Universität Münster

UNIVERSITY OF TWENTE.

# The AI-sylum Process

## Ethical Implications of AI-Implementation in German Asylum Procedures

Public Governance across Borders (B.Sc.)

**Bachelor Thesis**

By **Omnia Al-Hadari**

Supervisors:

**1st Dr. Su Yun Woo**

**2nd Dr. Veronica Junjan**

Date: 28.02.2025

Word Count: 10.975

# Abstract

This study explores the ethical implications of artificial intelligence (AI) in Germany's asylum procedures, focusing on its impact on public administration and asylum seekers' rights. Using a qualitative research design centred on textual analysis, the study examines AI tools adopted by the German Federal Office for Migration and Refugees (BAMF) since 2015. The research applies the Ethics Guidelines for Trustworthy AI by the EU's High-Level Expert Group on Artificial Intelligence (AI HLEG), focusing on four ethical principles—respect for human autonomy, prevention of harm, fairness, and explicability—operationalised through the Assessment List for Trustworthy AI (ALTAI). Findings highlight both the benefits of administrative efficiency and the risks of bias, lack of transparency, and potential rights violations. The study identifies critical ethical challenges, particularly regarding vulnerable groups such as asylum seekers. Recommendations are provided for policymakers to ensure alignment of AI implementation with ethical and human rights standards. This research contributes to understanding the intersection of AI, ethics, and public administration in sensitive, high-risk contexts, offering insights for improved governance and ethical compliance in the asylum process.



**Keywords**: Artificial Intelligence (AI), Public Administration (PA), Asylum, Migration, Germany, Ethics of AI, Human Rights

# Table of Contents

# 1. Introduction

The integration of artificial intelligence (AI) in public administration (PA) has gained significant attention in recent years (Madan & Ashok, 2023; Kossakowski, 2024). This growing trend is driven by the promise of increased efficiency, reduced workload, and faster decision-making in various bureaucratic processes (Wirtz et al., 2019). However, this is also backed by research that AI has the potential to minimise bureaucratic inefficiencies and improve service delivery across various sectors (Reis et al., 2020; Fedyk et al., 2022). As public institutions face rising demands, especially in areas like Germany's migration management, AI systems are increasingly seen as a solution to overburdened civil servants and public administration in general (Dataport, n.d.; BMI, n.d.-a).

Additionally, AI is regarded to bring forth ethical benefits that are both implied and explicitly addressed. Implied benefits include individual flourishing, collective well-being, wealth maximisation and value creation through increased efficiency (Stahl et al., 2023). Some specific ethical benefits of AI include its potential to alleviate humans from hazardous and unpleasant work; advance scientific research benefitting health, even provide better access to healthcare for the visually impaired; optimise agricultural practices to increase food accessibility; improve transportation and logistics etc. (Stahl et al., 2023).

Nonetheless, this optimistic outlook comes with challenges that require careful consideration. Discussions surrounding AI ethics have consistently highlighted concerns about poorly designed AI systems, their susceptibility to biases, the appropriateness of using AI in sensitive domains involving personal data, effects on human rights and many more (Wirtz et al., 2019). These challenges are particularly consequential in public administration, in which decisions profoundly impact individual lives, particularly in the context of asylum procedures. Specifically, concerns regarding transparency, fairness, and accountability arise when AI systems are involved in decisions affecting individuals' rights and legal status (Forster, 2022).

This is reflected in the recently enacted EU AI Act, a major regulator's first comprehensive regulation of AI (AI Act, n.d.). The EU AI Act categorises AI systems into four risk categories: AI systems classified as ´unacceptable risk´, are banned; those classified as ´high risk´ are subject to specific legal requirements; ´limited risk´ applications associated with a lack of transparency and hence have specific transparency obligations and ´minimal or no risk´ are largely unregulated and constitute the majority of AI systems used in the EU (AI Act, n.d.; European Commission, 2024). However, AI applications in migration, asylum and border management are classified as ´high-risk´ (Ozkul, 2023; European Commission, 2024).

Despite the associated risks, the integration of AI within German asylum procedures continues and increasingly influences asylum decision-making processes, with the Federal Office for Migration and Refugees (short: BAMF), being at the forefront (BAMF, n.d.). Discussions on AI implementation in PA mainly focus on its potential to enhance overall process efficiency and reduce the workload on public administrators. Yet its impact on the rights of asylum seekers is rarely highlighted by policymakers (i.e., BAMF, n.d.). AI tools significantly influence the rights of asylum seekers, which are often undermined by legal exceptions and rationalisations that

treat their rights as optional (Twigt, 2023). Even research examining the ethical challenges associated with the use of AI in German asylum procedures and their implications for public administration and policymakers in Germany remains relatively scarce.

This thesis therefore attempts to address this gap by offering an analysis of these ethical challenges related to AI implementation in asylum procedures, particularly within the context of German public administration. It will outline current challenges, propose recommendations for policymakers and contribute to the broader discourse on governing AI in sensitive or so-called ´high-risk´ areas. This research intends to offer both theoretical insights and practical solutions for balancing the benefits and ethical risks of AI adoption.

The aim of this thesis is therefore to answer the following research question:

***What are the ethical implications of the procedural implementation of AI in the asylum process for Germany's public administration (2015–beyond)?***

This explanatory research question seeks to explain how the use of AI in asylum procedures affects the ethical dimensions of Germany's public administration. It investigates the intersection between AI technologies, PA and ethics within a particularly sensitive domain, focusing also on how the introduction of AI impacts the rights of asylum seekers. This timeframe was selected to focus the analysis on AI tools implemented since the onset of the refugee crisis in 2015 in Germany. During this period, Germany experienced a substantial influx of refugees and asylum seekers (Herbert & Schönhagen, 2020). In response, the BAMF accelerated the development and implementation of digital tools, including AI, to enhance the efficiency of asylum procedures (Biselli, 2024; Migration Data Pool, 2022). This process is ongoing, with Chancellor Olaf Scholz emphasising the future role of AI in the decision-making of asylum applications (Kossakowski, 2024).

To answer the research question effectively, the study first employs elements of descriptive research to delineate the current state of AI implementation and its tools in the German asylum process. This will be answered through the following sub-question (see Section 4):

1. *How has AI been procedurally implemented in Germany's asylum process since 2015?*

Secondly, is to explore the broader implications, through the investigation of challenges emerging from AI tools used in asylum processes and what they imply for policymakers (Sections 4 and 5). Hence, the formulation of the last two sub-questions:

2. *What are the main ethical challenges when implementing AI in the asylum process?*

The ethical challenges of the collected AI tools will be identified and analysed using the EUs Ethics Guidelines of Trustworthy AI as a guide. This framework has yet to be integrated into research discussing AI in PA and asylum procedures, which provides another gap in the current research landscape. The connection between this framework, AI and ethics will be discussed in the theory section of this paper (see Section 2). The objective of this analysis is not to conduct an in-depth evaluation of each tool but to identify the main challenges across the board. This identification will offer insights essential for developing policy recommendations by the end of the analysis (see Section 5). And lastly:

3. *What recommendations can be given to German policymakers regarding the implementation of AI tools in the asylum process to ensure the human rights of asylum seekers and an ethical PA procedure in general?*

The social relevance of this study lies in the profound impact that AI tools can have on vulnerable populations, particularly asylum seekers (Forster, 2022; European Commission, n.d.; European Commission, 2019; Bakiner, 2023). The use of AI in asylum procedures can influence critical decisions about people's rights and legal statuses, making it essential to ensure that these systems are designed and implemented ethically and transparently. Given the high stakes for individuals involved and the increased interest in AI within this context, understanding and addressing the ethical challenges is not only a matter of academic interest but also of societal responsibility.

From a scientific perspective, this research contributes to the growing body of knowledge on AI in public administration by combining insights from AI, ethics and asylum/migration. While much has been written about AI in public administration and ethical concerns in some high-risk areas, there is a lack of studies that investigate those concerns within Germany's asylum procedures whilst testing the practicality of one of the many assessment frameworks that seek to ensure the ethics of AI implementation (like the EU Ethics Guidelines of Trustworthy AI). Additionally, there is still the need for general AI principles that safeguard from negative human rights consequences, especially on the national and regional level that are tailored to the asylum context (Forster, 2022). By focusing on this intersection, the study adds to the scientific discourse on the governance of AI systems in sensitive areas, providing both theoretical insights **and** practical recommendations.

# 2. Theory

In this section, specific elements and phrasings of the research question and theoretical concepts employed in this research will be further clarified to establish a foundation for the analysis.

## 2.1. Definition of AI

The ethical challenges in artificial intelligence already begin with the complexities surrounding its definition, or to be specific, the lack thereof.

The term "Artificial Intelligence" (AI) has been widely explored over several decades, yet no universally accepted definition has been agreed upon (Wirtz et al., 2019; Stahl et al., 2023). This definitional ambiguity already poses significant challenges for addressing ethical dilemmas in AI, as the discourse on ethics is inherently linked to the foundational understanding of the concept itself.

The fundamental problem is that AI is yet to be understood in its entirety and no definition seems to fully capture all aspects relevant to it (Wirtz et al, 2019; Stahl et al., 2023). This

ununified view on AI hinders the clear identification of ethical as well as human rights-based problems that emerge from its implementation, making it even more difficult to solve them (Stahl et al, 2023). It is mentioned that the "(…) uncertainty of the term AI poses a problem for the academic discourse as well as practical and policy interventions." (Stahl et al., 2023, p. 2).

To facilitate this research, a broader conceptualisation of artificial intelligence (AI) is nevertheless necessary. Wirtz et al. (2019) have extracted key aspects of AI that have been commonly highlighted across various definitions. According to them "(…) AI refers to the capability of a computer system to show human-like intelligent behavior characterized by certain core competencies, including perception, understanding, action, and learning." and "(…) AI application refers to the integration of AI technology into a computer application field with human computer interaction and data interaction." (Wirtz et al, 2019, p. 599).

AI generally can be divided into three forms, namely ANI, AGI and ASI (Wirtz et al, 2019). Artificial narrow intelligence (ANI) is designed and trained by humans to perform a specific task, this is also called Narrow AI or Weak AI (Wirtz et al., 2019; Emmert-Streib, 2024; Siao & Wang, 2020). Artificial general intelligence (AGI) possesses additionally the ability to learn on its own and transmit its results and skills to other tasks without human assistance, thus human-level intelligence (Wirtz et al., 2019; Emmert-Streib, 2024). Lastly, artificial super intelligence (ASI) describes software that even exceeds the human mind (Wirtz et al., 2019). Currently, only ANI technologies have been developed and many are in the developing stages between ANI and AGI (Wirtz et al., 2019; Abonamah et al., 2021). AI applications within asylum procedures are hence still considered narrow AI, as they are usually designed to perform specific tasks.

Artificial Intelligence (AI) systems are designed to solve complex problems and perform tasks in the digital or physical world. They process and analyse data to act intelligently, with some systems interacting physically with their environment, such as robots (embodied AI) (Deutscher Bundestag, 2020). AI can be classified into rule-based systems, which follow predefined rules created by experts, and learning systems, which use data to improve their performance through training (Deutscher Bundestag, 2020). Machine Learning (ML) is central to learning systems, enabling them to adapt and refine their behaviour. A specific type of ML, Deep Learning (DL), utilises layered neural networks to process large amounts of data with minimal human intervention, often outperforming traditional systems but being harder to understand (Deutscher Bundestag, 2020).

Rule-based systems use straightforward if-then logic for decision-making, requiring clear rules and facts, which can be difficult due to language ambiguity (Csuk, 2025). While they yield explainable results, they struggle to adapt to changes. Machine learning systems, on the other hand, can automatically identify patterns in large datasets, allowing for continuous learning and efficient information processing (Csuk, 2025). Deep learning, a subset of machine learning, utilizes artificial neural networks to model the human brain, enabling even greater accuracy and data handling. However, maintaining data quality and avoiding biases is crucial for reliable outcomes in machine learning, which often functions as a "black box," making decision reasoning difficult to understand (Csuk, 2025). Efforts in Explainable AI aim to improve

transparency and clarity in these systems, helping users comprehend decision-making processes, though full transparency remains a challenge (Csuk, 2025).

Training AI involves three primary methods: supervised learning, where data is labeled to guide the system; unsupervised learning, which identifies patterns in unlabeled data; and reinforcement learning, where the system learns through trial and error to achieve a set goal (Deutscher Bundestag, 2020). The efficacy of these approaches is contingent upon the quality and quantity of data, which must be detailed, accurate and correctly labelled (Deutscher Bundestag, 2020). Data sources could include raw data (direct observations), secondary data (processed from raw data), and synthetic data (artificially generated to simulate real-world scenarios) (Deutscher Bundestag, 2020). The quality and types of data employed in training AI systems are relevant when addressing the ethical challenges of AI in subsequent sections of this paper.

This conceptual understanding of AI will form the basis for examining AI adoption in the asylum process. However, it is essential to keep the ethical problems posed by the definitional ambiguity in mind and that AI encompasses more than what was presented in this section.

## 2.2. Understanding Ethics

Ethics is a broad and complex concept, which makes it necessary to provide a foundational understanding of it and how it relates to AI to understand the points put forth in this paper. Simply described, ethics can be understood in three interconnected forms: as an idea, as a discipline, and in practice and it can be broadly divided into normative, meta and applied ethics (Dimmock & Fisher, 2017).

