

FAccT-Check on AI regulation: Systematic Evaluation of AI Regulation on the Example of the Legislation on the Use of AI in the Public Sector in the German Federal State of Schleswig-Holstein

Katharina Simbeck
HTW Berlin
simbeck@htw-berlin.de

ABSTRACT

In the framework of the current discussions about regulating Artificial Intelligence (AI) and machine learning (ML), the small Federal State of Schleswig-Holstein in Northern Germany hurries ahead and adopts legislation on the Use of AI in the public sector. The legislation aims on the one hand to enable the use of AI in the public sector by creating a legal framework and to limit its potential discriminatory effect on the other hand. Contrary to the European AI Act, which is valid for all companies and organizations in Europe, and contrary to the Chinese administrative rule on Internet information recommender systems, the Schleswig-Holstein “IT Deployment Law” (ITDL) would therefore only apply to public administrations and agencies in the federal state. The legislation addresses several AI risks, including fairness and transparency, and mitigates them with approaches quite different from the proposed European AI Act (AIA). In this paper, the legislation will be systematically reviewed and discussed with regards to its definition of AI, risk handling, fairness, accountability, and transparency.

CCS CONCEPTS

• Artificial Intelligence; • Law;

KEYWORDS

AI regulation, AI fairness, AI transparency

ACM Reference Format:

Katharina Simbeck. 2022. FAccT-Check on AI regulation: Systematic Evaluation of AI Regulation on the Example of the Legislation on the Use of AI in the Public Sector in the German Federal State of Schleswig-Holstein. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*, June 21–24, 2022, Seoul, Republic of Korea. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3531146.3533076>

1 INTRODUCTION

In recent years, it has been repeatedly shown that data-driven information technologies that use machine learning are not fair, non-discriminatory, and transparent [2, 3, 24]. Machine learning systems are designed to recognize relationships in data, model them,

and apply them to new data. If the data used to train the model contains undesirable correlations (e.g., women earn less than men), these correlations are “learned” by the model and applied to new data (e.g., for credit decisions). For this reason, there are numerous initiatives to develop methods for fair [24] and transparent [17] machine learning on the one hand, and ethical principles for its use on the other [1, 30]. Various legislation initiatives have started to regulate the use of Artificial Intelligence (AI). In this paper, the legislative initiative “IT deployment law” in the German federal state of Schleswig-Holstein [28] will be systematically discussed, as it is one of the first initiatives to regulate Artificial Intelligence. While much smaller in scope, both regionally and because of its limitation to the public sector, it is comparable to the European Union’s proposed AI Act (AIA) [13] and the Chinese Internet Information Service Algorithm Recommendation Management Regulations [6].

As the discussed legislation specifically refers to the use of AI in the public sector, it is important to state, that Germany is not well advanced in the use of technology in public sector, it only ranks 25th, well below average, in OECD’s 2019 Digital Government index [25]. In the UN’s E-Government index, it ranks well below fellow EU states, such as Denmark, Austria, France, or Belgium [31].

The following sections will first review the regulatory and standard-setting initiatives for AI, then the elements of the IT deployment law will be introduced (section 2) and discussed (section 3). This Article proposes to systematically assess AI regulation with regards to its scope/definition of AI, risk handling, fairness, accountability, transparency, and process changes. The understanding of the term AI has changed over time and depends on the context, therefore regulations will have to specify their scope. AI technologies have a wide range of useful applications, and their use should not be rejected on principle. Regulation of AI becomes necessary when the risks associated with the technology are considered unacceptable for society in general or in certain cases. For this reason, regulation needs to define which applications are always permitted, which are never permitted, and which are permitted under certain conditions. Specific technological consequences relate to fairness, accountability, and accessibility and shall be discussed in detail as well.

1.1 Discussions about the Regulation of AI

Legal regulations specifically on the use of AI that go beyond, for example, regulations on data protection or liability under existing laws are new. The legislation on the use of AI in the public sector in Schleswig-Holstein is was passed in March 2022. The European Union published a legislative proposal for an Artificial Intelligence



This work is licensed under a Creative Commons Attribution International 4.0 License.

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9352-2/22/06.
<https://doi.org/10.1145/3531146.3533076>

Act (AIA) in April 2021 [13], which will be referred to several times below. China passed an administrative rule on recommender systems which will be in effect as of March 2022.

1.1.1 The Legislation in the German Federal State of Schleswig-Holstein. In September 2021 the State Government of Schleswig-Holstein, namely the Ministry of Energy Transition, Agriculture, Environment, Nature, and Digitalization submitted a draft legislative package to promote digitization and the provision of open data and to enable the use of data-driven information technologies in public administration (Digitalization Act) [28] which was enacted by the state parliament on March 16th 2022 [34]. The small federal state of Schleswig-Holstein in northern Germany with a population of almost 3 million inhabitants contributes about 100 billion Euro to the German GDP (comparable to the GDPs of Slovakia or Puerto Rico). The Schleswig-Holstein state parliament is governed by a coalition of Christian Democrats, Greens, and Liberals.

