_in_AI_EN.pdf/\$FILE/3_Discussion_Paper_-_Accountability_in_AI_EN.pdf">https://www.ic.gc.ca/eic/site/133.nsf/vwapj/3_Discussion_Paper_-_Accountability_in_AI_EN.pdf/\$FILE/3_Discussion_Paper_-_Accountability_in_AI_EN.pdf</a>
71	Related Projects	Unbias	<a href="https://unbias.wp.horizon.ac.uk/initiatives-and-research-projects/">https://unbias.wp.horizon.ac.uk/initiatives-and-research-projects/</a>
72	Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems	IEEE	<a href="https://ieeexplore.ieee.org/document/8662743">https://ieeexplore.ieee.org/document/8662743</a>
73	Mapping AI Governance	Nesta	<a href="https://www.nesta.org.uk/data-visualisation-and-interactive/mapping-ai-governance/">https://www.nesta.org.uk/data-visualisation-and-interactive/mapping-ai-governance/</a>
74	From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices	Jessica Morley, Luciano Floridi, Libby Kinsey, Anat Elhalal	<a href="https://medium.com/@jessicamorley/applied-ai-ethics-reading-resource-list-ed9312499c0a">https://medium.com/@jessicamorley/applied-ai-ethics-reading-resource-list-ed9312499c0a</a>
75	AI Global Surveillance (AIGS) Index	Carnegie Endowment	<a href="https://carnegieendowment.org/files/AI_Global_Surveillance_Index1.pdf">https://carnegieendowment.org/files/AI_Global_Surveillance_Index1.pdf</a>
76	2020 and the European National Strategies for Artificial Intelligence	Alex Moltzau	<a href="https://medium.com/@alexmoltzau/2020-and-the-european-national-strategies-for-artificial-intelligence-6b66a6b2adf9">https://medium.com/@alexmoltzau/2020-and-the-european-national-strategies-for-artificial-intelligence-6b66a6b2adf9</a>
77	A survey of the European Union's artificial intelligence ecosystem	Charlotte Stix	<a href="https://83d6fa69-2c07-4589-82a1-386547d3715c.filesusr.com/ugd/ff3afe_1513c6bf-2d81400eac182642105d4d6f.pdf">https://83d6fa69-2c07-4589-82a1-386547d3715c.filesusr.com/ugd/ff3afe_1513c6bf-2d81400eac182642105d4d6f.pdf</a>
78	Publications	AI Now Institute	<a href="https://ainowinstitute.org/reports.html">https://ainowinstitute.org/reports.html</a>
79	The AI Policy Landscape	Matt Chessen	<a href="https://medium.com/artificial-intelligence-policy-laws-and-ethics/the-ai-landscape-ea8a8b3c3d5d">https://medium.com/artificial-intelligence-policy-laws-and-ethics/the-ai-landscape-ea8a8b3c3d5d</a>
80	Artificial Intelligence for the American People	US Government	<a href="https://www.whitehouse.gov/ai/ai-american-innovation/">https://www.whitehouse.gov/ai/ai-american-innovation/</a>
81	AI Ethicist	AI Ethicist	<a href="https://www.aiethicist.org/">https://www.aiethicist.org/</a>
82	White Paper: Artificial intelligence across industries	International Electrotechnical Commission	<a href="https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/?wpdm-dl=1795&amp;ind=1540285831573">https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/?wpdm-dl=1795&amp;ind=1540285831573</a>
83	Digital Policies Frameworks	Council of Europe	<a href="https://docs.google.com/spreadsheets/d/1mU2brATV_fgD5MRGfT2ASOFepAI1piwWhGm0VCT22_U/edit#gid=0">https://docs.google.com/spreadsheets/d/1mU2brATV_fgD5MRGfT2ASOFepAI1piwWhGm0VCT22_U/edit#gid=0</a>
84	New Tech, New Threats, and New Governance Challenges: An Opportunity to Craft Smarter Responses?	Carnegie Endowment for International Peace	<a href="https://carnegieendowment.org/2019/08/28/new-tech-new-threats-and-new-governance-challenges-opportunity-to-craft-smarter-responses-pub-79736">https://carnegieendowment.org/2019/08/28/new-tech-new-threats-and-new-governance-challenges-opportunity-to-craft-smarter-responses-pub-79736</a>