Ethics as an idea is concerned with the way the world ought to be and can be defined as "(…) moral principles governing the behaviors or actions of an individual or a group of individuals (…)." (Siao & Wang, 2020, p. 75). To offer justification for those principles and beliefs, we enter the world of ethics as "(…) a philosophical discipline that explores what counts as right or wrong, good or bad and on what grounds such judgements are made." (Stahl et al., 2023, p. 2).

Normative ethics focuses on creating theories guiding moral judgements concerning our behaviour, such as virtue theory, deontology, consequentialism/utilitarianism or even rights-based ethics (Dimmock & Fisher, 2017; LaFolette, 2014). Metaethics is concerned with how we engage in ethics. In contrast, applied ethics is ethics in practice, meaning it is adapted to specific contexts to answer certain moral questions (i.e. if euthanasia or meat-eating are morally right or wrong) (Dimmock & Fisher, 2017). Ethics is applied in many areas creating sub-disciplines of applied ethics, such as neuroethics and biomedical ethics (Stahl et al, 2023). Applied to AI, ethics of technology, computer ethics, information ethics but also robot ethics (roboethics) and machine ethics are relevant (Stahl et al., 2023; Siao & Wang, 2020). The ethics of AI, however, is still an emerging field that explores ethical principles, guidelines, and regulations related to AI, along with the moral obligations of AI systems and their developers (Siao & Wang, 2020).

When analysing and assessing the ethics of AI and its applications, various principles and frameworks are designed and presented (Fjeld et al., 2020; Jones, 2023; Stahl et al., 2023; Siao & Wang, 2020). Many of these frameworks are grounded in or inspired by human rights, which is why they are frequently discussed together in research (Fjeld et al., 2020; Stahl et al., 2023). Jones (2023) argues that while ethics and human rights are distinct fields, their integration is vital for effective AI governance, as they complement each other. Therefore, discussing them in tandem is essential.

Additionally, ethics, as a branch of philosophy, does not provide a universal set of norms; instead, it allows for multiple interpretations, which can often be ambiguous (Jones, 2023). In contrast, human rights establish international norms derived from ethical principles and serve as a "crystallisation" of ethical values, helping to avoid the criticisms typically directed at ethics or at other codes of AI ethics that lack a firm grounding (Jones, 2023; Yeung et al., 2020). Given that AI also significantly affects human lives and experiences, it is crucial to evaluate it considering these rights as well (Jones, 2023)."

## 2.3. Understanding Human Rights

Human rights are fundamental rights that apply to all individuals, regardless of race, sex, nationality, ethnicity, language, religion, or other status Polok et al., 2023). They are inherent, inalienable, interdependent, and universal, with non-derogable rights such as the right to life and the prohibition of torture, which cannot be limited (Polok et al., 2023). These rights aim to uphold human dignity, allowing individuals to live freely and independently while being protected from fear and want. States are responsible for safeguarding these rights and must implement legal protections against misuse (Polok et al., 2023). Ultimately, these rights are granted to every individual by virtue of their inherent humanity (Krennerich, 2024).

The Universal Declaration of Human Rights (UDHR), adopted in 1948, is a pivotal document that outlines these rights, albeit as a non-binding instrument (United Nations, n.d.; Polok et al., 2023). It was significant for establishing the individual right to asylum for the first time (Oltmer, 2017). Following this, the 1951 Convention Relating to the Status of Refugees (Refugee Convention) was created to define the rights and obligations of states concerning refugees fleeing violence (Oltmer, 2017; UNHCR, n.d.). Initially addressing post-World War II conditions in Europe, the scope of the Convention was expanded in 1967 to address global refugee movements (Oltmer, 2017; UNHCR, n.d.).

In 1966, two important covenants were established: the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social, and Cultural Rights (ICESCR). Together with the UDHR, these documents form what is known as the "International Bill of Human Rights" (Polok et al., 2023). They transcribe the rights identified in the 1948 declaration, with some exceptions (Krennerich, 2024). In Europe, the European Convention on Human Rights (ECHR) established by the Council of Europe, signed in 1950 and effective from 1953, became the first instrument to legally bind certain rights from the UDHR (ECHR, n.d.).

There is also the EU Charter of Fundamental Rights and Germany's constitution (*Grundgesetz*) which exemplifies a rare instance where national regulation incorporates the right to asylum found in the UDHR (Oltmer, 2017, UNHR, n.d.). It is therefore essential for every member state of the Council of Europe and the EU, of which Germany is both a part of, to develop human rights-compliant asylum procedures as highlighted in the report of the special rapporteur and Committee on Migration, Refugees and Displaced Persons in the Parliamentary Assembly of the Council of Europe (2024).

## 2.4. Ethics Guidelines for Trustworthy AI

Numerous models and guidelines have been created to address ethical concerns in AI across various sectors, with over 160 identified by Algorithm Watch in 2019. Critics argue that private sector models may act as "ethics washing," serving to evade legal regulation or functioning as public relations tools, making them unsuitable for policy analysis (Molnar, 2021; Radu, 2021; Algorithm Watch, 2020; Jones, 2023). In contrast, government agencies and international organizations develop similar guidelines emphasising human agency, data protection, and accountability (OECD, n.d.; European Commission, 2019; UNESCO, 2023; BAMS, 2022). This research focuses on models that adopt a human rights-based approach, highlighting the distinction between ethics and law.

The EU has been in the forefront globally in regulating technology including AI, through for instance the GDPR and the AI Act (Bakiner, 2023; AI Act, n.d.). The AI Act was built upon the work of the Commission's High-Level Expert Group on Artificial Intelligence (AI HLEG), which is an expert group appointed by the European Commission to advise the Commission on its AI strategy in 2018 (European Commission, n.d.; AI HLEG, 2019). This group was tasked to present two deliverables: firstly, ethics guidelines on AI and secondly policy and investment recommendations (Hickman & Petrin, 2021). In 2019 the expert group published the ´Ethics Guidelines for Trustworthy Artificial Intelligence´ based on a draft document published by the AI HLEG and feedback from public consultation from 500 contributors (Hickman & Petrin, 2021; AI HLEG, 2019).

Given Germany's position within the EU and the specific emphasis of these EU Guidelines on ethics, developed by the same expert group that contributed to the AI Act, it seemed rather plausible to examine this framework more closely.

The ´Ethics Guidelines for Trustworthy AI´ outline a comprehensive framework based on three pillars that AI systems must be lawful, ethical, and robust to be considered trustworthy (AI HLEG, 2019). Specifically, Trustworthy AI must comply with all applicable laws (lawful), adhere to ethical principles and values (ethical), and ensure robustness both technically and socially. While the framework acknowledges the importance of the lawful component and its overlap with the ethical, it is not the focus, nor is it evaluated with these principles (see Figure 1). The focus of these guidelines is the ethical and robust dimensions, which are analysed by four ethical principles, namely: respect for human autonomy, prevention of harm, fairness, and explicability.

From these four ethical principles, seven key requirements (often referred to as principles, though distinct from the four core ethical principles) are derived. These requirements guide the operationalisation of ethical AI systems and are built on fundamental rights as outlined in the Charter of Fundamental Rights of the European Union (EU Charter) and relevant international human rights law. Importantly, the ethical principles are presented without a hierarchical order, highlighting that each principle holds equal significance and that the seven principles must be upheld throughout the entire lifecycle of AI systems (AI HLEG, 2019).

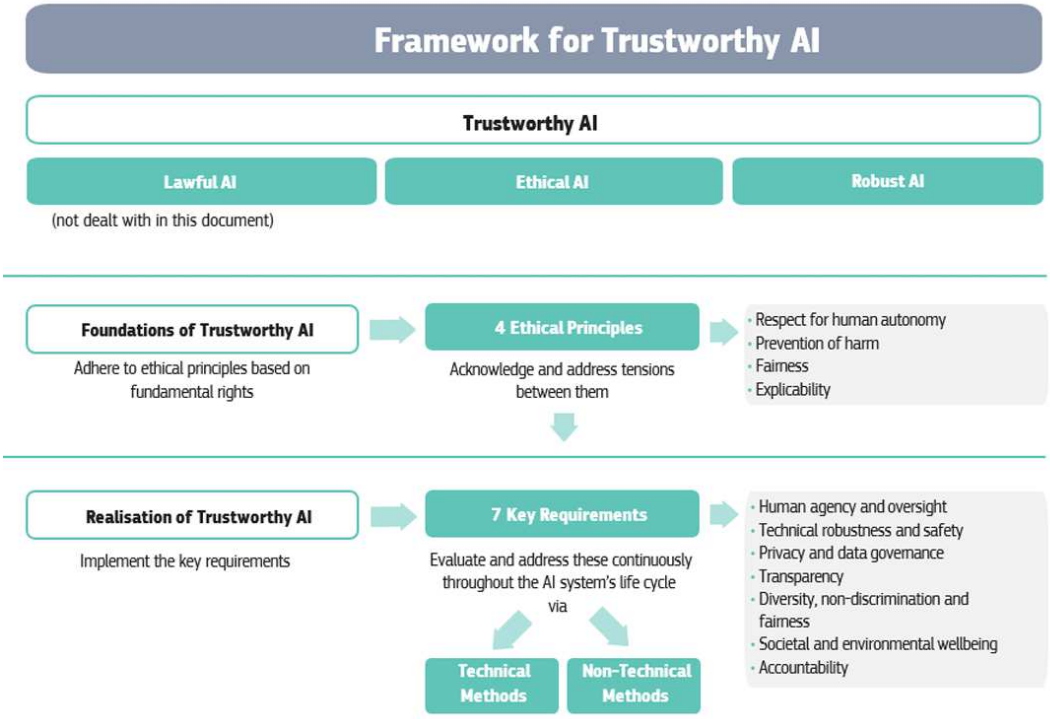


*Figure 1: Framework of Trustworthy AI (European Commission, 2019)*

This framework adopts a clear human-rights-based approach, emphasising the distinction between legal compliance and ethical adherence. It is mentioned that Trustworthy AI must comply with the law, but since legal frameworks often struggle to keep pace with technological advancements, it is essential for AI systems to also align with ethical norms (AI HLEG, 2019). While many legal obligations are rooted in human rights, ethical principles extend beyond mere legal compliance to ensure broader protection and fairness.

It is also mentioned that special attention must be directed towards vulnerable groups and AI systems with significant impacts on individuals, particularly in situations of power or information asymmetries, which is the case with asylum seekers and government authorities (AI HLEG, 2019).

These seven principles are each evaluated and further broken down in the Assessment List presented by AI HLEG. The Assessment List for Trustworthy AI (ALTAI), introduced by the AI HLEG in 2020, is a self-assessment tool designed to accompany the Ethics Guidelines for Trustworthy AI. It provides a detailed breakdown of the seven principles for trustworthy AI, each accompanied by a set of questions that guide the assessment process (AI HLEG, 2020). The framework also includes definitions of key terms, ensuring clarity for stakeholders

involved in the evaluation. To ensure a thorough and effective assessment, the framework recommends the participation of multiple stakeholders or a multidisciplinary team, including AI developers, data scientists, legal compliance officers and management professionals. External support is encouraged if any questions remain unclear during the process (AI HLEG, 2020).

Despite its comprehensive nature, the Ethics Guidelines for Trustworthy AI and the ALTAI have notable limitations. Firstly, to date, the framework has not been applied in a context similar to this research paper, which may lead to unforeseen challenges during its use. Additionally, the framework acknowledges that some of its principles may conflict with one another, particularly when applied in complex or context-specific scenarios (AI HLEG, 2019). A significant limitation of this research is also due to scope constraints, meaning not all questions within the assessment tool will be comprehensively addressed. Instead, the study will focus on summarising key questions presented in ALTAI to be able to highlight ethical concerns most relevant to the research context (see Section 3 and Appendix). While this approach may limit the breadth of the analysis, it aligns with the aim of uncovering pressing ethical challenges rather than achieving exhaustive compliance with the framework.

In conclusion, the framework used to guide this research in evaluating the ethical challenges of AI in asylum procedures in Germany is a human-rights-based framework emphasising that AI systems must be lawful, ethical, and robust, with a focus on ethical and robust dimensions. These are derived from four core ethical principles, namely respect for human autonomy, prevention of harm, fairness, and explicability, and operationalised through seven key requirements to ensure trustworthy AI throughout its lifecycle.

When talking about the principle of respect for human autonomy, the EU highlights the importance of fundamental rights that uphold human autonomy, ensuring individuals retain self-determination and meaningful choices in their interactions with AI systems. These systems should be designed to enhance human capabilities, support human oversight, and contribute to the creation of meaningful work in alignment with these rights. The principle of prevention of harm asserts that AI systems must not cause or exacerbate harm to individuals, emphasising the importance of protecting human dignity and integrity while ensuring safety, security, and technical robustness. Additionally, it underscores the necessity of including vulnerable persons in the development, deployment, and use of AI systems, particularly in contexts where power or information asymmetries may lead to negative impacts. The principle of fairness in AI encompasses both substantive and procedural dimensions, highlighting the equitable distribution of benefits and protection against bias, along with ensuring transparency and the right to contest decisions made by AI systems and their operators. This principle advocates for equal opportunities and a balanced consideration of competing interests in the development and deployment of AI technologies. Lastly, the principle of explicability is vital for fostering user trust in AI systems, necessitating transparent processes and clear communication about the system's capabilities and decision-making. In cases where decisions stem from "black box" algorithms, alternative measures such as traceability and auditability should be implemented, with the level of explicability required tailored to the context and potential consequences of errors.