A major pillar of the draft legislative package is the “IT-Einsatz Gesetz (ITEG)” (IT deployment law) which intends to enable and regulate the use of data-driven information technologies in the public administration of the German federal state of Schleswig-Holstein [28]. As the explanatory memorandum states, the law intends to address risks resulting from training data, but also risks from lack of explainability, traceability and transparency of administrative action. The memorandum explicitly refers to the prevention of biases and discrimination in data-based administrative decisions. On the other hand, the law is supposed to enable the administration to benefit from the use of AI technologies.

The draft law was discussed in the plenary session of the Schleswig-Holstein state parliament in September 2021 and referred to the Environment, Agriculture, and Digitization Committee. In the months of October 2021 to January 2022, the parliamentary committees requested written and oral expert comments on the draft. The law was adopted on March 16th by the state parliament.

1.1.2 The European AI Act (AIA). In contrast, the draft AIA deals with the regulation of all artificial intelligence (AI) systems used in the European Union, by private and public organizations [13]. Due to the primacy of European Union law, the Schleswig-Holstein law would be overruled by the AIA in case of conflict once both are in effect.

In the AIA, artificial intelligence is quite broadly defined in Annex I as “(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference, and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods.”

The AIA pursues a risk-based approach where systems with an unacceptable risk (Article 5: behavioral manipulation exploitation of vulnerabilities that cause harm, social scoring, remote biometric identification in public spaces) are forbidden, high-risk systems (Article 6, annex III: safety systems i.e. in aviation, specific applications, i.e. in critical infrastructure, education, human resources management, law enforcement) are regulated, and low risk or risk-free

systems are unregulated. Additionally, Article 52 defines transparency obligations: natural persons need to be aware, that they are communicating with an AI system or subject to emotion recognition/biometric categorization; generated or manipulated media need to be disclosed.

AI applications for product security and biometric identification, critical infrastructure, education, human resources, access to essential private or public services, law enforcement, migration/asylum, border security, and legal and democratic processes are defined as high-risk applications (Annex III). Extensive auditing and monitoring mechanisms are required for them (conformity assessments under Article 43 and monitoring systems under Article 61). While the conformity assessment is conducted before the system is brought to market, the monitoring takes place throughout the lifetime of the AI product in order to identify, document, and address failures of the system [32]. Conformity assessments rely on detailed documentation of system purpose, design, use, functionality and can either be conducted internally or by an external auditor (“notified body”).

The expected requirement of conformity assessments has initiated and accelerated various standard-setting initiatives on standards addressing the transparency, robustness, ethics, trustworthiness, fairness, and non-discrimination of AI systems, such as the German DIN (Deutsche Industrie Norm, German Industrial Norms Association) or the Institute of Electrical and Electronics Engineers (IEEE).

1.1.3 The Chinese Internet Information Service Algorithm Recommendation Management Regulations. In January 2022, the Chinese Administration has issued Regulations on the Administration of Algorithm Recommendations for Internet Information Services which aim to protect national security and “social public interests” as well as civil rights by regulating recommender algorithms [6]. The need for regulation is justified by discrimination through algorithms, addiction, social order, and “ideological security” [6, 7]. With the regulation, the authorities chose not to regulate artificial intelligence or machine learning in general, but specifically recommender systems. Recommender algorithms as defined as technologies used for “generation and synthesis, personalized pushing, sorting and selection, retrieval and filtering, and scheduling decisions to provide information to users” [7]. Recommender systems are required to “adhere to the mainstream value orientation” and “resist the spread of false information”, they must not be used to influence public opinion or contribute to monopoly creation [7]. Especially the elderly shall be protected from online fraud and minors from inappropriate content [18]. With regards to consumer rights, users must be provided with transparency about the use of algorithms and their underlying principles as well as with an option to choose information that is not personalized [7]. Algorithm operators are responsible for compliance with the new regulation and have to put regular review mechanisms in place.

1.2 Standards on fair, accountable, and transparent AI

In contrast to laws and regulations, adherence to norms and standards is voluntary, though often encouraged or required by partners or authorities. Norms and standards reduce transaction costs and time as well as risks by providing clear definitions, requirements, deliverables, or pre-requisites for goods, services, quality, and compatibility criteria. With regards to fair, accountable, and transparent AI systems there are several national and international standards, mostly work in process, that will describe a state of the art. The proposed AIA (Articles 9-15) requires conformity assessments that will rely on well-defined standards that don't exist yet. The ITDL enables a federal authority to define applicable norms.

The so-called delegation principle ensures that local interested parties can contribute to standardization processes without language barriers or high travel costs. National mirror committees work in parallel to international level technical committees and send experts to international consultations.

1.2.1 The German and EU AI standardization initiatives. In Germany, the DIN organization leads the national AI standard-setting process which is laid out until 2030 in the first version of the national standard-setting roadmap published in November 2020 supported by the German Federal Ministry of Economics [11]. Currently, the roadmap is being updated; bias, fairness, and robustness will continue to be discussed within the focus area “audit and certification” [10].