## Appendix 2 – Keyword search

**Table 27 - Search Terms for Keyword Search**

Types of Soft Law	Themes and Applications		Technology of Interest
Principles	Agriculture	Military	Artificial intelligence
	Artificial General Intelligence	Monitoring	
Certification	Aviation	Music	
	Bias	National Security	
Voluntary programs	Contestability	News	
	Deepfakes	Outreach	
Recommendations	Diversity	Policing	
	Due Process	Privacy	
Strategy	Education	Private sector development	
	Ethics	Public goods	
Professional guidelines	Explainability	Robots	
	Fairness	Safety	
Code of conduct	Finance	Security	
	Health	Sensors	
Private standards	Human resources	Social improvement	
	Labor	Transparency	
Best practices	Marketing	Transportation	
	Meaningful Human Control	Urban Planning	
Public-private partnerships		Values	

Appendix 3 – Screening of soft law programs





## Appendix 4 – Codebook for database

Database Variables	
Field	Description
Name of soft law program	Self-explanatory
Name of sponsoring organization	Self-explanatory
Year	Year the program was published
Types of Soft Law Program	
Recommendations and strategies:	<p>The common denominator for these programs is their description of different varieties of actionable items.</p> <p><b>Strategies</b> are roadmaps that highlight the direction an organization wishes to pursue and the actions needed to reach it.</p> <p><b>Recommendations</b> are a set of ideas, proposals, or evidence-based actions to improve the status quo. To qualify as such, they should be imposed on a particular target audience. The important element of a recommendation is that they must be actionable items targeted at stakeholders who have a reason to follow them. Recommendations by an entity without any relationship to its target are not included. For instance, recommendations to policymakers by a non-profit unassociated or with funding from a government entity would not qualify.</p>
Principles:	<p>Broad statements that serve as high-level norms for any type of organization. They are unlikely to include actionable items, instead they are an organization's aspirational objectives.</p> <p>Examples:</p> <ul style="list-style-type: none"> <li>• Technology company defines principles it will abide by in the development of products.</li> <li>• Non-profit defines its core beliefs related to AI research.</li> </ul>
Professional guidelines or codes of conduct:	<p>Expectations of behavior applicable to individuals that work with AI applications or methods.</p> <p>Example:</p> <ul style="list-style-type: none"> <li>• Non-profit creates a pledge for AI tech workers with the aim of protecting consumers from facial recognition technology.</li> <li>• A private sector firm releases a code of conduct for its employees and customers.</li> </ul>
Standards:	<p>Programs requiring an independent third party to monitor and/or verify the compliance or performance of an entity with specific criteria and in a limited scope. They are created by standard setting organizations (SSO) such as IEEE, BSI, among others.</p> <p>Example:</p> <ul style="list-style-type: none"> <li>• Standards set by the ISO that homologate the vocabulary and definitions utilized in the AI field.</li> </ul>
Certification or voluntary programs:	<p>A certification ranges from a "seal of approval" or a statement mentioning compliance to a set of pre-defined characteristics. Educational certifications should not be included in the database. On the other hand, a voluntary program is generally a government initiative that invites non-government entities (private sector and non-profits) to comply with a set of actions or guidelines.</p> <p>Examples:</p> <ul style="list-style-type: none"> <li>• A seal of quality on an AI application that creates a signal for consumers regarding its characteristics.</li> </ul>