From those four ethical principles the seven requirements are derived: Human Agency and Oversight; Technical Robustness and Safety; Privacy and Data Governance; Transparency; Diversity, Non-discrimination and Fairness; Societal and Environmental Wellbeing and lastly Accountability. These will be further explained in the methods section.

# 3. Methods

In the subsequent section, the methodological framework employed in this research to address the posed research question will be outlined. This will encompass an overview of the research design, a concise case description and an outline of the data analysis process.

## 3.1. Research Design

To address the research, this study employs a qualitative research design centred on textual analysis (Given, 2008) to investigate the ethical guidelines surrounding AI implementation in the public sector, with a particular focus on asylum procedures in Germany. A qualitative textual approach is particularly well-suited for analysing a variety of texts, which is also the main source of data. By concentrating on Germany as a single case study, the research facilitates a thorough examination whilst avoiding the limitations of a broader scope. The aim is not to draw comparative conclusions but to explore specific dimensions within this critical context. Given Germany's pivotal role in Europe's refugee migration and intake, it serves as an exemplary case for this inquiry.

The ethical implications will be assessed by first analysing the compliance of AI systems in asylum procedures with the Ethics Guidelines for Trustworthy AI which includes pre-established principles. By applying a predefined ethical framework, the research is structured to evaluate AI tools systematically.

Nevertheless, certain challenges arise when choosing this approach. Textual analysis is particularly susceptible to the unconscious biases and perspectives of the researcher. To enhance the trustworthiness and credibility of textual analysis (Given, 2008), several measures can be implemented. These include defining thematic categories clearly, ensuring transparent coding and maintaining detailed documentation. It is essential that the categories used for identifying content are well-defined and that the coding process is meticulously recorded. Furthermore, it is advisable to select data that is known for its careful wording, such as finalised policy papers (Radu, 2021).

Hence, the aim is to provide a clear coding scheme, detailed documentation and that the unclear terms used in the framework are defined using the AI HLEG presented in their finalised papers attached in the appendix. Hence, the objective is to create a clear coding scheme and provide comprehensive documentation. Additionally, any ambiguous terms used in the framework will be defined according to the definitions from the AI HLEG, as presented in their finalised papers (see Appendix No.1). This approach aims to ensure clarity and consistency throughout the framework.

While complete eradication of bias in interpretive and subjective research designs, such as textual analysis, is rather unrealistic (Given, 2008), the objective remains to promote transparency in the coding scheme and documentation, thereby enabling other researchers to comprehend how specific findings were obtained.

## 3.2. Case Description

This research will be conducted as a single-case study of Germany's asylum process. The reason for that is the significant changes over the past decade, which accelerated technological as well as AI adoption in this process.

Germany is a country of immigration, with around 30% of its population having a migration background as of 2023 (BpB, 2021). Despite its long history of immigration, it was not until the 2000s that this reality was officially acknowledged, leading to liberalisation of migration laws, particularly for skilled migrants, while asylum has remained a contentious issue, especially following the significant influx of refugees in 2015 (BpB, 2021). In Germany, the number of asylum applicants rose dramatically from an annual average of 34,000 between 2003 and 2013 to 173,000 in 2014 (Herbert & Schönehagen, 2020). By mid-2015, the government adjusted its forecast to 800,000 refugees for that year, predominantly from Syria, Iraq, and Afghanistan (Herbert & Schönehagen, 2020). On August 21, 2015, the BAMF issued an internal directive to not reject Syrian refugees without registration in another EU country, reflecting the humanitarian crisis of the Syrian refugees and the overwhelmed administrative capacities (Herbert & Schönehagen, 2020). The situation escalated when news of the directive spread and falsely understood to mean that Germany would let refugees in without checking, leading to large-scale movements of refugees to Germany (Herbert & Schönehagen, 2020). On August 28, 2015, a truck in Burgenland, Austria, was discovered containing 71 deceased refugees from Syria and Iraq, while the subsequent photo of Alan Kurdi, a drowned two-year-old boy, sparked global outrage and highlighted the refugee crisis in Europe (Herbert & Schönehagen, 2020). In response, German media and Chancellor Angela Merkel emphasised the moral obligation to assist refugees, culminating in the “March of Hope” as thousands of individuals departed Budapest to seek refuge in Austria and Germany (Herbert & Schönehagen, 2020). By the summer of 2016, approximately 1.4 million refugees had entered Germany (Herbert & Schönehagen, 2020).

This crisis has not only profoundly shaped migration policies and political discourse in the country making the so-called ´refugee question´ has been the most contentious issue in Germany and the EU, but it has also placed significant pressure on Germany's asylum infrastructure (Herbert & Schönehagen, 2020). Thereby setting the stage for technological and administrative changes, including technological adoption, to manage the influx. The BAMF has since 2017 introduced many IT assistance systems and can be considered one of the forerunners in this particular field in the EU (Biselli, 2024).

The German asylum procedures, in which the BAMF plays a crucial role, can be summarised in the following (derived from the official document by the BAMF (2024)):

1. Arrival and Registration

Asylum seekers must report to a state organisation, such as border police or reception centres, where personal data, including photographs and fingerprints, is collected and stored in the Central Register of Foreigners (AZR). This information is checked against national and EU databases like EURODAC under the Dublin III Regulation to determine responsibility for the asylum claim, and applicants receive proof of arrival to reside and access state benefits.

2. Initial Distribution of Asylum Seekers (EASY)

Applicants are distributed across Germany to reception facilities using the EASY quota system, which ensures fair allocation based on each federal state's capacity. These facilities provide accommodation, basic needs, and guidance on subsequent steps.

3. Personal Asylum Applications

Applicants file their asylum applications at a BAMF branch, often with interpreter assistance, and provide details on their identity and reasons for fleeing. BAMF also assesses whether the Dublin III Regulation applies, potentially transferring applicants to another responsible EU country. (Note: Dublin III Regulation assessing previous registration in other EU countries was temporarily suspended for Syrian refugees.)

4. The Personal Interview (Most Important)

Applicants explain their reasons for seeking protection and provide evidence of threats to their safety or freedom during a detailed interview conducted by trained BAMF officials, with interpreters present. The credibility of claims is a key factor in the decision-making process.

5. The Decision of the Federal Office

BAMF issues decisions granting one of four outcomes: Asylum, Refugee Protection, Subsidiary Protection, or a National Deportation Ban or rejecting the application. Those denied protection (Asylum, Refugee Protection, Subsidiary Protection) may not qualify due to factors such as criminal history or threats to public safety. National Deportation Ban is issued when returning the applicant would violate their human rights (put forth in ECHR) or put their life and liberty at risk.

6. Appeals against the Decision

Rejected applicants can appeal through the administrative court system, allowing judicial review of BAMF's decision. While appeals are pending, applicants may remain in Germany except in expedited cases.

7. The Outcome of the Asylum Proceedings

Outcomes include granting protection, temporary suspension of deportation, or repatriation. Rejected applicants may receive assistance for voluntary return or face deportation if no other legal grounds for residency apply.

Special procedures exist for vulnerable groups, such as unaccompanied minors, individuals with disabilities, or victims of trauma. Applications from nationals of safe countries of origin

undergo accelerated procedures, as claims from these applicants are generally deemed less likely to succeed. Similarly, asylum seekers arriving at airports without valid documents are subject to fast-track reviews, where decisions are made within days to determine admissibility.

The analysis seeks to review the existence of AI adoption within these steps of the asylum procedure based on secondary data.

## 3.3. Data Collection

This study will exclusively analyse secondary data which includes scientific literature, policy documents, governmental websites and news articles/blog posts that include one or more of these topics: *AI, AI in Public Administration, AI and Ethics, AI in asylum processes, AI and BAMF*, *Human Rights and AI.* These texts should provide the basis to address the ethical implications as well as the descriptive analysis of AI implementation in the German asylum process. To collect this data Google, Google Scholar, the official BAMF website, and relevant news outlets and organisations specializing in AI and asylum issues will be utilised to search both in German and English.

Lastly, the proposed recommendations for policymakers at the end of my research will be mainly based on the result of the analysis itself, however, previous scientific research and other secondary sources might also be considered if necessary.

## 3.4. Data Analysis

The analysis is divided into two sections: the collection of tools and the ethical implications analysis.

The first step involves gathering all applications utilised in the German asylum process. This phase serves as the foundation for the analysis, providing essential context about the asylum procedure itself. All identified AI tools will be compiled into a structured Excel file (see Figure 2) that includes a coding table for easy reference. Additionally, all sources will be documented to ensure transparency and support for the subsequent analysis.

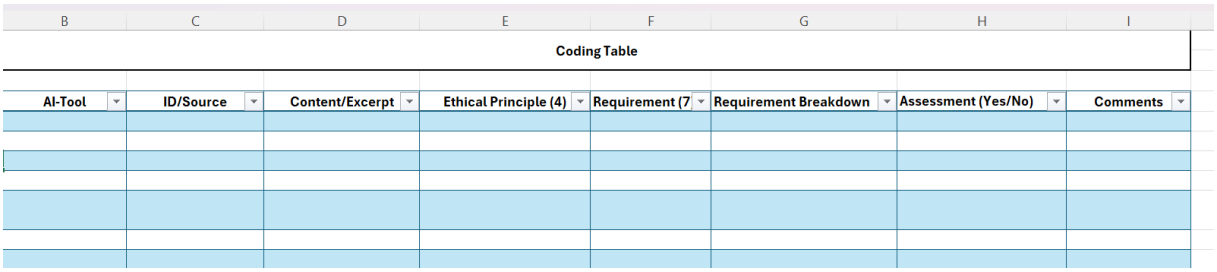


*Figure 2: Example of the Excel file (might slightly differ)*

The second phase of the analysis focuses on assessing how these tools function and whether they adhere to the seven established requirements that were derived from the four ethical principles established in the Ethics Guidelines for Trustworthy AI. For the ethical implications analysis, a coding table that consists of the four ethical principles as the main themes, then these seven requirements will further break down each of those principles. These seven requirements are explained in detail within the Assessment List for Trustworthy AI (ALTAI) (AI HLEG, 2020). This Assessment List however was revised and summarised to better fit the scope of the analysis, and this revised Assessment List is added in the appendix (see Appendix No.1).

The coding table, derived from the ALTAI, will again break down each requirement into manageable components (see Figure 3). These components are then called "Requirement-Breakdown" in the coding table and the Excel file. The ALTAI provides explanations and multiple questions for each requirement breakdown, which were subsequently revised and condensed into 2 to 5 questions maximum (see Appendix No.1). These revised questions are also incorporated into the Excel file and should be answered accordingly.

Key sections of the relevant texts will then be examined, highlighted and added to the Excel file to identify how well the tools align with these principles using an assessment mark of 0-2 (0 indicating non-compliance/no information; 1: partial compliance/little information; 2: compliance/ sufficient information). After the assessment the tools can reach a maximum of 52 points. The extracted information will also be incorporated into Sections 4 and 5 of the thesis. The Excel file will be used to filter and organise data by each principle and tool, supporting a comprehensive discussion of the findings.

It is anticipated that some principles or questions may be challenging to fully assess due to limitations in the available information. Consequently, this model is designed to serve as a flexible guide rather than a rigid, step-by-step framework, though the analysis will aim to adhere to it as closely as possible. Meaning if there is a lack of information/transparency in a particular system this will be incorporated in the grading system of (0-2).

**Coding Table** (Ethical Implications Analysis)

| **Theme (4 Ethical Principles)** | **Requirements (7 Principles of Trustworthy AI)** | **Requirement-Breakdown (from ALTAI)** |
|---|---|---|
| **Respect for human autonomy** | Human agency and oversight **(Principle 1)** | Human Agency and Autonomy |
| | | Human Oversight |
| **Prevention of harm** | Technical robustness and safety **(Principle 2)** | Resilience to Attack and Security |
| | | General safety |
| | | Accuracy |
| | | Reliability, Fall-back Plans and Reproducibility |
| | Privacy and Data Governance **(Principle 3)** | Privacy |
| | | Data Governance |

| | | |
|---|---|---|
| | | |
| | Societal and Environmental Well-Being **(Principle 6)** | ➔ See *Fairness* |
| **Fairness** | Diversity, Non-Discrimination and Fairness **(Principle 5)** | Avoidance of Unfair Bias |
| | | Accessibility and Universal Design |
| | | Stakeholder Participation |
| | Societal and Environmental Well-Being **(Principle 6)** | Environmental Well-being |
| | | Impact on Work and Skills |
| | | Impact on Society at large or Democracy |
| | Accountability **(Principle 7)** | Auditability |
| | | Risk Management |
| **Explicability** | Transparency (**Principle 4**) | Traceability |
| | | Explainability |
| | | Communication |

*Figure 3: Coding Table for Ethical Implications Analysis derived from Assessment List for Trustworthy AI (ALTAI) for self-assessment (AI HLEG, 2020)*

# 4. Analysis

In this section, sub-questions 1 and 2 will be addressed, in which the history of AI implementation in asylum procedures, as well as a small section on the EU, will be presented followed by the description of the adopted AI tools with their analysis based on the Ethics Guidelines for Trustworthy AI.