The relevant European standard setters are Comité Européen de Normalisation (CEN - European Committee for Standardization) and Comité Européen de Normalisation Électrotechnique (CENELEC - European Committee for Electrotechnical Standardization). The German roadmap [11], as well as the European Commissions White Paper on AI [15], have pushed the European standardization initiative by CEN and CENELEC [4]. In accordance with the AIA proposal, the European Commission is expected to issues a standardization request on AI, this is already reflected in the preliminary draft Annual Union Work Programme for standardization 2022 under the title “Safe and trustworthy Artificial Intelligence systems” [14]. The standard is supposed to address „safety and trustworthiness, including documentation, transparency, robustness, accuracy, human oversight, data training, and testing” of AI systems, respecting European fundamental values and human rights [14].

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have established a joint technical committee on Artificial Intelligence that has been working on various AI-related standards [21].

1.2.2 UK Algorithmic Transparency Standards. The UK Algorithmic Transparency Standard is specifically designed to provide guidelines for public sector organizations that use algorithmic tools for decision-making [5]. The UK's Central Digital and Data Office provides a template to systematically collect information about algorithms used in the public sector. As a next step, the Algorithmic Transparency Standard collection is supposed to be published. The UK Algorithmic Transparency Standard initiative is a result of the British National Data Strategy [8]. While the Data Strategy was

launched in September 2020, the collection has just begun and no data has been published yet.

1.2.3 IEEE Standards for Artificial Intelligence Systems. The leading technical professional organization IEEE is developing standards for the development and certification of AI as well, including standards on AI explainability and algorithmic bias considerations [19]. It has already developed an AI certification, the Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) [20] and demonstrated a proof-of-concept application in the public sector [27].

1.3 Framework for assessing Fairness, Accountability, and Transparency of AI legislation

In this paper a framework is proposed to systematically evaluate how AI legislation regulates the fields of fairness, accountability, and transparency (Figure 1).

As AI is an undefined legal concept, as a first step regulation needs to be reviewed for how AI is defined in the regulation, and where the regulation is to be applied to. The ethical issues around fairness, accountability, and transparency arise at the intersection of technology and its use in society, therefore legislation needs to be assessed in how it evaluates and approaches risks from the use of AI technologies. Artificial Intelligence is a broad technological concept that is associated with numerous ethical issues for which many guidelines have been proposed[33]. The most commonly discussed issues around the use of AI technology are privacy, fairness, accountability, and transparency [33]. While privacy is already heavily regulated (in Europe with the General Data Protection Regulation GDPR [16]), the other issues are not regulated yet.

2 ELEMENTS OF THE LEGISLATION IN SCHLESWIG-HOLSTEIN

The "IT Deployment Law" (for short: ITDL, in German: "IT-Einsatzgesetz", short ITEG) is part of a larger legislative initiative on Open Data and the digitization of the administration. In the following, the elements and structure of the law will be presented, in general, and specifically with regards to the definition of AI, the risk-based approach, and the assignment of responsibilities.

ITDL includes 13 Articles:

Article 1 states the purpose and scope of the law and defines that “data-driven information technology” may only be used while ensuring “the right to informational self-determination, and the principles of primacy of human action, human oversight and accountability, transparency, technical robustness and security, diversity, non-discrimination, fairness, and societal and ecological well-being”.

Article 2 in principle permits the use of data-driven information technology but lists four exceptions under which the use is prohibited.

Article 3 defines “data-driven information technology” and chatbots and separates three levels of automation (assisting systems, delegation, autonomous decisions). The process for assigning levels of automation is described in Article 5.

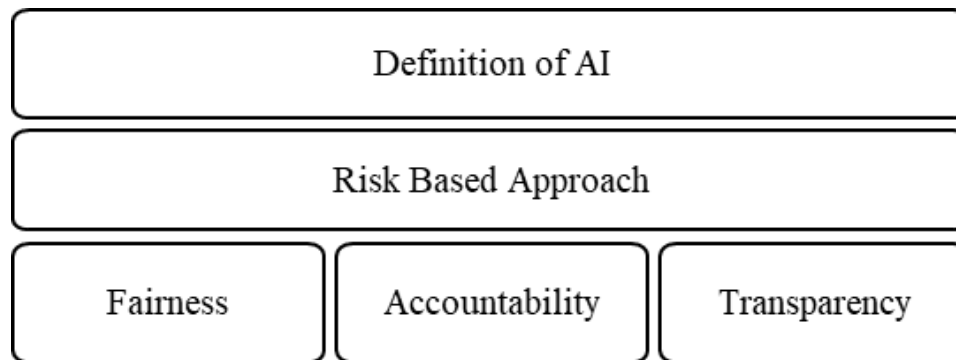


Figure 1: Framework for FAccT check of AI legislation.

Article 4 establishes that the administration that uses the data-driven technology is responsible for their consequences.

Article 6 sets standards on transparency for the use of data-based information technologies which are a prerequisite for the validity of administrative acts.

Article 7 establishes the principles of human supervision and the primacy of human decisions.

Article 8 permits the documented use of administrative data for the development and training of data-driven information technology. Personal data must be pseudonymized „provided that this does not prevent the purpose.”

Article 9 sets requirements for measures to ensure controllability and risk management depending on the level of automation. Article 10 sets requirements for measures to ensure safety, robustness, and resilience.