<b>Voluntary moratorium or ban:</b>	<p>Call of action to avoid or cease the usage of an AI application or method. Generally, these can be expected of technologies that cause harm or negatively affect individuals.</p> <p>Example:</p> <ul style="list-style-type: none"> <li>• Non-profit advocates for a ban on the utilization of autonomous weapon systems.</li> </ul>
<b>Partnerships:</b>	<p>An initiative in which two or more entities collaborate to advance a particular purpose.</p> <p>Examples:</p> <ul style="list-style-type: none"> <li>• Facebook and a University partner to create a research center dedicated to AI research.</li> <li>• A public-private partnership between government and the private sector geared towards exploring the military applications of AI.</li> <li>• Several research centers ban together to share resources to focus on a problem or research question.</li> </ul>
<b>Led by:</b>	<p>Notes the type of organization responsible for the soft law program or were major funders of it. These are catalogued as:</p> <p>Ps = Private Sector Gov = Government Np = Non-Profit</p> <p>Note: Thus far, we have cataloged non-profits that represent the interest of a sector as Np and Ps. Standards organizations usually involve the participation of all parties (Gov, Np, Ps).</p>
<b>Level of government involved</b>	<p>In case a government entity appears in the led by column, readers will find whether this includes:</p> <p>National: federal or national authority State: provincial or state authority Local: local authority Multilevel: combination of government levels</p>
<b>Participated in:</b>	<p>Only includes organizations that had a secondary role in the development of a program. In other words, they participated, but did not lead the effort (e.g. were interviewed for it, etc.). Utilizes the same abbreviations for organizations as above.</p>
<b>Notes on participation</b>	<p>Any notes that provide context regarding participating organizations. If multi-party, a note will be included describing the two or more types of organizations that organized the program.</p>
<b>Purpose</b>	<p>Describes the objective of the program.</p>
<b>Influence</b>	<p>Identifies if a program is meant to be applied only to the organization that created it (e.g. Google Principles) or external parties (e.g. OECD principles). Its implementation in the database would be in the form of a binary indicator:</p> <ul style="list-style-type: none"> <li>• 1 if it applies just for it itself.</li> <li>• 2 if it applies to itself and/or other entities.</li> </ul> <p>When cataloguing influence for government entities (or any other similar situations), this database considers the hierarchy of the institutions. For instance, if a branch of government creates a recommendation for a different branch (executive to legislative), this would classify as a 2. If the executive branch of government creates a mechanism with recommendations for units within said branch, it would be classified as a 1.</p>
<b>Compliance / enforceability / implementation: yes or no?</b>	<p>If the program mentions any means to regulate compliance or enforcement, then this variable will be a Yes.</p>

<b>Notes on: compliance / enforceability / implementation</b>		Describes any mention of actions meant to require the organization to comply with, enforce, or implement the mechanism.
<b>Country/region of origin</b>		Denotes the origin of the program.
<b>Jurisdiction</b>		Identifies the program's geographic scope of influence.
<b>Themes</b>		
<b>Accountability</b>		
<b>General mentions of accountability</b>		Mentions the term accountability in a general sense. They stand out by their lack of detail or explanation as to how the entity harnesses the term in their soft law program.
<b>Processes to ensure accountability</b>		Mechanisms or actions employed by the organization to ensure accountability in the use of AI applications or methods.
<b>What entity is ultimately responsible for the actions of AI</b>	<b>Organizational responsibility</b>	Responsibility and/or accountability of organizations due to the decision-making of AI applications and methods.
	<b>Human responsibility</b>	Responsibility and/or accountability of organizations due to the decision-making of AI applications and methods.
<b>Artificial general intelligence (AGI)</b>		
<b>Development and governance of AGI</b>		As a technology in development, statements in this sub-theme focus on actions that assist in making AGI a reality or prepare humanity for its consequences.
<b>Goals and traits of AGI</b>		Discussion of the goals or possible traits of AGI systems. This can mention calls for the technology to be safe or describe the human-like qualities that we hope or fear will be present in AGI systems.
<b>Bias</b>		
<b>General mention of discrimination or bias</b>		General statements mentioning the words discrimination or bias.
<b>Diversity</b>		Considering diverse groups of people to reflect the population or the inclusion of individuals with a variety of characteristics (e.g. skills or demographics) in the creation of AI systems.
<b>Inclusion</b>		Providing minority populations with access to the use and development of AI systems. In other words, the empowerment of disadvantaged groups.
<b>Mechanisms for entities</b>		Actions taken by entities to limit the bias and/or discrimination by AI systems.
<b>Environment</b>		
<b>Environmental protection generally</b>		Broad normative statements about the relationship between AI and the environment.