## 4.1. Migration, AI and the EU

Before discussing the history of AI implementation in German asylum procedures, it is important to understand the broader use of technological and AI tools in migration and asylum management across the EU that directly influence Germany's practices as a member state.

The migration crisis of 2015–2016 catalysed the EU and its member states to develop new technologies to manage migration and asylum processes (Ozkul, 2023). This context gave rise to tools like ITFLOWS (2020–2023), which developed the EUMigraTool (EMT) (Ozkul, 2023). EUMigraTool predicts migration flows to specific European countries and analyses migration drivers, patterns, public sentiment, and tension risks between migrants and EU citizens (Ozkul, 2023).

Another tool is the Early Warning and Preparedness System (EWPS), developed by the European Asylum Support Office (EASO), now operating under the European Union Agency for Asylum (EUAA) (Ozkul, 2023). EWPS forecasts migration patterns to EU territories using data from four key sources: GDELT, which monitors daily events by country of origin; Google Trends, which tracks weekly online search trends; Frontex, which provides statistics on irregular border crossings; and the agency's internal data on asylum applications and recognition rates across EU Member States, Norway, Switzerland, and the UK (Ozkul, 2023).

Individual EU Member States have also started to develop their own systems, like the Predictive Migration Analysis project by Germany (BMI, 2023). The Predictive Migration Analysis (VoMa) project by the BAMF aims to develop a stand-alone AI-based IT tool and methodology to analyse migration trends in a predictive and timely manner, creating scenarios for potential migration.

Forecasting tools and centralised systems reflect the EU's dual priorities of resource planning and border control. On one hand, they help authorities and humanitarian organisations allocate resources to prepare for displaced people and on the other hand, they risk reinforcing surveillance and securitisation at borders, as seen in the post-crisis environment (Ozkul, 2023). These tools, including EUMigraTool, demonstrate how technology is shaping migration management in Europe, with implications for both efficiency and ethics (Ozkul, 2023).

Germany's predictive system was excluded from the analysis and assessment, as it operates outside the formal asylum procedure. However, it is pertinent to mention it for contextual clarity. In this context, it is also worth mentioning that the German government has adopted an AI strategy for 2018-2025 and spent 3.5 billion on its implementation (Deutscher Bundestag, 2023-b). Thus, the increase in the implementation of AI in the asylum procedure can also be seen as a broader development for the entire German government and its agencies.

## 4.2. AI in German Asylum Procedures

While researching the AI tools currently implemented in the German asylum system, many automated systems emerged. However, it is important to note that not every automated system qualifies as AI; some are merely IT systems, and it is often unclear whether they involve IT, AI, or both. The Federal Office for Migration and Refugees (BAMF) has also implemented various IT assistance systems, such as the name transliteration system (TraLitA), but it is uncertain whether these systems employ AI (Ozkul, 2023). Both the TraLitA program and the Mobile Phone Data Analysis (AmD) system were excluded from the analysis due to ambiguity regarding the use of AI. Initially, AmD was included but because of growing concerns about its

classification as AI it was reconsidered. Notably, the German government did not mention AmD in its list of AI tools when requested (Deutscher Bundestag, 2023-b). The government explained that not all systems with AI components are necessarily included, as some may be difficult to detect and that they focused on explicit AI systems (Deutscher Bundestag, 2023-b). It was said that only systems are mentioned in which AI is used specifically and explicitly for (partially) automated decisions or pattern recognition. Furthermore, certain systems may not be mentioned if there are reasonable grounds to suggest that the security of the state could be jeopardised. However, since AmD has already been referenced by the BAMF and the German government in other contexts, it can be assumed that AmD is not subject to these restrictions (BAMF, 2024). Further examination of other sources also failed to clarify whether AI was involved in these tools or if they were solely, IT systems. Consequently, AmD was also excluded from consideration.

A similar issue arises with the Dialect Identification Assistance System (DIAS). Although this system was not included in the list provided by the German government, other credible sources, including the European Migration Network (EMN), have indicated that it is indeed an AI system (EMN, 2022).

Many of the AI and IT systems implemented in the asylum process in Germany are in the registration and application phase to support the interview and the final decision on the application, which can be seen in Figure 4, and have been developed since 2017 (BAMF, 2024).

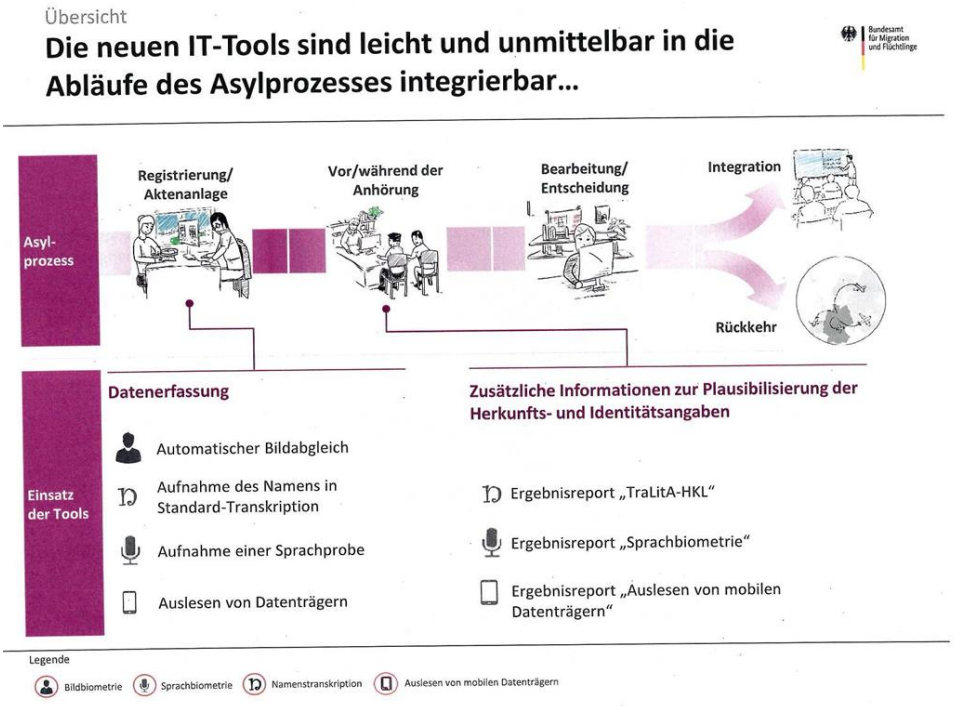


*Figure 4: Derived from training material published by BAMF (2017); ´Sprachbiometrie´ is the DIAS, ´Auslese von (mobilen) Datenträgern´ is AMD.*

The BAMF mentioned in 2019 that AI methods are used in three projects, including the ZPE (Central Inbox) project, the EGVP (Electronic Court and Administrative Mailbox) and the so-called 'Profile Analysis' project (Biselli & Meister, 2019). However, upon further research, specific information on explicit tools could only be found for the Profile Analysis, which was,

however, renamed to ASS (Assistance system for safety messages) (Deutscher Bundestag, 2023-a). The Profile Analysis or ASS is one clear example of an explicit AI system, that was also highlighted by the Bundestag, news articles and research (Deutscher Bundestag, 2023-b). Considering that not all AI systems are publicly communicated protecting state security it can be assumed that way more might be involved in the asylum system but were not mentioned (Deutscher Bundestag, 2023-b).

Additionally, there were other systems mentioned to explicitly include AI, namely the Asylaktendurchdringungsassistenten or short: ADA. ADA is being piloted 2024 at the administrative court Karlsruhe's Kompetenz- und Innovationszentrum Asyl as part of a comprehensive initiative to modernise and accelerate asylum procedures (MJ BW, 2024). ADA, short for Asylaktendurchdringungsassistent, is introduced as a support tool that automatically scans asylum case files to identify specific information relevant for processing and marks these findings with a digital sticky note (MJ BW, 2024). This marking enables required data to be entered more quickly into the justice systems' databases and pre-structures the files, thereby facilitating an immediate start to case processing (MJ BW, 2024). ADA is also not the only assistance system used in the administrative courts in Baden-Württemberg with a similar tool, also the "Herkunftslandinformationsassistenten" (HekLA), translated to ´Country of origin information assistant' provides a quick and easy full-text search option for the state-wide asylum database of the Administrative Court of Baden-Württemberg (Bundesrat, 2024). Initial results indicate an acceleration in processing asylum applications in Baden-Württemberg, likely influenced by such systems (Redaktion Filstalexpress, 2025). However, after having searched all the pages that could be found on the ADA tool, the information that could be collected is rather scarce. Through careful consideration, this tool was also excluded from the assessment entirely for the mentioned reasons (Note: HekLA was only found later on in the analysis and the same problem (even more emphasised) arises with this tool).

BAMF is currently exploring the development of a tool called Bamf-GPT to assist the ministry in efficiently locating the necessary information within the extensive range of BAMF documents, such as country documentation, intranet articles, and instructions (Heeger, 2024). Currently, this tool is still in the testing phase, and it is uncertain whether it will be implemented in the asylum process (Heeger, 2024; Deutscher Bundestag, 2023-b). Since the system is still being tested and its potential role in assisting with the asylum process is unknown, it was not included in the analysis.

Hence, in the following sections, more detailed descriptions of the AI tools DIAS and ASS are provided.

### 4.2.1. Dialect Identification Assistance System (DIAS)

Germany is the first in Europe to introduce a dialect identification assistance system (DIAS) for processing asylum applications (Ozkul, 2023). Since September 2017, the BAMF has been using DIAS, following an initial trial in Bamberg earlier that year (Ozkul, 2023; Amnesty Digital, 2024; Zomignani Barboza, 2021; Deutscher Bundestag, 2023-a). Its legal basis is Section 16, Paragraph 1, Sentence 3 of the Asylum Act, which allows audio recordings of applicants'

oral statements to determine their origin, provided they are informed in advance (Ozkul, 2023). The reasoning behind the usage of DIAS presented by BAMF is that it addresses the lack of ID documents among asylum seekers, improves efficiency, identifies fraudulent claims and provides evidence for returns, required for the origin countries to accept the rejected applicants (Ozkul, 2023).

When using this tool, often referred to as ´speech biometrics´, asylum seekers are required to describe an image in their native language for about two minutes while speaking into a phone (Amnesty Digital, 2024). What they say is recorded, and the speech samples are then automatically compared with a language model (Amnesty Digital, 2024). This process calculates the probabilities of the spoken dialect, which are documented in a results report (Amnesty Digital, 2024). Results are documented as a PDF-file and stored in the applicant's electronic case file (Ozkul, 2023). The exact process is also illustrated in the Figure 5.

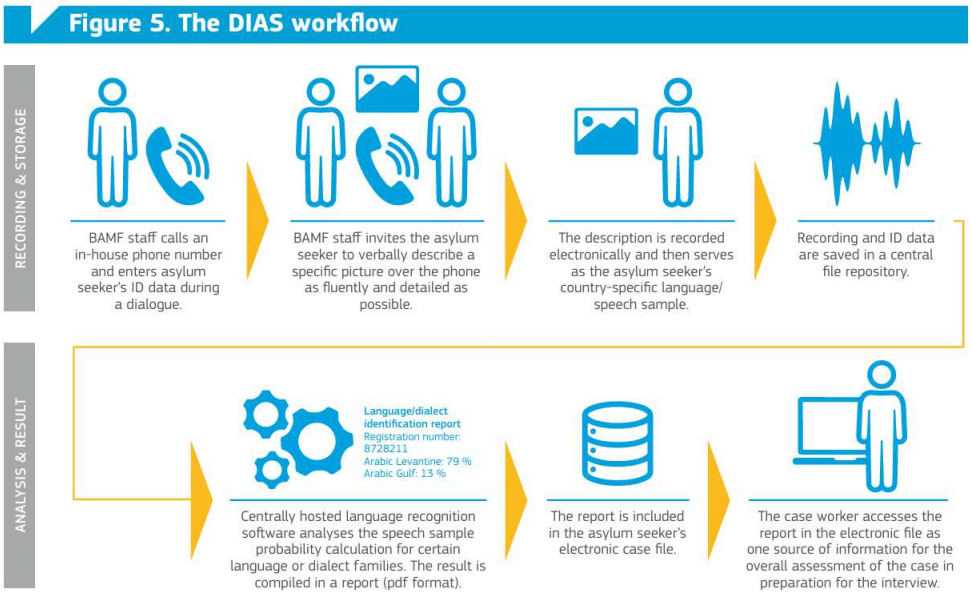


*Figure 5: (EMN, 2022, p.9.)*

Initially capable of recognising only Arabic dialects (Maghrebian, Levantine, Iraqi, Egyptian, Gulf), the system has been, according to Ozkul (2023) based on a response by the Bundestag, expanded since July 2022 to include Dari, Persian/Farsi, and Pashto, with Kurdish under development (Ozkul, 2023). In the more recent response by the Bundestag in November 2023 however, it was said that only the Arabic dialects (Maghrebian, Levantine, Iraqi, Egyptian, Gulf), Dari and Farsi (Persian) are used (Deutscher Bundestag, 2023-a). Meaning: Pashto was excluded, and Kurdish was not yet implemented. Reasons for their exclusion were not provided.