Article 11 authorizes the highest state authority responsible for interdepartmental information and communications technology to issue ordinances for minimum standards, for example on IT security, training data, or documentation.

Article 12 establishes a new instrument of legal remedy – the “AI reprimand” which enables addressees of administrative acts to request that the decision be revoked.

Article 13 requires the state government to submit a report on the effects of the law after four years.

2.1 Definition of artificial intelligence in ITDL

It is to be noted that in the legal text of ITDL use of the terms “artificial intelligence” or “machine learning” is avoided. Instead, the term “data-driven information technology” is used. A definition of the technologies to be regulated can thus be found in Article 1(3) of ITDL: “Data-driven information technologies include automated procedures that independently compare or interpret existing, measured, perceived or combined data from one or more data sources in order to solve complex tasks and objectives. The selection of which data is considered with which weighting is made based on previous evaluations of the procedure or on reference data and predefined evaluations.” Further in Article 3(1) data-driven information technology is defined as “...a specialized application (...) that is used to efficiently solve a specific task or a complex problem based on a data set using specialized systems, such as artificial neural networks and machine learning, and that evolves decision-making

parameters without active intervention.” The definition in Article 3(1) therefore does not include data-based systems in general but only those where parameters evolve automatically. In addition, in Article 11(2), “speech recognition, text classification, pattern recognition, data analysis, or artificial intelligence (AI) applications” are explicitly mentioned as “forms of data-driven information technologies.” Neural networks, machine learning, or speech recognition are therefore given as examples of data-driven information technology, but the regulated technology is broadened to include all data-driven information technologies.

The chatbot technology is singled out further in the ITDL. In Article 3(1), a chatbot is defined as a “text or speech-based basic service (...) that (...) can initiate a dialogue between a human and a technical system by using databases and interfaces in addressable language.”

2.2 The Risk-based approach in ITDL

The handling of risks depends on the importance of the risk, on the danger of possible consequences. In the ITDL draft, this is done in three stages:

- prohibited applications Article 2 (2),
- mitigation approaches linked to the automation level (performance of task by employees Article 5(2), AI reprimand according to Article 12),
- mitigation approaches requiring an administrative decision (Article 7, Article 6 (4), and (5)).

In Article 2 (2) the draft ITDL, defines a number of areas of application where the use of AI is generally not permitted: 1) “in the exercise of direct force against the life and physical integrity of natural persons”, 2) “for the purpose of assessing the personality, work performance, physical and mental resilience, cognitive or emotional abilities of individuals, and for making predictions about the delinquency of individuals or groups of individuals”, 3) “for the mass identification of persons (. . .) based on biometric characteristics”, and 4) in administrative acts “in which there is discretion or scope for assessment”.

In Article 3(2) three automation levels of data-driven information technologies are defined: assistance system (systems that propose relevant information or decisions), delegation (systems that continuously process well-defined tasks), and autonomous decision

(systems that process tasks and solve arising issues and take decisions even in unforeseen situations). Data-driven information technologies used in state government are to be assigned to one of the automation levels (Article 5(1)). The administration must ensure, that a failure of the automated system is recognized, that the system is halted appropriately, and that the tasks can be taken over by civil servants instead (Article 5 (2)). Depending on the level of automation, adequate measures to ensure the controllability of the technology are to be taken (Article 9 (1)), especially ensuring the precedence of human decisions, “methods for reviewing and verifying the decision-making processes”, “the creation of rights and role concepts to correct decisions”, “ways to reduce the level of automation”, processes to identify problems, reduction of the level of automation, and training of employees. Article 12 introduces a new legal reprimand that allows addressees of data-driven information technologies of automation levels 2 and 3 a right to review the decision (AI reprimand).

Article 7 (2) states that decisions prepared or made by data-driven information technology must be modifiable by public servants. The impact on personal data protection needs to be assessed before the implementation of the technology (Article 9 (2)).

Article 10 regulates the necessity of measures for security, robustness, and resilience.

2.3 Fairness in ITDL

The ITDL proposal enumerates fairness as one of the main principles for the use of AI in the public sector (Article 1(2)). However, it does not define fairness. In the legislative memorandum fairness is brought into context with the need of good faith in administrative action and with Article 3 of the German Basic Law (constitution). Article 3 states that all “persons shall be equal before the law” which implies “that equals are treated equally and unequals unequally” according to the legislative memorandum. The law prohibits the uses of AI in areas of application that have been shown to be prone to unfairness, such as personality assessments and delinquency prediction (Article 2(2)).

2.4 Accountability in ITDL

The societal consequences of the use of AI systems arise in the interplay between the algorithm used (responsibility of the provider), the data used (responsibility of the provider and/or deploying organization), and the use/interpretation of the results (responsibility of the deploying organization). It is therefore essential to clearly define accountability when using AI systems: is it with the individual system operator, the data provider, the software supplier, or the organization using the system?

In ITDG Article 4 responsibility for the use of data-driven information technologies it is clearly defined, that the responsibility lies with the administration using the system. If several administrations use the same system, they are jointly responsible and have to agree on lead responsible and contact person (Article 4 (2)). If administrations employ third parties, they need to guarantee that those follow all necessary regulations (Article 4 (3)).

