Evaluation of Public Services Through the Lens of Digital Ethics

Narek Andreasyane^{a*}, *Daniele* Buson^b, *José* Mancera^c, *Edy* Portmann^d, *Luis* Terán^e.

^{a}PhD Student, Human-IST Institute, University of Fribourg, narek.andreasyan@unifr.ch

^{b}Master Student Lucerne Business School, daniele.buson@stud.hslu.ch

^{c}Ph.D. Candidate in Computer Science and Artificial Intelligence, University of Fribourg, jose.mancera@hslu.ch, ORCID: 0000-0003-3837-6524

^{d}Human-IST Institute, University of Fribourg, Switzerland, edy.portmann@unifr.ch, ORCID: 0000-0001-6448-1139

^{e}Lucerne University of Applied Sciences and Arts, luis.teran@unifr.ch, ORCID: 0000-0002-0503-511X

Abstract. The rapidly growing popularity of digital services requires robust frameworks to identify and address associated ethical concerns. This paper presents a structured framework for assessing ethical concerns of digital services, offering a scalable and adaptable tool to assess concerns, including data misuse, cybersecurity, transparency, inequality, and sustainability. The framework employs a customized Delphi method to gather diverse expert insights, translating them into quantifiable metrics through a mathematical model. These metrics inform structured surveys, generating actionable outputs, including visual summaries, static recommendations, and AI-driven insights. To illustrate the framework's application, we detail its implementation in the context of electronic voting (e-voting). By addressing key ethical challenges, mainly privacy, transparency, and inclusivity, this use case demonstrates the framework's utility in analyzing complex digital services. The study highlights the importance of balancing technological innovation with ethical accountability, providing a practical approach to ensuring transparency and trust in public digital services.

Keywords. digital ethics, humanistic, AI ethics, ethical concerns, LLM, NLP

1. Introduction

As digital services become integral to modern society, addressing their ethical implications has become increasingly critical. For example the Facebook–Cambridge Analytica scandal in 2018 highlighted major ethical concerns regarding privacy and consent in digital services. Personal data from millions of Facebook users was collected without their explicit permission and used for political advertising, sparking global debates about data misuse and democratic integrity (The Guardian, 2018). Similarly, Google faced criticism when it was discovered that their Street View cars collected data from unsecured Wi-Fi networks during their image capture process. This raised serious privacy concerns, as users were unaware that their data was being intercepted and stored (The Guardian, 2010; Wired, 2012). Another significant case is Clearview AI, which built a powerful facial recognition tool by scraping billions of images from social media platforms without user consent. This practice ignited concerns over mass surveillance, potential misuse by law enforcement, and the lack of transparency in how facial recognition data is gathered and applied (Hart, 2024; The Guardian, 2022). These cases collectively emphasize the urgent need for ethical oversight, informed consent, and accountability in digital services.

This requires a comprehensive approach that spans the entire digital service lifecycle, encompassing design, development, and deployment, as outlined by (Cath et al., 2018). Building on these foundations, this paper

presents a framework for assessing ethical concerns in digital services. The framework enables organizations to identify risks, implement actionable improvements, and expand on previous research in the field (Andreasyan et al., 2024; Teran et al., 2021; Wallimann-Helmer et al., 2021). The proposed framework utilizes a customized Delphi method (see subsection 3.2 and also previous work (Andreasyan et al., 2024)), fostering iterative discussions among experts to capture diverse perspectives rather than consensus. Unlike the traditional Delphi method, our version does not aim to achieve consensus among participants but rather to foster diverse perspectives that allow participants to reconsider their conclusions after hearing others' viewpoints. These expert insights inform structured surveys that allow service providers, such as Swiss Post employees, to evaluate their services against ethical criteria derived from expert input and scientific literature. A web-based application processes the survey responses to generate practical outputs, including visual summaries, tailored recommendations, and insights from a domain-specific large language model (LLM) (future development). These tools assist decision-makers in prioritizing and addressing ethical challenges. The framework serves two main purposes: evaluating ethical considerations for proposed digital services and diagnosing ethical issues in existing ones.