The DIAS tool uses the Nuance Speech Suite, developed by Atos, a multinational company specialising in digital transformation and speech recognition, with capabilities tailored for forensic and intelligence purposes (Ozkul, 2023). DIAS classifies dialects by analysing phonemes and generating probabilities (e.g., 60% Arabic Levantine, 20% Arabic Gulf) (Ozkul, 2023). Training data was sourced from the Linguistic Data Consortium (LDC), Clickworker GmbH, and anonymised speech samples from BAMF itself, covering 19 languages to enhance dialect distinctions (Ozkul, 2023).

From 2017 to August 2022, the system cost approximately 5.03 million euros to purchase, adapt, and expand (Ozkul, 2023). By mid-2022, DIAS had analysed tens of thousands of speech samples annually, supporting BAMF's decision-making process (Ozkul, 2023). However, its reports are considered complementary evidence and not the sole determinant of a case's outcome (Ozkul, 2023). If a DIAS result contradicts an applicant's account, they are given the opportunity to address it during the asylum hearing (Ozkul, 2023). Despite its efficiency, DIAS does not automate identification or credibility assessments but provides automated evidence to assist caseworkers in their evaluations (Ozkul, 2023).

In June 2018, the DIAS tool won the "Best Digitisation Project" award for its innovative contribution to modernising public administration (Ozkul, 2023). BAMF highlighted its pioneering role in implementing this technology, noting that other European states have shown great interest in its practices (Ozkul, 2023). As a leading institution in digitalisation, BAMF has sought to collaborate with partner institutions to further develop these systems and exchange knowledge (Ozkul, 2023). It has called for increased cooperation at a European level, encouraging other institutions to join its efforts (Ozkul, 2023). BAMF has shared anonymised language samples with the Netherlands for linguistic testing in a pilot project and exchanged information about the DIAS tool with Austria, Finland, Norway, Sweden, Lithuania, Greece, Switzerland, and the Netherlands (Ozkul, 2023). By 2022, European-level collaboration was being discussed and tested in pilot projects involving these countries, yet it seems no other country has implemented it to date (Ozkul, 2023).

### 4.2.2. Assistenzsystem für Sicherheitsmeldungen (ASS)

The Assistenzsystem für Sicherheitsmeldungen (ASS) is an AI-based system used by the German Federal Office for Migration and Refugees (BAMF) to assist in identifying security-relevant information within asylum hearing transcripts (BAMF, 2024-b); Hessischer Flüchtlingsrat (hfr), 2024; Deutscher Bundestag, 2023-b; Deutscher Bundestag, 2023-a). Initially developed under the name "Profilanalyse," the system was introduced as a pilot project in 2017 and became fully operational across all BAMF outposts in 2022 (Deutscher Bundestag, 2023-; Biselli, 2023). BAMF legally bases the implementation of this system on many statutes mandating BAMF to forward security-related findings to relevant authorities (Deutscher Bundestag, 2023-a).[1]

ASS works in the following way: Once an asylum interview has been saved and completed in the BAMF asylum procedure system the interview gets automatically forwarded to ASS and converted into a machine-readable form if required. It then scans asylum hearing transcripts for potential security-related content such as references to war crimes, organised crime, or state security threats (see Figure 6) (BAMF, 2024-b). It does not autonomously decide on reporting but serves as a decision-support tool for BAMF asylum officers (BAMF, 2024-b). The system employs a two-stage analysis process to identify security-relevant text in asylum hearing

[1] §18 Abs. 1, 1a, 2 Bundesverfassungsgesetz, § 23 Abs. 3 Gsetz über den Bundesnachrichten-dienst, § 23 Abs. 3 gesetz über den Bundesnachrichtdienst, § 9 Bundeskriminalamtgesetz, § 8 Abs. 3 Asylgesetz (Deutscher Bundestag, 2022)

transcripts. First, it uses a rule-based system with semantic rules and specialised dictionaries to detect keywords and highlight a five-sentence text window around them (Deutscher Bundestag, 2023-a). In the second stage, the highlighted text is analysed using supervised machine learning models, specifically fine-tuned BERT models, which classify the content based on predefined reporting criteria (Deutscher Bundestag, 2023-a). In the second step it used a similar technology to LLM, which could be called "Small Language Model" (Heeger, 2024). The system, developed by SVA/IBM, is not fully automated and relies on manually labeled, anonymised training data (Deutscher Bundestag, 2023-b; Deutscher Bundestag, 2022). If relevant information is detected, the caseworker/asylum officer is notified by email and it is highlighted for caseworkers/asylum officer in a web interface of the intranet of BAMF, where they classify passages as "helpful and relevant," "not relevant," or "helpful but not reportable" (Deutscher Bundestag, 2023-a). The caseworker can also manually highlight other passages deemed relevant (Deutscher Bundestag, 2023-a). Each decision has to be further explained in the report by the officer before sending it to the BAMF central unit which is responsible for cooperation with ethe security authorities for further inspection

Once flagged, a human reviews the text before forwarding it to BAMF's central security coordination unit (Deutscher Bundestag, 2023-a). Another review is conducted before reporting to security agencies, ensuring a multi-step validation process (Heeger, 2024). The system does not store or process personal data such as gender or nationality to prevent bias, and training data consists of anonymised excerpts from past hearings, supervised by BAMF's data protection officer (Deutscher Bundestag, 2023-a).

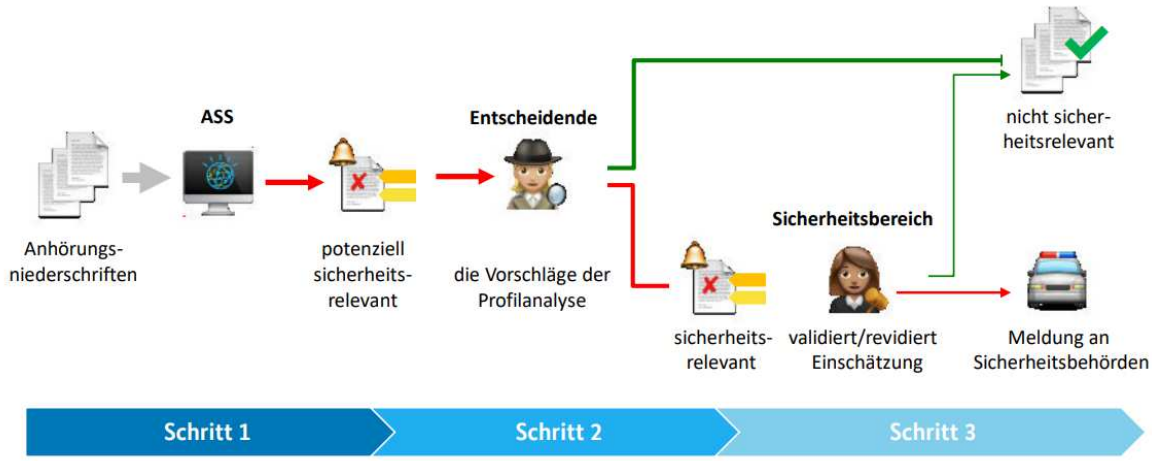


*Figure 6: Illustration of the ASS System: Steps from Hearing Protocol to Reporting to Security Authorities. (BAMF, 2024, p.2)*

ASS was externally developed by SVA/IBM and received €18 million in funding between 2018 and 2021 before full implementation in 2022 (Deutscher Bundestag, 2023). It operates within BAMF's private cloud infrastructure at ITZ Bund for data security (Heeger, 2024). The system is continuously updated, with refinements based on caseworker feedback to improve classification accuracy (Deutscher Bundestag, 2023). The number of information that was reported in 2018 was 14.769 and it continuously decreases until 2023 with 7.225 that were reported back to security authorities by BAMF using ASS (Deutscher Bundestag, 2023-a)

The system has been recognized for its societal relevance, securing third place in the 2024 eGovernment competition (BAMF, 2024). While it aims to enhance efficiency in identifying security threats, BAMF maintains the position that all final decisions remain with human officers, preserving the four-eyes principle in the reporting process (Heeger, 2024).

# 5. Discussion

This section will address the second and third sub-questions by presenting the results of the Ethics Implication analysis based on the Ethics Guidelines for Trustworthy AI for both DIAS and ASS. It will cover the challenges, implications, and policy recommendations identified through this analysis.

## 5.1. Ethical Challenges in AI-sylum

The ethical challenges presented by the AI systems DIAS and ASS, when assessed through the framework of trustworthy AI guidelines, are intricate and raise considerable concerns.

A primary issue is the risk associated with automation bias and confirmation bias. Although the BAMF maintains that AI systems only assist and do not decide, in practice, decision-makers may unintentionally assign excessive weight to AI-generated results, particularly under conditions of time pressure and cognitive bias. This can lead to an overreliance on these outputs, thereby reinforcing pre-existing assumptions and prejudices. Although for ASS, BAMF mentioned that training is provided to combat automation and confirmation bias it is unclear what that exactly entails. For DIAS, no information could be found on such training.

Another significant concern pertains to informed consent. In numerous instances, applicants are seldom made aware of the AI tools employed in the decision-making process. This lack of transparent communication effectively results in coerced consent, as asylum seekers feel compelled to accept these processes, recognising that refusal could jeopardise their application. Such a deficit of voluntary, informed participation undermines the ethical principle of autonomy. It is also unclear whether asylum officers are fully aware of the use of AI in those systems, although it can be assumed that they are aware to some extent. Applicants need to be informed that an AI system will be used to analyse their hearings for security-relevant information, rather than simply stating that security-relevant information will be forwarded without mentioning the AI component. Asylum applicants are informed to some extent why DIAS is used and regarding ASS they are completely unaware.

Regarding ASS for instance a grading mark of 31 (out of 52) was received and many things can be highlighted regarding this system. Information is broadly selected purposefully to ensure that no important details are lost. This means that even information which may be deemed of low relevance to security will be highlighted by the system, which is why more than half of the reports are rejected by asylum officers afterwards. This process is intentional. It was criticised by parliamentarian Clara Bünger from the Party Die Linke, who consistently asked the

government to clearly define and limit what information is collected by the system to avoid endangering the rights of asylum seekers.

Additionally, it can be questioned whether this system actually enhances efficiency or merely creates more work, as it highlights a broad spectrum of information that then needs to be verified by two separate instances. There is also the issue that there are no actual statistics on the relevance of the reported notifications and what they entail, leading to a significant lack of transparency.

This transparency issue is complicated; on the one hand, it is necessary to have transparent processes to assess the ethics and efficacy of such a system. On the other hand, security apparatuses enjoy certain protections due to the potential to harm the state. Nonetheless, there must be some independent body that analyses the effectiveness of such a system.

A positive aspect of this system is that two human officers are reviewing the results produced by the AI. Moreover, from the first instance, it is communicated to the officers that the responsibility lies with them and it is not sufficient to just forward the text passage, they must also formulate an actual report. Therefore, there seems to be a thorough review in place.

Despite these challenges, using AI only as an aid in decision-making is a promising approach. Ethical assessments highlight the importance of maintaining human oversight and autonomy, which would be threatened if these systems were granted decision-making powers. The ASS system illustrates this well, as it preserves human involvement through a two-step protocol that controls the security passages identified by the system, primarily serving to highlight relevant information.

In contrast, the DIAS system, although also providing insights for decision-makers, does not make decisions on its own. Instead, it delivers results that must be reviewed by asylum officers in light of the statements of the applicant, to assess whether it is reliable or not. Precisely in this case automation bias could be problem. A notable benefit of DIAS is the flexibility it offers, allowing asylum officers to stop, abort, or repeat the process. However, this option is generally accessible only to the officers.

The DIAS system is applied in specific cases, while the ASS system is employed for all hearing protocols, often without the asylum applicant's knowledge. In contrast, asylum applicants using the DIAS system are informed to some extent about its purpose and operation, while informed consent is notably absent in the ASS system. Overall, consent plays a limited role in the asylum procedure due to the power imbalance and the potential consequences of non-compliance.

## 5.2. Ethical Implications and Policy Recommendations

BAMF consistently emphasizes that AI systems function solely as assisting tools, lacking the authority to make final decisions. However, challenges such as confirmation bias, automation bias, and time constraints can unduly influence the responses generated by these systems.

Integrating artificial intelligence (AI) into Germany's asylum processes presents a valuable opportunity to address pressing issues. However, it is crucial that this implementation does not compromise the fundamental rights of asylum seekers. Although AI has the potential to streamline administrative processes and improve decision-making, its application in this sensitive context requires stringent ethical standards and robust oversight to ensure the protection of those affected.

The AI Act, which represents the first significant regulatory framework for artificial intelligence in the European Union, marks an important advancement. However, although it acknowledges migration and asylum as high-risk areas, it does not fully address the specific challenges and potential pitfalls inherent in this context.