ITDG in Article 7 further regulates the need for human supervision and the precedence of human decisions. Article 11 empowers

a supreme state authority to issue ordinances on IT security requirements, audit criteria for training data and proper operation, documentation, and emergency procedures. However, the authority is not responsible for systematically auditing the AI systems used in the state.

2.5 Transparency requirements in ITDL

In Article 6 (1) the disclosure of the algorithm is required “unless the protection of personal data, other rights of third parties or public confidentiality interests are in conflict therewith”. The disclosure requirement refers to “algorithm” not to source code. The disclosure needs to be complemented by a description of the functionality in plain language and include the basic operation and decision logic of the algorithm. The legislative memorandum points out that the disclosure of the algorithm shall help to evaluate whether the used type of algorithm is generally appropriate for the task at hand and to check the results of the algorithm. The law therefore does not require source code availability in an open repository.

Article 6 (2) extends to the need for records of processing activities for data-driven technologies according to Article 30 GDPR [16] to those not using personal information. In Article 6 (3) it is prescribed that when automated forms of communication are used, the parties involved must be informed about this. In administrative decisions, the addressee needs to be informed about the use of data-driven information technology, specifically the used algorithm and the level of automation (Article 6(4)).

3 DISCUSSION OF THE LEGISLATION

In the following the legislation on the regulation of AI in ITDL (Schleswig-Holstein) will be discussed with regards to its definition of AI, handling of risk, fairness, accountability, transparency, accessibility, and process changes. The regulations will be compared to regulations in AIA and the Chinese regulation of recommender systems. This approach will be useful also for future reviews and comparisons of AI regulation initiatives.

3.1 Definition of Artificial Intelligence

The law in Schleswig-Holstein avoids using the term „artificial intelligence“. However, the explanatory memorandum to the law starts with a statement on artificial intelligence and its positive impact for the advancement of modern information technology. The law also includes a newly created legal remedy (Article 12) which is named „AI reprimand“(“KI-Rüge”). The law therefore clearly aims to regulate AI, even though it avoids using the legally undefined term.

The draft AIA [13], on the other hand, does not avoid using the term “artificial intelligence” but defines it as “software developed using (...) techniques listed in Annex I (...) that can generate results (...) for a specific human-defined objective that influence the environment with which they interact”. Annex I enumerates machine learning, logic and knowledge-based approaches, and statistical approaches, including Bayesian estimation, search, and optimization. This broad definition has been criticized in various ways [12], but its all-encompassing effect is countered by the distinction of risk levels with varying degrees of regulation [29]. The Chinese regulation on recommender systems in contrast only applies to internet

information services that use “technologies such as generation and synthesis, personalized push, sorting and selection, retrieval and filtering, and scheduling decision-making to provide information to users” [6]. It is therefore limited to recommender systems.

The definition of AI used in Article 1(3) ITDL has the advantage of being less comprehensive than the definition in the AIA draft. The definition nevertheless bears the risk of being interpreted too comprehensively. For example, database-driven applications that trigger actions based on past values are conceivable without them being machine learning (e.g., automatic reordering of consumables in hospitals when stock levels reach low levels). It is questionable if those applications should be regulated.

Even though the AI definition in the law appears useful in its context, it is not clear why Article 3(1) uses a second definition and separates it from chatbots. Although Article 3(1) explicitly distinguishes between data-driven information technologies and chatbots, chatbots are listed in the explanatory memorandum as examples of data-driven information technologies. The law specifically refers to chatbots as it is an AI technology that is already used in public service. The lawmaker aims to ensure that chatbots continue to be made available by administrations, but that citizens always have an alternative form of communication and that it is clear to citizens that they are communicating with a machine (explanatory memorandum). Chatbots are often used as “intelligent forms” that query the appropriate data for an administrative process and hide unnecessary input fields. A public sector application of chatbots as intelligent forms is the U:DO chatbot, a private initiative to enable easy application for seasonal short-time allowance.¹ On the other hand, chatbots can function as a list of frequently asked questions, in which they offer the most suitable answers possible to questions posed in natural language. A public sector application of an FAQ-type chatbot is the chatbot “Bobbi” provided on the official website of the federal state of Berlin.² The use of chatbots as user interfaces appears to be unproblematic if conventional user interfaces (web pages, forms) are also offered in parallel and if the interface technology can be clearly recognized as a bot. While data-based language models are prone to discrimination (e.g., they might be harder to understand for people with accents or female voices), as an additional user interface they also improve accessibility by making offerings accessible to people with lower mobility, dyslexia, or other limitations. The use of AI techniques as additional user interfaces should therefore not be regulated, but only observed. Natural language processing technologies, especially in recognizing spoken language, should be tested on speech samples from linguistic minorities (with dialect/accents, very old/very young voices).