By focusing on critical ethical factors such as privacy, transparency, trust, and inclusivity, the framework promotes ethical accountability and fosters public trust in digital governance. These general ethical concerns (see subsection 4.2) align with the discussions in (Nabbosa and Kaar, 2020; Schoentgen and Wilkinson, 2021), which explore ethical issues in the digital era and societal digitalization. Figure 1 shows the results, namely consensus versus dissensus of the 75 responses of two customized Delphi methods after two rounds.

The results of these discussions are then integrated into a survey tool, where service providers, such as Swiss Post employees, can evaluate their own services based on ethical concerns raised by the experts and literature. A web-based application aggregates the survey responses and generates an output in the form of graphical summaries, recommendations with cited sources, and access to an LLM trained in the relevant domain to offer further insights into potential issues. Figure 2 illustrates the process flow of our framework. It provides an overview diagram highlighting key steps, including literature review, ethics assessment framework development, iterative evaluation, and finally generate ethical insights.
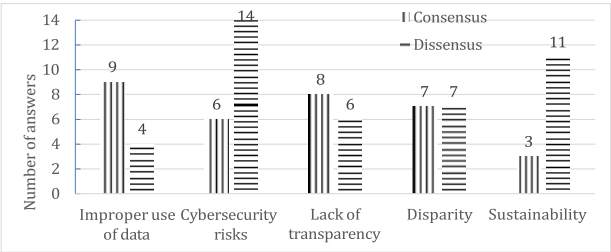

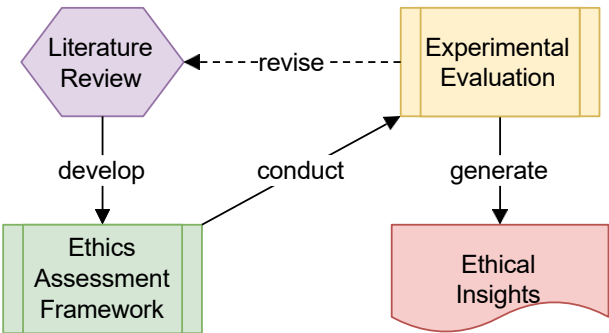Fig. 1 – Survey results for two rounded customized Delphi method.

Fig. 2 – Process flow for identifying and analyzing ethical concerns in public digital services.

1.1. Public Services

In digital governance, public services stand at the forefront of societal well-being, requiring efficient delivery and a strong ethical foundation. As governments and organizations adopt new technologies to administer services—ranging from healthcare to education—balancing innovation with robust ethical standards becomes paramount, which is stated in international standard (Olszewska et al., 2022), that it is crucial to address ethical concerns during and from the beginning of the digital platform's development. As highlighted by (Warner, 2023), public services play a crucial role in upholding societal values, necessitating transparent, inclusive, and accountable practices to maintain public trust.

We first examine the broader landscape of public services and emphasize the importance of aligning technological innovation with ethical principles to foster trust. We then delve into a specific application, the e-voting use case, which illustrates how our ethical evaluation framework can be adapted to safeguard critical democratic processes in the digital era. Future work will be dedicated to analyzing and implementing other public services (e.g., health and logistics, among others).

1.2. Evaluation Framework for Ethical Digital Services

We outline our ongoing efforts to develop a robust framework for evaluating the ethical concerns associated with digital services. This initiative builds upon our previous work, as outlined in (Andreasyan et al., 2024), which presented the theoretical foundation for addressing digital ethics challenges in public services, explicitly using the case of electronic voting. The current framework, conducted in collaboration with our implementation partner, Swiss Post, reflects our shared commitment to enhancing the transparency and trustworthiness of public digital services. Central to this mission is integrating expert insights and empirical data into a structured analytical process that supports creating and maintaining ethically sound digital solutions.

Our methodology begins with a diverse panel of experts participating in a modified Delphi method. Unlike traditional consensus-driven approaches, our adaptation of the Delphi method actively embraces a wide array of perspectives to ensure that differing viewpoints are captured and considered. This iterative process ensures that ethical standards are rigorously debated and refined. Following these expert panels, service providers complete structured surveys based on the ethical standards identified during the Delphi process. Responses are then processed through a Web-based application that consolidates the data into actionable insights, visual summaries, and tailored recommendations.