There needs to be the focus on the development of tailored ethical frameworks specifically designed for migration. These frameworks ought to extend beyond existing general principles to adequately consider the unique vulnerabilities and rights of asylum seekers. To ensure that these guidelines are more than mere aspirations, they need to be binding, with clear and enforceable standards. Moreover, the issue of enforceability remains pressing. Existing ethical frameworks for AI frequently lack binding power, functioning more as instruments for "ethics washing" or public relations rather than as enforceable guidelines that ensure accountability. This deficiency allows institutions to evade more stringent regulatory measures, creating significant gaps in oversight. In the absence of binding standards, the application of AI in sensitive domains such as asylum procedures risks perpetuating systemic biases and compromising the fundamental rights of individuals seeking protection.

It is also essential to establish an independent oversight body responsible for the continuous assessment and regulation of AI applications in asylum procedures. This body must operate free from political influence to guarantee transparency and accountability in its evaluations.

Moreover, achieving a global consensus on AI ethics, similar to the international processes that have shaped human rights norms, can strengthen the credibility and legitimacy of these policies. This collaborative approach would not only align national standards with international best practices but also ensure that the deployment of AI technologies respects and safeguards human dignity and rights across borders.

# 6. Conclusion

## 6.1. Answer to Research Question

The incorporation of artificial intelligence (AI) into Germany's asylum processes presents significant ethical challenges, particularly regarding transparency, bias, and accountability. This study focuses on two primary AI systems employed by Germany’s Federal Office for Migration and Refugees (BAMF): the Dialect Identification Assistance System (DIAS) and the Assistenzsystem für Sicherheitsmeldungen (ASS). Although these systems aim to improve efficiency, they raise important concerns about fairness and the safeguarding of asylum seekers’ rights.

DIAS, introduced in 2017, analyses the spoken dialects of asylum seekers to determine their claimed origins and assigns probabilities for dialect classifications to assist in decision-making. In contrast, ASS, implemented in 2022, examines asylum interview transcripts for security-related information, flagging potential threats for further review by human agents. Other AI tools, such as the Asylaktendurchdringungsassistent (ADA), were not included in this analysis due to insufficient data regarding their AI functionalities or decision-making roles.

A principal ethical concern is the opacity of AI-assisted decision-making. Many asylum seekers remain unaware of AI's influence on their cases, and there is limited understanding of how BAMF officers interpret AI-generated data. The risk of dialect misclassification with DIAS is particularly high due to insufficient training data, which may lead to erroneous conclusions about asylum seekers' backgrounds. Although training exists for the ASS system, the specifics remain ambiguous.

Moreover, there is a significant concern regarding automation bias, where caseworkers may overly rely on AI outputs rather than conducting independent assessments. This risk intensifies under time pressures, with AI conclusions potentially unduly influencing decisions. The use of AI in security assessments raises additional concerns about surveillance and the potential criminalization of vulnerable groups. AI's lack of contextual understanding can result in false positives, unjust scrutiny, or wrongful deportations. The absence of informed consent further complicates the matter, as asylum seekers often do not fully grasp the implications of AI assessments on their cases.

To mitigate these ethical concerns, BAMF must enhance transparency, accountability, and fairness. Asylum seekers should be clearly informed about the role of AI in their applications, and BAMF should publish comprehensive explanations of AI model training and testing procedures. Independent oversight is essential, with a human rights-focused ethics committee regularly auditing AI tools for ethical and legal compliance. AI models should incorporate diverse dialect samples to mitigate discrimination, and protocols for bias detection must be established to ensure equitable outcomes. Additionally, human oversight should remain integral to the decision-making process, with caseworkers trained to critically evaluate AI outputs.

While AI holds the potential to streamline asylum procedures, it also poses substantial ethical risks. Without robust safeguards, AI systems may lead to unjust outcomes and violations of rights. By fostering transparency, independent oversight, and a commitment to human rights, it is possible to harness AI in asylum decision-making while preserving the fundamental rights of those seeking protection.

## 6.2. Limitations of Research and Recommendations

The research underpinning this study is subject to several limitations, particularly in relation to data sources and the depth of technical expertise. One significant limitation is the reliance on secondary data, as the AI systems under investigation are not thoroughly discussed by the pertinent agencies in public forums. Consequently, the majority of available information is derived

from researchers and parliamentarians. For a more accurate analysis of these AI tools, direct engagement with the respective agencies would be necessary to obtain detailed primary data. This absence of readily accessible information constrains the depth of analysis and impairs the ability to draw robust conclusions regarding the ethical and operational dimensions of the AI systems in question.

Furthermore, the research is hindered by my lack of specialised expertise in artificial intelligence. While the study can provide an initial exploration of the ethical concerns and regulatory implications, a more comprehensive analysis would greatly benefit from the input of dedicated AI experts. It is strongly recommended that future research involve specialists in artificial intelligence to bridge this knowledge gap. Researchers should also seek to establish an academic consensus on AI ethics by integrating interdisciplinary perspectives. This endeavour should include a transition from secondary data to the conduct of primary research, such as interviews and direct consultations with technology developers, regulators, and other key stakeholders. A more nuanced approach will facilitate the development of ethical frameworks tailored to migration and asylum contexts and support the formulation of potential extensions or revisions to existing regulations, such as the AI Act.

# 7. List of References


Abonamah, A. A., Tariq, M. U., & Shilbayeh, S. (2021). On the Commoditization of Artificial Intelligence. *Frontiers in Psychology*, *12*, 696346. https://doi.org/10.3389/fpsyg.2021.696346.

AI HLEG. (2020). *Assessment list for trustworthy artificial intelligence (ALTAI) for self-assessment*. Retrieved November 7, 2024, from https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

AI HLEG. (2019). *Ethics guidelines for trustworthy AI*. Retrieved November 8, 2024, from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

AlgorithmWatch. (2020). *AI ethics guidelines global inventory upgrade*. *AlgorithmWatch*. Retrieved November 7, 2024, from https://algorithmwatch.org/en/ai-ethics-guidelines-inventory-upgrade-2020/.

Amnesty Digital (2024): Amnesty Digital. (2024). *KI und Migration*. Retrieved December 30, 2024, from https://amnesty-digital.de/ki-und-migration/.

Artificial Intelligence Act (AI Act). (n.d.). *The EU Artificial Intelligence Act*. Retrieved January 3, 2025, from https://artificialintelligenceact.eu/.

Bakiner, O. (2023). The promises and challenges of addressing artificial intelligence with human rights. *Big Data & Society*, *10*(2). https://doi.org/10.1177/20539517231205476.

Biselli, A. (2024). KI in Asylverfahren: Gefährlicher Tech-Populismus. *Netzpolitik*. Retrieved November 8, 2024, from https://netzpolitik.org/2024/ki-in-asylverfahren-gefaehrlicher-tech-populismus/.

Biselli, A., & Meister, A. (2019). *Asylbehörde sucht mit Künstlicher Intelligenz nach auffälligen Geflüchteten*. netzpolitik.org. Retrieved January 13, 2025, from https://netzpolitik.org/2019/asylbehoerde-sucht-mit-kuenstlicher-intelligenz-nach-auffaelligen-gefluechteten/.

Biselli, A. (2023). *KI-Tools im BAMF: Ist das sicherheitsrelevant?* Netzpolitik.org. Retrieved January 13, 2025, from https://netzpolitik.org/2023/ki-tools-im-bamf-ist-das-sicherheitsrelevant/.

Bundesministerium für Arbeit und Soziales (BMAS). (2022). Selbstverpflichtende Leitlinien für den KI-Einsatz in der behördlichen Praxis der Arbeits- und Sozialverwaltung. *Bundesministerium für Arbeit und Soziales*. Retrieved November 5, 2024, from https://www.bmas.de/DE/Service/Publikationen/Broschueren/a862-leitlinien-ki-einsatz-behoerdliche-praxis-arbeits-sozialverwaltung.html.

Bundesministerium des Innern und für Heimat (BMI). (n.d.-a). Daten und KI. *CIO Bund*. Retrieved November 7, 2024, from https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/datenpolitik/daten-und-ki/daten-und-ki-node.html.

Bundesministerium des Innern und für Heimat (BMI). (n.d.-b). Verwaltungsrecht. *Bundesministerium des Innern und für Heimat.* Retrieved November 7, 2024, from

https://www.bmi.bund.de/DE/themen/moderne-verwaltung/verwaltungsrecht/verwaltungsrecht-node.html.

Bundesministerium des Innern und für Heimat (BMI). (n.d.-c). *Verwaltungsverfahrensgesetz*. Retrieved November 7, 2024, from https://www.bmi.bund.de/DE/themen/moderne-verwaltung/verwaltungsrecht/verwaltungsverfahrensgesetz/verwaltungsverfahrensgesetz-node.html.

Bundesministerium für Innern und Heimat. (2023). *Bericht des BMI für die 220. IMK vom 6. bis 8. Dezember 2023 in Berlin zum Thema „Aktueller Stand zur Einführung eines Informationssystems zu Migrationsbewegungen“*. Bundesministerium für Inneres und Heimat. Retrieved February 28, 2025, from https://www.innenministerkonferenz.de/IMK/DE/termine/to-beschluesse/2023-12-08-06/anlage-zu-top-9.pdf?__blob=publicationFile&v=2.

Bundesministerium für Migration und Flüchtlinge (BAMF). (n.d.). *Digitalisierung des Asylverfahrens*. Retrieved November 5, 2024, from https://www.bamf.de/DE/Themen/Digitalisierung/DigitalesAsylverfahren/digitalesasylverfahren-node.html.

Bundesamt für Migration und Flüchtlinge (BAMF). (2024). *The stages of the German asylum procedure*. Retrieved January 1, 2025, from https://www.bamf.de/SharedDocs/Anlagen/EN/AsylFluechtlingsschutz/Asylverfahren/das-deutsche-asylverfahren.pdf?__blob=publicationFile&v=25.

Bundesamt für Migration und Flüchtlinge (BAMF). (2024-b). *ASA: Automatisierte Unterstützung von Asylverfahren – Digitalisierung in der Migrationspolitik*. Retrieved January 13, 2025, from https://egovernmentwettbewerb.de/wp-content/uploads/2024/06/BAMF_ASS_Kat_5.pdf.

Bundesamt für Migration und Flüchtlinge (BAMF). (2024-c). *Anfrage nach dem Informationsfreiheitsgesetz*. [Document uploaded in response to information request]. Frag den Staat. Retrieved January 13, 2025, from https://fragdenstaat.de/anfrage/aktuelle-dienstansweisungen-zum-einsatz-von-it-assistenzsystemen-und-kuenstlicher-intelligenz-im-bamf-und-schulungdokumente/905671/anhang/13b-ifg-1157-bescheid_geschwaerzt.pdf.

Bundesamt für Migration und Flüchtlinge (BAMF). (2024-d). *Sprach- und Dialekterkennung DIAS Release 3.1. Tutorial für AVS-Mitarbeitende in den Außenstellen***.** Retrieved January 13, 2025, from https://fragdenstaat.de/anfrage/aktuelle-dienstansweisungen-zum-einsatz-von-it-assistenzsystemen-und-kuenstlicher-intelligenz-im-bamf-und-schulungdokumente/#nachricht-903767**.**

Bundesamt für Migration und Flüchtlinge (BAMF). (2017) Integriertes Identitätsmanagement - Plausibilisieren, Datenqualität und Sicherheitsaspekte. Einführung in die neuen IT-Tools. Schulung AVS-Mitarbeiter, Entscheider und Volljuristen. [Document uploaded in response to information request]. Frag den Staat. Retrieved January 13, 2025, from https://fragdenstaat.de/dokumente/9653-schulung_idms_bamf/.

Bundesamt für Migration und Flüchtlinge (BAMF). (2020). *Belehrung für Erstantragsteller über Mitwirkungspflichten und allgemeine Verfahrenshinweise* [Document uploaded in response to information request]. Retrieved January 13, 2025, from

https://fragdenstaat.de/anfrage/belehrung-von-antragstellenden-sprachbiometrie-dias-namenstranskription-tralita/503448/anhang/D0179.pdf.

Bundesrat. (2024). *Bericht des AK I für die 222. IMK vom 04.12.2024 bis 06. 12.2024 in Rheinsberg zu den Maßnahmen zur Beschleunigung der Asylgerichtsverfahren in den Ländern* [PDF]. Bundesrat. Retrieved February 27, 2025, from https://www.bundesrat.de/IMK/DE/termine/to-beschluesse/2024-12-04-06/TOP_08.pdf?__blob=publicationFile&v=2.

Bundeszentrale für politische Bildung (BpB). (2021). *Deutschland*. Retrieved January 13, 2025, from https://www.bpb.de/themen/migration-integration/regionalprofile/deutschland/.

Csuk, A. (2025). Können Asylentscheidungen automatisiert werden? Zu Potenzialen und Grenzen des Einsatzes von Künstlicher Intelligenz im Asylverfahren. In H. de Haas, J. Mourão Permoser, P. Raghuram, E. Recchi, V. Kourabas, W. Sievers, R. Bauböck, M. Czaika, & A. Kraler (Eds.), *Drawing boundaries and crossing borders: Migration in Theorie und Praxis* (pp. 149–168). Austrian Academy of Sciences. https://doi.org/10.1553/978OEAW95559.