Accountability		
General mentions of accountability		Mentions the term accountability in a general sense. They stand out by their lack of detail or explanation as to how the entity harnesses the term in their soft law program.
Processes to ensure accountability		Mechanisms or actions employed by the organization to ensure accountability in the use of AI applications or methods.
What entity is ultimately responsible for the actions of AI	Organizational responsibility	Responsibility and/or accountability of organizations due to the decision-making of AI applications and methods.
	Human responsibility	Responsibility and/or accountability of organizations due to the decision-making of AI applications and methods.
Artificial general intelligence (AGI)		
Development and governance of AGI		As a technology in development, statements in this sub-theme focus on actions that assist in making AGI a reality or prepare humanity for its consequences.
Goals and traits of AGI		Discussion of the goals or possible traits of AGI systems. This can mention calls for the technology to be safe or describe the human-like qualities that we hope or fear will be present in AGI systems.
Bias		
General mention of discrimination or bias		General statements mentioning the words discrimination or bias.
Diversity		Considering diverse groups of people to reflect the population or the inclusion of individuals with a variety of characteristics (e.g. skills or demographics) in the creation of AI systems.
Inclusion		Providing minority populations with access to the use and development of AI systems. In other words, the empowerment of disadvantaged groups.
Mechanisms for entities		Actions taken by entities to limit the bias and/or discrimination by AI systems.
Environment		
Environmental protection generally		Broad normative statements about the relationship between AI and the environment.
Efficient use/conservation of resources		Covers the role of AI in improving the conservation of resources and avoiding waste. It may also mention the use of “smart” systems in areas such as agriculture and the harm caused by the procurement of inputs needed to manufacture AI components (e.g. semiconductors).
Disaster management		AI can be a powerful logistical tool that can help in directing resources and personnel during natural disasters. This sub-theme considers how AI or other smart systems may affect the management of natural disasters.

Education / displacement of labor		
Research	Research projects generally	General mentions of AI-related research are flagged in this sub-theme.
	Link between society and research	Linking research with society, which can include technology transfer opportunities, commercialization of AI discoveries, or partnerships with academia to bring research to the public, among others.
Education	AI talent	Procurement of human resources that are trained in the management, development, or teaching of AI systems.
	Skills / retraining	In order to adjust to a new labor market, individuals will need to complement their toolkit with new skills. This sub-theme does not mention any skills related to AI.
	AI literacy	Any educational effort related to improving AI literacy will be placed in this sub-theme. This consists of activities such as helping the public increase their understanding of how AI systems function and encouraging STEM and Computer Science Education.
	AI aiding education	Incorporating AI in assisting or complementing the educational system (e.g. applications of AI in the classroom or it providing custom educational materials).
Labor markets changes	Job loss and avoid job loss	Role of AI systems in replacing human jobs and the importance of preventing this trend.
	Solutions to job loss	Mechanisms or means of mitigating the harm caused by job loss due to AI systems (e.g. universal basic income or unemployment pay).
	AI helping people get jobs or improve their current job	Scenarios where AI applications and methods are beneficial in the procurement of employment or improve the efficiency in the completion of a task.
Ethics		
Protection of rights		Relationship between AI systems affecting the rights of individuals, examples include constitutional to human rights.
Ethical consequences		Examination of the negative ethical consequences brought about by the use of AI systems.
Mechanisms to ensure AI ethics		Actions to ensure that AI remains ethical, be it through the implementation of controls, creating an advisory board or a code of conduct, re-training, establishing ethical best practices, cross-sector collaboration, or ethics-by-design.
Values		This sub-theme mentions values as desirable characteristics for AI systems (e.g., friendliness, trust, dignity, etc.).
Culture		Interaction between AI systems and culture.
General mentions of AI ethics		Instances where the term “ethics” is mentioned, but not described in detail.
AI must be for a social benefit		AI systems should be developed or used for the benefit of society or for positive social change.