The systematic nature of our framework is designed to strengthen public trust in digital services, particularly within sensitive domains such as healthcare, education, and governance. The importance of ethical considerations in these areas is well-documented in the literature, with notable examples including the ethical use of digital trace data in education (Hakimi et al., 2021), the challenges of applying digital twins for personalized healthcare services (Huang et al., 2022), and the broader implications of digital innovation in public administration (Pakhnenko, Kuan, et al., 2023). Ethical lapses in these domains can significantly erode public confidence, making addressing these concerns from the design phase to deployment critical.

Through a cycle of expert-driven inquiry, data-gathering, and targeted recommendations, our framework aims to establish a scalable and adaptable model for ethical oversight that can be broadly applied to a diverse range of public digital services. By embedding ethical considerations into every stage of the process, we strive to create digital solutions that are functional, efficient, and aligned with the principles of fairness, transparency, and accountability.

1.3. Use Case: E-Voting

The first use case of our framework is in the context of e-voting systems by Swiss Post, a domain where ethical considerations are crucial. Electronic voting systems' integrity, transparency, and security are essential for protecting democratic processes and ensuring public confidence in electoral outcomes. By applying our framework to e-voting, we can identify and mitigate potential ethical risks such as data privacy breaches, voter manipulation, or transparency deficiencies.

In collaboration with the Swiss Post, we are currently developing this framework with their e-voting experts to evaluate and enhance the ethical aspects of their digital voting services. The process involves an iterative feedback cycle and refinement to adapt the framework precisely to the complexities of e-voting. This use case not only serves to demonstrate the framework's effectiveness but also sets a precedent for other digital services, such as e-health and logistics.

Through the e-voting pilot, we aim to establish a model for ethical digital governance that can be replicated and scaled across various public services. By focusing on critical issues such as privacy, security, and trust, the framework supports creating e-voting systems that voters can trust. Moreover, this initiative aligns with global efforts to strengthen the resilience of democratic institutions against the challenges posed by digitization and cyber threats.

In conclusion, the development and application of our ethical evaluation framework represent a significant advancement in the governance of digital services. By proactively addressing ethical concerns, we can safeguard public trust and ensure that digital innovations in public services, like e-voting, are beneficial and secure.

Ultimately, this work addresses the following research questions: What ethical concerns and general ethical factors are important in the e-voting use case? How can a theoretical framework and user-friendly prototype address those ethical concerns and general factors?

2. Related Works

As digital services shape modern life, ethical design has become crucial. This section outlines advancements in evaluating digital ethics and reviews frameworks that offer structured approaches to addressing ethical challenges. These tools provide practical insights into embedding ethical principles in digital service design.

Our tool, the Swiss Digital Ethics Compass (SDEC), differs from others by focusing on twelve ethical concerns rather than general ethical factors. This approach builds on literature emphasizing ethics in technology design and governance.

Two prominent frameworks informed our work. The TEDS framework (Joisten et al., 2022) focuses on "technoethical" concerns, especially around data privacy and breaches. The AI Digital Tool Product Lifecycle Governance Framework (Ortega et al., 2023) emphasizes embedding ethics and compliance by design through human-centered principles and risk management.

Other valuable contribution includes the Digital Ethics Canvas (Hardebolle et al., 2023), which evaluates ethical risks through six key principles (e.g., beneficence, fairness, sustainability) (see Figure 3). It was positively received in evaluations with novices (N=26) and experts (N=16). The Ethics Canvas (Lukianets et al., 2021), inspired by the Business Model Canvas (Osterwalder and Pigneur, 2010), guides users in identifying stakeholders and ethical concerns across three stages. The Data Ethics Canvas[1] helps organizations navigate ethical data use, while the Ethics Canvas[2] promotes responsible innovation across various project types.

Fig. 3 – Digital ethics canvas (Hardebolle et al., 2023).

The ALTAI framework (on AI, 2020) offers a self-assessment for AI systems based on seven key requirements like oversight, safety, and transparency. In contrast, the AI4Belgium tool[3] focuses on an organization's readiness for AI deployment, covering infrastructure and strategic alignment.