ITDL aims to enable the use of AI in public service and create a reliable legal environment for this purpose. However, it is likely that applications that meet the definition of Article 1 ITDL are already in use within the scope of the law. Possible examples of technologies that are probably already used and would be regulated could include:

- the functionality to automatically hide or replace the background in videoconference systems,

- the learning analytics functionality available in the learning management system Moodle (Moodle is used in numerous universities, schools, and other educational institutions; however, its learning analytics functionality might not be broadly used),
- automatic document recognition (e.g. on photos). This feature is implemented for example in the state app for health insurance claims,
- OCR (character recognition on photos/scans) for processing receipts e.g. in financial or social administrations,
- risk-dependent checking/sample selection, e.g. for fraud detection in financial authorities,
- the use of common technologies such as route planners, search engines,
- forecast models, e.g. for the weather,
- chatbots that may already exist.

The legislator should therefore consider cataloging and assessing the risks of existing technologies and defining a transition period for this. Many data driven applications within the definition of the law could be considered uncritical, if they don't create the risk of inconsistent administrative action (i.e. the functionality to blur the background in videoconferences or to automatically reorder office supplies). The legislator could consider narrowing Article 1 (1) further on a risk basis and, for example, exclude technologies at the user interface, such as chatbots, if alternatives are available for citizens.

Article 11 (2) lists a number of exemplary categories of data-driven information technologies and thus delimits speech recognition, text classification, pattern recognition, data analysis, and AI applications. Usually, however, the first three items are understood as subfields of AI. Data analytics, on the other hand, is a very broad term that can go beyond the Article 1(3) definition.

3.2 Risk-based approach

The draft law contains frequently cited elements to mitigate AI risks (transparency, primacy of human decision, non-discrimination). The areas of application prohibited under Article 2(2) overlap with practices prohibited under AIA, specifically the “mass identification of persons (. . .) on the basis of biometric characteristics”, personality assessments, and delinquency forecasts. Hereby the law reflects the societal discussion about AI risks, especially discrimination in applications in human resources, crime prevention, and predictive policing [23].

Items 1 (exercise of direct force), and 4 (administrative acts with discretion or scope for assessment) in Article 2(2) go beyond the prohibition of use under the AIA (Annex III) but relate to the narrower definition of AI in the ITEG and to the limited scope of use (public administration Schleswig-Holstein). The Chinese administrative rule on recommender systems prohibits its use for activities prohibited by laws such as endangering national security or disrupting social order (Article 7) nor to manipulate online public opinion (Article 14).

The stipulation that the use of AI is only possible where there is no room for discretion or judgment (Article 2 (2) point 4) clearly limits the possible uses of AI.

¹<https://kurzarbeit-einfach.de/>

²<https://service.berlin.de/chatbot/chatbot-bobbi-606279.php>

The draft of the AIA only regulates high-risk AI systems through its risk-based approach, for which a conformity assessment (audit) or post-market monitoring is required. Comparable to the conformity assessments according to the AIA is the audit criteria referred to under Article 11 (1) points 2 and 5. However, the necessary standardization processes, e.g. by German standard setter DIN, international standard setter ISO, or professional organization IEEE, have not yet been completed. ITDL recognizes this lack of audit standards and empowers a state authority to define audit requirements for data quality and proper operation (Article 11 (1)).

3.3 Fairness

Even though the law aims to ensure the fair application of AI systems and prohibits their use in sensitive areas of applications (personality assessments, delinquency prediction), it does not include specific requirements for ensuring fairness. Such requirements could include the training on AI fairness for individuals involved in the development or selection of AI systems or the obligation to ensure diversity with regards to gender, age groups, and other minorities in the conception, development, or testing of the systems.

The AIA on the other hand specifically requires that training, validation, and testing data sets used for high-risk applications shall be relevant, representative, that they “shall have the appropriate statistical properties” and be examined for possible biases (Article 10). It also requires providers of high-risk systems that continuously learn to watch out for biased outputs and unfavorable feedback loops (Article 15).

The Chinese administrative rule on internet recommender systems does not put specific emphasis on fairness but just generally requires the systems to follow, among others, the “principle of fairness” (Article 4).

The use of AI technologies can improve the accessibility of public administration services and thus improve fairness. Examples include the integration of multilingualism, e.g., through the automated translation of information offerings, read-aloud function for websites, or the automatic generation of subtitles in videos/video conferences. Involving disability advocates in exploring meaningful uses for AI could therefore be recommended.

3.4 Accountability

The clear definition of accountability in ITDL is to be welcomed. However, in contrast to the European AI Act and Chinese administrative rule, IT deployment law does not provide for a supervisory authority for data-driven information technologies, in this respect, it falls short of the requirements of the European draft.

A new and separate legal remedy is created for decisions made by data-based information technology (“KI-Rüge”). This could contribute to preserving the principle of accountability and the principle of controllability of automated decisions.

The Chinese administrative rule on recommender systems clearly defines which authorities are responsible for the supervision of legal compliance (Article 3) and that providers of recommendation services are responsible to establish management systems and technical measures for auditing, reviewing, and monitoring the compliance of the system (Article 7).

3.5 Transparency

The transparency requirement in Article 6 (1) is very broad on the one hand, requiring the publication of the used algorithm. If the transparency requirement is not met, the administrative decision becomes null and void. On the other hand, the regulation allows for broad exceptions to the transparency requirement (in case of protection of personal data, rights of third parties, public confidentiality interests). Also, there is no requirement to publish or explain the training data used. However, transparency about the data used is necessary in some cases to understand the system.