Health	
Standards and oversight	Creation of standards or oversight mechanisms for AI in the healthcare industry. Examples include organizing an international task force or expert group to develop AI healthcare standards or ensuring that communities understand the benefits of standardization.
Access	Ensuring access to healthcare services related to AI.
Benefits of AI in healthcare writ large (to hospital, people, and field)	Any mention of AI benefiting healthcare should be included here, be it to hospitals, clinicians, or the healthcare field generally. Examples may be improved patient outcomes, new treatment methods, and improved screening procedures for diseases.
Patient health	Highlights patient safety when AI is used in the healthcare industry.
Collaboration of clinicians and developers	Clinicians and developers should work together to ensure AI is utilized and safely implemented as a tool for the healthcare industry to use.
Meaningful Human Control	
Human control and involvement in AI decision-making	Covers broad statements about humans being in control of AI systems. Includes discussion of the level of control or involvement of humans in the decision-making process using terms such as “human-in-the-loop” or “human-over-the-loop”.
User consent	Emphasis on procuring the consent of users/individuals prior to engaging with an AI system.
Right to refuse	Mechanisms describing how a user has a right to, at any time, not be subject to the decision of an AI system.
Human autonomy	The right of a human to make their own decisions without undue influence. Discusses the role of AI systems in persuading or nudging the decisions of individuals.
Right to redress or review	Individuals have a right/ability to seek an explanation for decisions made by AI systems, have these overturned, or dispute them after the fact.
Development and feedback	Stakeholders of an AI system (e.g. the public) should be able to contribute to its development.
Privacy	
Individual privacy	Any statements about the importance of user, customer, or general privacy concerns.
Privacy protection mechanisms	Systems or mechanisms in place to protect privacy. Common examples include data encryption, minimizing data collection, purging old data, privacy policies to inform users, requests for user consent, etc. This category does not include any mechanism reliant on AI systems to accomplish these goals.
AI privacy systems	AI-enabled privacy protection mechanisms. AI systems are used for data anonymization or pseudonymization in most cases. Example: “AI can be used to anonymize data prior to analysis to protect user privacy”



<b>Compliance with laws</b>		Includes regulations and laws that institutions must follow to ensure privacy and data protection. Common example in the EU is the GDPR.
<b>Private Sector Development</b>		
<b>Entrepreneurship</b>		Efforts to support AI startups or investments in the advancement of the AI industry with a focus on medium and small enterprises.
<b>Competitiveness</b>		Any action related to improving the competitiveness of an entity and AI systems.
<b>General mention of support for private sector development</b>		The role of government on improving the private sector's position with respect to AI systems that is not related to competitiveness.
<b>Private sector and non-government action</b>		Actions taken by entities outside of government to improve the market conditions of AI systems. These include organizational actions such as the creation of internal governance structure, creating performance indicators, or developing strategies to recognize the potential of AI.
<b>Social responsibility</b>		Private sector actions that are aligned with corporate social responsibility goals (e.g. creating sustainable development goals).
<b>Safety</b>		
<b>Restriction on AWS</b>		Examples of nations, open letters signed by NGOs, or experts calling for a moratorium, ban or other restrictions for Autonomous Weapons Systems (AWS).
<b>Military use of AI</b>		Describes future and current use of AI by militaries. This includes not just weapons, but reconnaissance, surveillance and logistics.
<b>National security</b>		The use of AI by a nation for intelligence gathering and processing. It also covers the use and risks of AI on national security.
<b>General safety</b>		Normative statements about AI being safe and avoiding or minimizing physical harm to people. Mentions of human safety and risk in relation to use of AI systems belong in this category.
<b>Safety mechanisms</b>		Specific uses of AI or safety measures taken to make the technology safer. This includes instituting procedures or processes to ensure safety, as well as putting forward standards and regulations to mandate AI safety.
<b>Law enforcement</b>		The use of AI by law enforcement, whether for surveillance through facial recognition, or any other use.
<b>Security/Reliability</b>		
<b>Security</b>	<b>Cyber security</b>	Cyber security encompasses both risks and protections to AI system integrity. This can mean hacking into an AI system or steps taken to prevent such adversarial attacks. This does not include examples that explicitly involve protecting user data or data sets used by AI. It primarily focuses on actions taken by individuals and institutions to prevent attacks, and does not include the use of AI systems for threat detection or other measures.