Dataport. (n.d.). *KI ist die Zukunft der öffentlichen Verwaltung.* Retrieved November 5, 2024, from https://www.dataport.de/unsere-themen/kuenstliche-intelligenz/.

Deutsche Presse-Agentur (DPA). (2024). *Digitaler Klebezettel: KI hilft Verwaltungsgericht bei Asylverfahren*. Süddeutsche Zeitung. Retrieved January 13, 2025, from https://www.sueddeutsche.de/panorama/digitaler-klebezettel-ki-hilft-verwaltungsgericht-bei-asylverfahren-dpa.urn-newsml-dpa-com-20090101-240801-930-191224.

Deutscher Bundestag. (2023-a). *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Clara Bünger, Nicole Gohlke, Gökay Akbulut, weiterer Abgeordneter und der Fraktion DIE LINKE. – Drucksache 20/8586 - Einsatz von IT-Assistenzsystemen und Künstlicher Intelligenz im Bundesamt für Migration und Flüchtlinge.* (Drucksache 20/9419). Retrieved January 13, 2025, from https://dserver.bundestag.de/btd/20/094/2009419.pdf.

Deutscher Bundestag. (2023-b). *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Anke Domscheit-Berg, Dr. Petra Sitte, Nicole Gohlke, weiterer Abgeordneter und der Fraktion DIE LINKE. – Drucksache 20/6401 – Einsatz Künstlicher Intelligenz im Geschäftsbereich der Bundesregierung* (Drucksache 20/6862). Retrieved February 13, 2025, from https://dserver.bundestag.de/btd/20/068/2006862.pdf.

Deutscher Bundestag. (2022). *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Anke Domscheit-Berg, Dr. Petra Sitte, Nicole Gohlke, weiterer Abgeordneter und der Fraktion DIE LINKE. – Drucksache 20/317 – Künstliche Intelligenz im Geschäftsbereich der Bundesregierung* (Drucksache 20/430). Abgerufen am 24. Februar 2025, von https://dserver.bundestag.de/btd/20/004/2000430.pdf.

Deutscher Bundestag. (2020). *Unterrichtung durch die Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale* (Drucksache 19/23700). Retrieved January 13, 2025, from https://dserver.bundestag.de/btd/19/237/1923700.pdf.

Dimmock, M., & Fisher, A. (2017). Ethics for A-Level. In *Open Book Publishers*. https://doi.org/10.11647/obp.0125.

Emmert-Streib, F. (2024). Is ChatGPT the way toward artificial general intelligence. *Discover Artificial Intelligence, 4*(32). https://doi.org/10.1007/s44163-024-00126-3.

European Commission. (n.d.). *High-level expert group on artificial interlligence*. Digital Strategy. European Commission. Retrieved January 9, 2025, from https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.

European Commission. (2024). *Regulatory framework on artificial intelligence.* Retrieved November 5, 2024, from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

European Court of Human Rights (ECHR). (n.d.). *European Convention on Human Rights*. Retrieved January 9, 2025, from https://www.echr.coe.int/european-convention-on-human-rights.

European Migration Network (EMN). (2022). *The use of digitalisation and artificial intelligence in migration management: Joint EMN-OECD inform*. Brussels, Belgium: European Migration Network. Retrieved January 13, 2025, from https://www.oecd.org/content/dam/oecd/en/topics/policy-issues/migration/EMN-OECD-INFORM-FEB-2022-The-use-of-Digitalisation-and-AI-in-Migration-Management.pdf.

Fedyk, A., Hodson, J., Khimich, N., Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies, 27*, 938–985. https://doi.org/10.1007/s11142-022-09697-x.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication*, (2020-1).

Forster, R. (2022). *Refugee protection in the artificial intelligence era*. Chatham House. Retrieved November 4, 2024, from https://www.chathamhouse.org/sites/default/files/2022-09/2022-09-07-refugee-protection-artificial-intelligence-era-forster.pdf.

Gonzalez Riedel, D., & Idema, S. (2024). Understanding intersection between EU's AI Act and privacy compliance. *Compact*. Retrieved November 7, 2024, from https://www.compact.nl/articles/understanding-intersection-between-eus-ai-act-and-privacy-compliance/.

Heeger, V. (2024). *So setzt das Bamf Künstliche Intelligenz ein*. Tagesspiegel Background. Retrieved January 13, 2025, from https://background.tagesspiegel.de/digitalisierung-und-ki/briefing/so-setzt-das-bamf-kuenstliche-intelligenz-ein.

Herbert, U., & Schönhagen, J. (2020). *Vor dem 5. September: Die "Flüchtlingskrise" 2015 im historischen Kontext*. Bundeszentrale für politische Bildung. Retrieved January 13, 2025, from https://www.bpb.de/shop/zeitschriften/apuz/312832/vor-dem-5-september/.

Hessischer Flüchtlingsrat (hfr). (2024). *Digitale Kontrollhelfer: Von der Bezahlkarte zum Ausländerzentralregister*. Retrieved January 13, 2025, from https://fr-hessen.de/wp-

content/uploads/2024/05/2024-05-22-Digitale-Kontrollhelfer-Von-der-Bezahlkarte-zum-Auslaenderzentralregister.pdf.

Hickman, E., & Petrin, M. (2021). Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *European Business Organization Law Review*, *22*(4), 593–625. https://doi.org/10.1007/s40804-021-00224-0.

Jones, K. (2023). *AI governance and human rights*. Chatham House. Retrieved December 30, 2024, from https://www.chathamhouse.org/sites/default/files/2023-01/2023-01-10-AI-governance-human-rights-jones.pdf.

Kossakowski, S. (2024). *Scholz zu Gast beim BAMF: Mehr KI-Einsatz bei Asylverfahren.* BR24. Retrieved November 5, 2024 from https://www.br.de/nachrichten/bayern/scholz-zu-gast-beim-bamf-mehr-ki-einsatz-bei-asylverfahren,UHwOYT3.

Krennerich, M. (2024). Human rights and human rights politics. In *Interdisciplinary studies in human rights* (pp. 1–13). https://doi.org/10.1007/978-3-031-57026-1_1.

United Nations. (n.d.). *Universal Declaration of Human Rights*. United Nations. Retrieved January 9, 2025, from https://www.un.org/en/about-us/universal-declaration-of-human-rights.

LaFollette, H. (2014). *Ethics in practice : An anthology*. John Wiley & Sons, Incorporated.

Madan, R., & Ashok, M. (2023). AI adoption and diffusion in public administration: a systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), Article 101774. https://doi.org/10.1016/j.giq.2022.101774.

Migration Data Portal. (2022). *BAMF: Identitätsmanagement durch künstliche Intelligenz*. Retrieved November 5, 2024, from https://www.migrationdataportal.org/de/did-profile/bamf-identitaetsmanagement-durch-kuenstliche-intelligenz.

Ministerium der Justiz und für Migration Baden-Württemberg (MJ BW). (2024). *Einrichtung von Asylkammern an Verwaltungsgerichten* [Pressemitteilung]. Abgerufen am 27. Februar 2025, von https://jum.baden-wuerttemberg.de/de/presse-service/presse/pressemitteilung/pid/einrichtung-von-asylkammern-an-verwaltungsgerichten-1.

Molnar, P. (2021). Chapter 10: Robots and refugees: The human rights impacts of artificial intelligence and automated decision-making in migration. In *Research handbook on international migration and digital technology*. Cheltenham, UK: Edward Elgar Publishing. https://doi.org/10.4337/9781839100611.00019.

Nyman, L. (2022). Potential Legal Challenges under the European Convention on Human Rights on the Use of Artificial Intelligence in Asylum Procedures. Retrieved November 5, 2024, from https://www.diva-portal.org/smash/get/diva2:1687906/FULLTEXT01.pdf.

Oltmer, J. (2017). Germany and Global Refugees: A History of the Present. *CESifo DICE report*, 14, 26-31.

Organisation for Economic Co-operation and Development (OECD). (n.d.). *OECD AI principles*. Retrieved November 4, 2024, from https://oecd.ai/en/ai-principles.

Ozkul, D. (2023). *Automating immigration and asylum: The uses of new technologies in migration and asylum governance in Europe*. Algorithmic Fairness for Asylum Seekers and Refugees (AFAR) Project, Refugee Studies Centre, University of Oxford.

Parliamentary Assembly of the Council of Europe. (2024). *Ensuring human rights-compliant asylum procedures* (Report No. 15997). Retrieved January 13, 2025, from https://rm.coe.int/ensuring-human-rights-compliant-asylum-procedures/1680afd824.

Polok, B., el-Taj, H., & Rana, A. A. (2023). Balancing potential and peril: The ethical implications of artificial intelligence on human rights. *Multicultural Education, 9*(6), 94–101. https://doi.org/10.2139/ssrn.4484386.

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society, 40*(2), 178-193. https://doi.org/10.1080/14494035.2021.1929728.

Redaktion Filstalexpress. (2025). *Land bei Verkürzung asylgerichtlicher Verfahren in Spitzengruppe*. Filstalexpress. Retrieved February 27, 2025, from https://filstalexpress.de/politik/183584/land-bei-verkuerzung-asylgerichtlicher-verfahren-in-spitzengruppe/.

Reis, J., Amorim, M., Cohen, Y., & Rodrigues, M. (2020). Artificial intelligence in service delivery systems: A systematic literature review. In Á. Rocha, H. Adeli, L. Reis, S. Costanzo, I. Orovic, & F. Moreira (Eds.), *Trends and Innovations in Information Systems and Technologies. WorldCIST 2020.* Advances in Intelligent Systems and Computing (Vol. 1159). Springer, Cham. https://doi.org/10.1007/978-3-030-45688-7_23.

Siau, K., & Wang, W. (2020). Artificial Intelligence (AI) Ethics. *Journal of Database Management*, *31*(2), 74–87. https://doi.org/10.4018/jdm.2020040105.

Stahl, B. C., Brooks, L., Hatzakis, T., Santiago, N., & Wright, D. (2023). Exploring ethics and human rights in artificial intelligence–A Delphi study. *Technological Forecasting and Social Change*, *191*, 122502.

Twigt, M. (2023). Doing refugee right(s) with technologies? Humanitarian crises and the multiplication of "exceptional" legal states. *Refugee Survey Quarterly, 43*(1), 1–21. https://doi.org/10.1093/rsq/hdad020.

UNESCO. (2023). *Ethical impact assessment*. Global AI Ethics and Governance Observatory. Retrieved November 7, 2024, from https://unesdoc.unesco.org/ark:/48223/pf0000386276.

UNHCR. (n.d.). *The 1951 Refugee Convention*. Retrieved January 13, 2025, from https://www.unhcr.org/about-unhcr/overview/1951-refugee-convention.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. International Journal of Public Administration, 42(7), 596–615.

Yeung, K., Howes, A., & Pogrebna, G. (2020). AI governance by human rights–centered design, deliberation, and oversight: An end to ethics washing. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford handbook of ethics of AI* (online edn). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780190067397.013.5.

# 8. Appendix

## Appendix No. 1: Assessment List (Revised)

### 1. Principle: Human Agency and Oversight

The AI system should support human agency and decision-making and support the user's agency to foster a democratic and equitable society while upholding fundamental rights through human oversight.

1.1. Human Agency and Autonomy:

Question to ask:

- *Is there awareness of AI involvement by end-users or other subjects (here: by civil servants, asylum seekers)?*
- *Is it clear whether an outcome/decision is a result of an algorithmic decision?*
- *Are strategies to avoid over-reliance/disproportionate attachment/addiction by end-users or other subjects on AI systems in place?*
- *Are there steps in place to avoid AI interference in the decision-making process leading to unintended outcomes?*
- *Are there strategies to avoid manipulating human behaviour?*

1.2. Human Oversight:

Questions to ask:

- *Is there some sort of human oversight (Human-in-the-loop; Human-on-the-loop; Human-in-command)[2]?*
- *Have users (here: civil servants, applicators) been given training on how to oversee?*
- *Are there detection and response mechanisms in place for undesirable outcomes affecting end-users or other subjects (here: civil servants, asylum seekers)?*
- *Is there the possibility to stop or abort a process when needed?*

### 2. Principle: Technical Robustness and Safety

"AI systems are developed with a preventative approach to risks and that they behave reliably and as intended while minimising unintentional and unexpected harm as well as preventing it where possible." (AI HLEG, 2020, p.9)

2.1. Resilience to Attack and Security

Questions to ask:

- *Are there protections against and measures to ensure safety from risks (misuse, technical faults, cyberattacks) for humans/society?*
- *Are the systems tested for vulnerabilities? (i.e.: types of attacks like cyberattacks; types of vulnerabilities)*
- *Are there strategies in place to ensure regular updates and communication with end-users about security?*

[2] Human-in-the-loop (HITL): Human intervention in every decision cycle of the sys-tem possible (usually not possible and not desirable, s. p.27); Human-on-the-loop (HOTL): Human intervention during the design cycle of the sys-tem and monitoring of the system's operation possible; Human-in-command (HIC): Oversee all activity of the AI system and its effects and decision-making in every step possible (including broader economic, societal, legal and ethical impact)

2.2. General Safety

Questions to ask:

- *Have you identified and managed the risks (i.e. malicious use/misuse), threats (design fault, technical faults, environmental threats) and potential impacts of the AI system?*
- *Are those risks, threats and potential impacts clearly communicated to end-users and others?*
- *Are there strategies to ensure and evaluate whether the AI system is safe, reliable, and can handle changes or faults effectively?*

2.3. Accuracy[3]

Questions to ask:

- *Have you ensured the AI system uses high-quality, up-to-date data (data that is complete and representative of the environment the system will be deployed in)?*
    - "Could a low level of accuracy of the AI system result in critical, adversarial or damaging consequences?" (AI HLEG, 2020, p. 10)
- *Do you regularly check and improve the AI systems and the data's accuracy?*
- *Are end-users and/or others informed on how accurate the AI system is?*

2.4. Reliability, Fall-back Plans and Reproducibility

Questions to ask:

- *Can the AI system cause harm if it does not work properly?*
- *Are there constant mechanisms in place to check if it works as expected in different situations?*
- *Are there clear plans to handle mistakes and test the system to make sure it works reliably, even when it gives unexpected results?*

## 3. Principle: Privacy and Data Governance

Protecting privacy, a fundamental right affected by AI requires adequate data governance to ensure data quality and integrity, relevance, secure access, and processing methods that safeguard privacy.