The currently regulation risks leading to bureaucratization if an inventory of relevant systems in public service needs to be taken and transparency for every system needs to be documented. This could be avoided by applying the transparency requirement conditionally, e.g., if the system is used to prepare administrative acts.

3.6 Consequences for public servants

In the written expert consultation on ITDL, trade unions expressed their concern that the use of AI will lead to additional burdens for employees, e.g. in the context of system introduction processes [22]. At the same time, there are fears that jobs will be graded lower, that employees will not have the necessary qualifications and that experts will not be consulted to a sufficient extent due to the shortage of specialists. In the view of the trade unions, the use of AI systems must not lead to automated work monitoring, the collection of data on employees, or to the further dissolution of the boundaries of work. The unions also emphasize that employees should be closely involved in co-determination processes and that far-reaching works council agreements should be concluded. The union expects that AI systems could put employees under pressure to accept the system’s decision-making proposals unchanged; the employer must therefore assume responsibility for corresponding administrative acts [9].

4 CONCLUSION

In this paper, a framework for discussing AI regulation was proposed and applied to the recently enacted German federal regulation of AI in the public sector in Schleswig-Holstein IT Deployment Law ITDL. As regulation needs to be limited in scope, in a first step the applied definition of AI was reviewed and compared to definitions in the proposed European AI Act. The ITDL provides a useful definition of AI. In a next step, it was discussed how the proposal addresses risks by prohibiting its application in certain areas such as predictive policing and requiring authorities to assign automation levels for automated decision-making processes. Further, ITDL regulations on fairness, accountability, and transparency of AI systems in the public sector were reviewed.

Due to the proven discrimination potential of machine learning systems, regulation in ethically critical areas of application (human resources, education system, decisions on social benefits, in the penal system or judiciary) is desirable in principle.

In Germany and in Europe in general, AI regulation and standardization are considered as means to strengthen joint European values and improve competitiveness [14]. The AI regulation in Schleswig-Holstein on the use of AI in the public sector reflects this.

It aims to both enable and constrain the use of AI technologies in the public sector, thus digitalizing and modernizing public service, while providing discrimination-free, fair, and transparent services to citizens. The law reflects the political strategy “mythology” that of AI as a “benevolent force of national progress” that will contribute to both economic success and public welfare if only tamed well enough [26].

ACKNOWLEDGMENTS

This research has been funded by the Federal Ministry of Education and Research of Germany under grant number 16DHB4002.