<b>Security</b>	<b>Data protection</b>	Data protection encompasses steps taken to prevent user data from being misappropriated either by an AI system, third parties, or developers.
	<b>Role of AI in cyber security</b>	Use of AI in threat detection, countering adversarial attacks, or to perpetrate cyber-attacks.
<b>Reliability</b>	<b>Reliability generally</b>	Covers broad normative statements about reliability, interoperability, or trustworthiness of AI systems.
	<b>Data quality</b>	Statements about data provenance and other factors that might affect quality of input data.
	<b>Verifiability</b>	Oversight mechanisms including audits, third party checks, among others, that allow an AI system's functionality to be confirmed.
	<b>Controlling for failure</b>	Mention of systems meant to offset failures by AI systems.
<b>Role of Government</b>		
<b>Role of government as a promoter of AI</b>		General government support in funding, removing barriers, developing a strategy, or other actions related to AI systems.
<b>Non-government governance efforts (private sector and non-profits)</b>		General support by non-government actors in funding, removing barriers, developing a strategy, or other actions related to AI systems.
<b>Cooperation/engagement</b>		Stakeholder engagement, public-private partnerships, creation of alliances, or participation in multilateral fora related to AI systems.
<b>Transparency and Explainability</b>		
<b>Individuals should be informed about their interaction with an AI system</b>		Suggestion or requirement to inform or confirm if individuals are interacting with AI systems.
<b>Consequential decisions should be transparent and explainable</b>		Decisions by AI systems should be made available to the people affected by them.
<b>Sharing information/open access</b>		Emphasis on sharing information and databases related to AI systems in order to advance their development or increase the knowledge of stakeholders.
<b>General transparency</b>		General mention of transparency with regards to AI systems.
<b>Open lines of communication with the public</b>		Efforts to increase the awareness surrounding AI systems to the public or generally creating open lines of communication amongst stakeholders.
<b>Data provenance</b>		Indication of where the data used by AI systems originated and discussions regarding maintaining transparency as to data usage and training.
<b>Transparency and Explainability</b>		
<b>Infrastructure</b>		Any element, physical or nonphysical, that supports urban planning and AI transportation systems.

<b>Role of the operator</b>	Role of an individual in controlling an AI transportation system and the communication between the operator and the system. This can include mentions of the levels of autonomous vehicle (1-5), as these inherently assign what sorts of tasks are to be handled by the vehicle and operator, or specific measures to allow human takeover of the vehicle.
<b>Governance of AI in transportation and urban areas</b>	Topics where government or other organizations set requirements for the operation of autonomous vehicles and infrastructure. This can include the creation of “test beds” for field testing of autonomous systems or licensing requirements.
<b>Autonomous aircraft</b>	Use of autonomous aircraft for civilian purposes. This can include drone AI for deliveries or transportation of individuals.
<b>Traffic safety and management</b>	Systems for reducing traffic congestion, reducing risk of accidents, use of inter-vehicle communication, and any other means of improving the flow or safety of traffic.
<b>Urban planning efficiency</b>	Efficiency here includes sustainable urban planning and resource management related to AI. This does not include traffic efficiency as this pertains to the traffic management sub-theme.







**ASU** Sandra Day O'Connor  
College of Law  
Arizona State University