3.1. Privacy

Questions to ask:

- *Are the impacts of the AI system on the right to privacy, the right to physical, mental and/or moral integrity and the right to data protection considered?*
    - *(Have you thought about how the AI system affects privacy and the physical, mental, and moral well-being of people?)*
- *Have you set up ways to report and address privacy issues that might come up while using the AI system?*

3.2. Data Governance

Questions to ask:

[3] "The goal of an AI model is to learn patterns that generalize well for unseen data. It is important to check if a trained AI model is performing well on unseen examples that have not been used for training the model. To do this, the model is used to predict the answer on the test dataset and then the predicted target is compared to the actual answer. The concept of accuracy is used to evaluate the predictive capability of the AI model. Informally, accuracy is the fraction of predictions the model got right. A number of metrics are used in machine learning (ML) to measure the predictive accuracy of a model. The choice of the accuracy metric to be used depends on the ML task." (AI HLEG, 2020, p. 23)

- *If the AI system uses personal data: Are there steps taken/strategies employed to ensure personal data is kept safely (i.e. oversight mechanisms for data processing, encryption, limited access, data minimisation) throughout the life cycle of the AI system?*
- *„Did you implement the right to withdraw consent, the right to object and the right to be forgotten into the development of the AI system?" (AI HLEG, 2020, p.12)*
- *Are any privacy risks considered that might emerge from the data the AI system uses or creates?*
- *Does the system align with relevant standards and protocols (i.e. ISO, IEEE)?*

## 4. Principle: Transparency

### 4.1. Traceability[4]

Questions to ask:

- *Can the AI system's decisions be traced back to the data and rules used?*
- *Are there measures to regularly check the quality of its input and output data?*
- *Do you keep clear records of the AI system's decisions to ensure transparency and build trust?*

### 4.2. Explainability[5]

**Attention – Problem:**

Blackboxes: "An explanation as to why a model has generated a particular output or decision (and what combination of input factors contributed to that) is not always possible." (AI HLEG, 2020, p. 14). → If the system still respects fundamental rights, "(…) other explainability measures (e.g. traceability, auditability and transparent communication on the AI system's capabilities) (…) may be required (…)" (ebd.).

Questions to ask:

- *Do end-users (and others?) understand how the AI system makes decisions?*
- *Do you check regularly if end-users and others (civil servants and asylum seekers) understand the AI systems' decisions?*

### 4.3. Communication

Questions to ask:

- *In the case of an interactive AI system (i.e.: chatbots, robolawyers): Do users know they are interacting with an AI system, and are its purpose, benefits, and limitations explained?*
- *Did you establish mechanisms to inform users about the purpose, criteria and limitations of the decision(s) generated by the AI system?*

## 5. Principle: Diversity, Non-Discrimination and Fairness

AI systems must promote inclusion and diversity, avoid bias or discrimination, and ensure accessibility for all people, including those with disabilities, throughout their lifecycle.

### 5.1. Avoidance of Unfair Bias

Questions to ask:

[4] "Ability to track the journey of a data input through all stages of sampling, labelling, processing and decision making". (AI HLEG, 2020, p. 29).

[5] "Feature of an AI system that is intelligible to non-experts. An AI system is intelligible if its functionality and operations can be explained non technically to a person not skilled in the art." (AI HLEG, 2020, p.26).

- *Are there steps taken to prevent unfair bias in the AI system (algorithm design, use of input and output data) by using diverse data, testing for different user groups, and monitoring the system for bias throughout its life cycle?*
- *Do you have ways to report and fix issues related to bias or unfairness*
- *Have you consulted different communities to ensure the AI system is fair for everyone?*

5.2. Accessibility[6] and Universal Design[7]

Questions to ask:

- *Is the AI system designed to be easy to use for everyone, including people with disabilities or those at risk of being left out?*
- *Did you involve different groups of users when creating the AI system and check if it could unfairly affect any specific group?*

5.3. Stakeholder Participation

Questions to ask:

"Did you consider a mechanism to include the participation of the widest range of possible stakeholders in the AI system's design and development?"

- *Are people who might be affected by the AI system consulted when designing and building it?*
- *Is there a way to get feedback from them regularly, even after the AI system is in use?*

## 6. Principle: Societal and Environmental Well-Being

AI systems should consider society, the environment, and future generations as stakeholders, ensuring they enhance well-being, respect democratic values, avoid harm, and promote sustainability while supporting global concerns like the Sustainable Development Goals.

6.1. Environmental Well-Being

Questions to ask:

- *Could the AI system harm the environment?*
- *Are there ways to measure and reduce the environmental impact of the AI system during its development, use, and entire lifecycle (i.e.: energy used, carbon emissions)?*

6.2. Impact on Work and Skills

Questions to ask:

[6] "Extent to which products, systems, services, environments and facilities can be used by people from a population with the widest range of user needs, characteristics and capabilities to achieve identified goals in identified contexts of use (which includes direct use or use supported by assistive technologies)." (AI HLEG, 2020, p.23).

[7] "Terms such as "Design for All", "Universal Design", "accessible design", "barrier-free design", "inclusive design" and "transgenerational design" are often used interchangeably with the same meaning. These concepts have been developed by different stakeholders working to deliver high levels of accessibility. A parallel development of human centred design emerged within ergonomics focusing on usability. These related concepts are expressed in the human rights perspective of the Design for All approach. The Design for All approach focuses on user involvement and experiences during the design and development process to achieve accessibility and usability. It should be applied from the earliest possible time, and throughout all stages in the life of products and services which are intended for mainstream use. A Design for All approach also focuses on user requirements and interoperability between products and services across the end-to-end chain of use to reach inclusive and non-stigmatizing solutions." (AI HLEG, 2020, p. 30)

- Are workers and their representatives informed and consulted about the AI system, how it works, and its impact on their jobs prior to its introduction?
- Were measures provided to avoid de-skilling of workers due to the AI system and to ensure the up-/re-skilling of workers to be able to use the system?
    - Was training provided to help workers learn new skills or avoid losing important ones because of the AI system? (de-skilling avoided; up/re-skilling of workers provided

### 6.3. Impact on Society at Large or Democracy

Questions to ask:

- *Were negative societal impacts assessed that affect stakeholders and society at large (not just end-users and others)?*
- *Were steps taken to minimise societal harm and to ensure the AI system does not negatively affect democratic processes?*

## 7. Principle: Accountability

"This term refers to the idea that one is responsible for their action – and as a corollary their consequences – and must be able to explain their aims, motivations, and reasons. Accountability has several dimensions. Accountability is sometimes required by law. For example, the General Data Protection Regulation (GDPR) requires organisations that process personal data to ensure security measures are in place to prevent data breaches and report if these fail. But accountability might also express an ethical standard, and fall short of legal consequences. Some tech firms that do not invest in facial recognition technology in spite of the absence of a ban or technological moratorium might do so out of ethical accountability considerations." (AI HLEG, 2020, p. 23)

### 7.1. Auditability[8]

Questions to ask:

- *Were ways created to make the AI system auditable, like tracking its development, data sources, and impacts?*
- *Can independent third parties audit the AI system?*

### 7.2. Risk Management[9]

Attention – Problem:

Tensions might rise between different aspects, leading to trade-offs which might be necessary. This requires individual evaluation*.:*

"This entails that relevant interests and values implicated by the AI system should be identified and that, if conflict arises, trade-offs should be explicitly acknowledged and evaluated in terms of their risk to safety and ethical principles, including fundamental rights. Any decision about which trade-off to make should be well reasoned and properly documented. When adverse impact occurs, accessible mechanisms should be foreseen that ensure adequate redress." (AI HLEG, 2020, p. 21).

Questions to ask:

[8] "Auditability refers to the ability of an AI system to undergo the assessment of the system's algorithms, data and design processes. This does not necessarily imply that information about business models and Intellectual Property related to the AI system must always be openly available. Ensuring traceability and logging mechanisms from the early design phase of the AI system can help enable the system's auditability." (AI HLEG, 2020, p. 25).

[9] "Both the ability to report on actions or decisions that contribute to the AI system's outcome, and to respond to the consequences of such an outcome, must be ensured. Identifying, assessing, documenting and minimising the potential negative impacts of AI systems (…). Due protection must be available for whistle-blowers, NGOs, trade unions or other entities when reporting legitimate concerns about an AI system."

- *Were third parties (i.e.: suppliers, end-users, subjects) included to assist in reporting risks, biases, vulnerabilities, or ethical concerns about the AI system (preferably even beyond the developmental phase)?*
- *Was there a risk training?*
- *Was the establishment of processes like an ethics review board or redress mechanisms to monitor risks, manage trade-offs, and ensure accountability throughout the AI system's lifecycle considered?*

## Appendix No. 2: Ethical Implications Analysis for ASS

| Ethical Principles | Seven Principles | Requirement-Breakdown | ASS |
|---|---|---|---|
| **Respect for human autonomy** | | *(Sum of T : Number of Principles)* | **1,1** |
| | ***Human agency and oversight*** | *(Sum of R : Number of Requirements) = T* | 1,1 |
| | *(Total Assessment : Number of Questions) = R* | ***Human Agency and Autonomy*** | 1,3 |
| | *(Total Assessment : Number of Questions) = R* | ***Human Oversight*** | 0,9 |
| **Prevention of harm** | | *(Sum of T : Number of Principles)* | **0,5** |
| | ***Technical robustness and safety*** | *(Sum of R : Number of Requirements) = T* | 0,6 |
| | *(Total Assessment : Number of Questions) = R* | ***Resilience to Attack and Security*** | 0,5 |
| | *(Total Assessment : Number of Questions) = R* | ***General safety*** | 0,8 |
| | *(Total Assessment : Number of Questions) = R* | ***Accuracy*** | 0,7 |
| | *(Total Assessment : Number of Questions) = R* | ***Reliability, Fall-back Plans and Reproducibility*** | 0,5 |
| | ***Privacy and Data Governance*** | *(Sum of R : Number of Requirements) = T* | 0,75 |
| | *(Total Assessment : Number of Questions) = R* | ***Privacy*** | 1 |
| | *(Total Assessment : Number of Questions) = R* | ***Data Governance*** | 0,5 |
| | ***Societal and Environmental Well-Being*** | *See Fairness* | 0,2 |
| **Fairness** | | *(Sum of T : Number of Principles)* | **0,3** |
| | ***Diversity, Non-Discrimination and Fairness*** | *(Sum of R : Number of Requirements) = T* | 0,25 |
| | *(Total Assessment : Number of Questions) = R* | ***Avoidance of Unfair Bias*** | 0,75 |
| | *(Total Assessment : Number of Questions) = R* | ***Accessibility and Universal Design*** | 0 |
| | *(Total Assessment : Number of Questions) = R* | ***Stakeholder Participation*** | 0 |

| | | | |
|---|---|---|---|
| | ***Societal and Environmental Well-Being*** | *(Sum of R : Number of Requirements) = T* | 0,2 |
| | *(Total Assessment : Number of Questions) = R* | ***Environmental Well-being*** | 0 |
| | *(Total Assessment : Number of Questions) = R* | ***Impact on Work and Skills*** | 0,5 |
| | *(Total Assessment : Number of Questions) = R* | ***Impact on Society at large or Democracy*** | 0 |
| | ***Accountability*** | *(Sum of R : Number of Requirements) = T* | 0,5 |
| | *(Total Assessment : Number of Questions) = R* | ***Auditability*** | 0,5 |
| | *(Total Assessment : Number of Questions) = R* | ***Risk Management*** | 0,5 |
| **Explicability** | | *(Sum of T : Number of Principles)* | **0,75** |
| | ***Transparency*** | *(Sum of R : Number of Requirements) = T* | 0,75 |
| | *(Total Assessment : Number of Questions) = R* | ***Traceability*** | 1 |
| | *(Total Assessment : Number of Questions) = R* | ***Explainability*** | 0,25 |
| | *(Total Assessment : Number of Questions) = R* | ***Communication*** | 1 |
| **Total Assessment Number (Max. 52)** | | | 31 |