REFERENCES

- [1] Algorithmwatch. AI Ethics Guidelines Global Inventory. Retrieved 13.11.21 from <https://inventory.algorithmwatch.org/>.
- [2] Julia Angwin and Terry Parris Jr. 2016. Facebook lets advertisers exclude users by race. ProPublica blog 28.
- [3] Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings.
- [4] CEN-CENELEC. CEN-CENELECT Topics: Artificial Intelligence. Retrieved January 21, 2022 from <https://www.cenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/>.
- [5] Central Digital and Data Office. 2021. <https://www.gov.uk/government/collections/algorithmic-transparency-standard> (2021). Retrieved 11.1.22 from <https://www.gov.uk/government/collections/algorithmic-transparency-standard>.
- [6] China. 2022. Internet Information Service Algorithm Recommendation Management Regulations (2022). Retrieved January 19, 2022 (via Google Translate) from http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm.
- [7] Cyberspace Administration of China. 2022. The State Internet Information Office and other four departments issued the “Internet Information Service Algorithm Recommendation Management Regulations” (2022). Retrieved January 19, 2022 (via Google Translate) from http://www.cac.gov.cn/2022-01/04/c_1642894606258238.htm.
- [8] Department for Digital, Culture, Media & Sport. 2020. National Data Strategy (2020). Retrieved from <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>.
- [9] DGB Bezirk Nord. Stellungnahme zum Gesetzentwurf zur Förderung der Digitalisierung und Bereitstellung von offenen Daten und zur Ermöglichung des Einsatzes von datengetriebenen Informationstechnologien in der Verwaltung (Digitalisierungsgesetz). Umdruck 19/6843. Retrieved 12.1.22 from <http://www.landtag.ltsh.de/infothek/wahl19/umdrucke/06800/umdruck-19-06843.pdf>.
- [10] DIN. Normungroadmap Künstliche Intelligenz Ausgabe 2. Retrieved January 21, 2022 from <https://din.one/pages/viewpage.action?pageId=33620030>.
- [11] DIN. 2020. Deutsche Normierungroadmap Künstliche Intelligenz (November 2020). Retrieved January 21, 2022 from <https://www.din.de/resource/blob/772438/6b5ac6680543eff9fe372603514be3e6/normungroadmap-ki-data.pdf>.
- [12] Martin Ebers, Veronica R. S. Hoch, Frank Rosenkranz, Hannah Ruschemeier, and Björn Steinrötter. 2021. The European Commission’s Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS). J 4, 4, 589–603. DOI: <https://doi.org/10.3390/j4040043>.
- [13] EC. 2021. Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts. AIA.
- [14] European Commission. Preliminary draft Annual Union Work Programme for standardization 2022. Retrieved January 21, 2022 from <https://www.agoria.be/en/legislation/standardisation/ec-annual-union-work-programme-2022>.
- [15] European Commission. 2020. WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust (February 2020). Retrieved January 21, 2022 from https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.
- [16] European Union. General Data Protection Regulation. GDPR.
- [17] David Gunning, Mark Stefik, Jaesik Choi, Timothy Miller, Simone Stumpf, and Guang-Zhong Yang. 2019. XAI-Explainable artificial intelligence. Science robotics 4, 37. DOI: <https://doi.org/10.1126/scirobotics.aay7120>.
- [18] Arendse Huld. 2022. China Passes Sweeping Recommendation Algorithm Regulations (2022). Retrieved January 11, 2022 from <https://www.china-briefing.com/news/china-passes-sweeping-recommendation-algorithm-regulations/>.
- [19] IEEE. Artificial Intelligence Systems (AIS) Related Standards. IEEE portfolio of AIS technology and impact standards and standards projects. Retrieved from <https://standards.ieee.org/initiatives/artificial-intelligence-systems/standards.html>.
- [20] IEEE. CertifAIED. Retrieved February 21, 2022 from <https://engagestandards.ieee.org/ieeecertifaiied.html>.
- [21] ISO. ISO/IEC JTC 1/SC 42 Artificial intelligence. Retrieved February 21, 2022 from <https://www.iso.org/committee/6794475/x/catalogue/p/1/u/1/w/0/d/0>.
- [22] komba-gewerkschaft schleswig-holstein. 2021. Stellungnahme zum Gesetzentwurf zur Förderung der Digitalisierung und Bereitstellung von offenen Daten und zur Ermöglichung des Einsatzes von datengetriebenen Informationstechnologien in der Verwaltung (Digitalisierungsgesetz). Umdruck 19/6835 (2021). Retrieved 12.1.22 from <http://www.landtag.ltsh.de/infothek/wahl19/umdrucke/06800/umdruck-19-06835.pdf>.
- [23] Lea Köstler and Ringo Ossewaarde. 2021. The making of AI society: AI futures frames in German political and media discourses. AI & Soc 62, 6, 1930. DOI: <https://doi.org/10.1007/s00146-021-01161-9>.
- [24] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. ACM Comput. Surv. 54, 6, 1–35. DOI: <https://doi.org/10.1145/3457607>.
- [25] OECD. 2021. Government at a Glance 2021. OECD.
- [26] Marinus Ossewaarde and Erdener Gulenc. 2020. National Varieties of Artificial Intelligence Discourses: Myth, Utopianism, and Solutionism in West European Policy Expectations. Computer 53, 11, 53–61. DOI: <https://doi.org/10.1109/MC.2020.2992290>.
- [27] Dietmar Schabus. The IEEE CertifAIED Framework for AI Ethics Applied to the City of Vienna. Retrieved February 21, 2022 from <https://beyondstandards.ieee.org/the-ieee-certifaiied-framework-for-ai-ethics-applied-to-the-city-of-vienna/>.
- [28] Schleswig-Holsteinischer Landtag. 2021. Entwurf eines Gesetzes zur Förderung der Digitalisierung und Bereitstellung von offenen Daten und zur Ermöglichung des Einsatzes von datengetriebenen Informationstechnologien in der Verwaltung (Digitalisierungsgesetz). ITEG.
- [29] Sebastian F. Schwemer, Letizia Tomada, and Tommaso Pasini. 2021. Legal AI Systems in the EU’s proposed Artificial Intelligence Act. In Joint Proceedings of the Workshops on Automated Semantic Analysis of Information in Legal Text (ASAIL & LegalAIIA 2021).
- [30] Katharina Simbeck. 2019. HR Analytics and Ethics. IBM J. Res. & Dev., 1. DOI: <https://doi.org/10.1147/JRD.2019.2915067>.
- [31] UNO. 2020. United Nations E-Government Survey 2020. Digital Government in the Decade for Sustainable Development (2020). Retrieved 21.1.22 from <https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2020.</bib>>
- [32] Jakob Mökander, Maria Axente, Federico Casolari, and Luciano Floridi. 2021. Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation. Minds & Machines (2021). <https://doi.org/10.1007/s11023-021-09577-4>.
- [33] Thilo Hagendorff. 2020. The Ethics of AI Ethics: An Evaluation of Guidelines. Minds & Machines 30, 99–120 (2020). <https://doi.org/10.1007/s11023-020-09517-8</bib>>
- [34] Schleswig-Holsteinischer Landtag. 2022. Gesetz über die Möglichkeit des Einsatzes von datengetriebenen Informationstechnologien bei öffentlich-rechtlicher Verwaltungstätigkeit (IT-Einsatz-Gesetz – ITEG). Retrieved 6.5.2022 from https://www.schleswig-holstein.de/DE/Landesregierung/IV/Service/GVOBl/GVOBl/2022/gvobl_5_2022.pdf?__blob=publicationFile&v=2