In the Figure 4 is presented the AI System Ethics Self-Assessment Tool (Dubai, 2023) from Digital Dubai, which evaluates digital systems based on fairness, accountability, transparency, and explainability. It produces an ethics score based on system impact and mitigation measures.

---

[1]Data Ethics Canvas by the Open Data Institute (ODI): https://github.com/theodi/data-ethics-canvas

[2]The Online Ethics Canvas: https://www.ethicscanvas.org

[3]Online platform for assessing the trustworthiness of AI implementation: https://altai.ai4belgium.be/
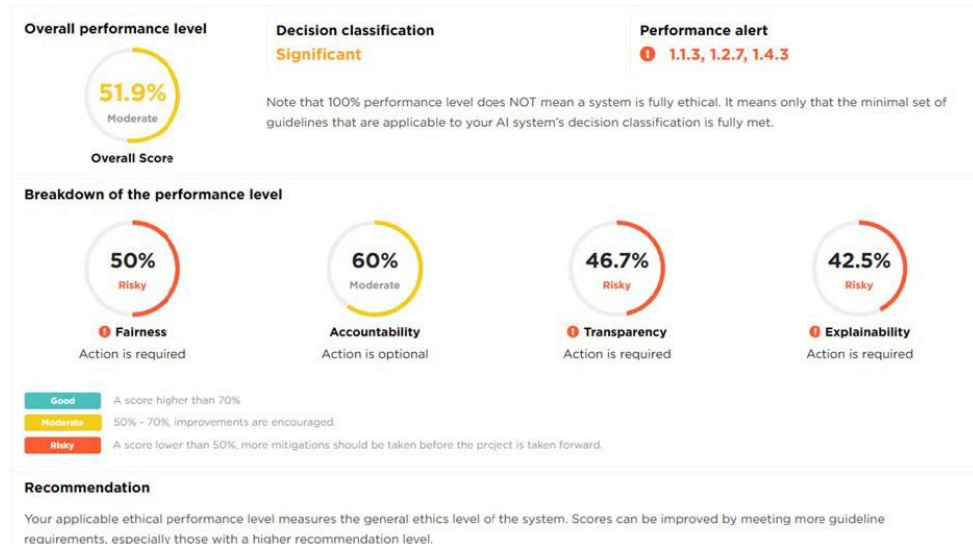
Fig. 4 – AI System Ethics Self-Assessment Tool (Dubai, 2023).

Our approach stands out by analyzing twelve ethical concerns found in our previous work (Andreasyan et al., 2024; Teran et al., 2021; Wallimann-Helmer et al., 2021):

- **Autonomy:** lack of decision-making ability due to external influence.
- **Discrimination:** unjustified differentiation for irrelevant reasons.
- **Domination:** actual or potential control over others.
- **Exclusion:** unfair denial of access to benefits.
- **Exploitation:** unfair advantage taken of individuals or groups.
- **Inequality:** unjustified unequal treatment or outcomes.
- **Justice:** biased or unfair distribution.
- **Privacy:** exposure without adequate protection.
- **Responsibility:** failure to acknowledge or prepare for negative impacts.
- **Trust:** weakened relationships among stakeholders.
- **Dignity:** disregard for human respect and rights.
- **Truth:** misrepresentation of identity or actions.

This approach enables deeper, context-specific evaluation of AI impacts, enhancing accountability and inclusivity. By addressing specific concerns often overlooked in traditional models, our framework promotes responsible and ethically grounded public digital systems.

3. Framework

This section outlines a comprehensive three-step process for evaluating ethical concerns in digital services. By integrating expert consultations, structured surveys, and advanced analytical tools, this framework ensures a robust assessment of critical issues such as privacy, transparency, and security. Leveraging proper methodologies such as the Delphi method and interactive web technologies facilitates informed decision-making while promoting explainability and transparency. The subsequent sections detail each step, supported by visual and technical insights, enhancing the ethical evaluation process. An overview of the process is provided in Figure 5.

3.1. System Architecture

Figure 6 illustrates the system architecture organized into two primary components: the user and technical layers.
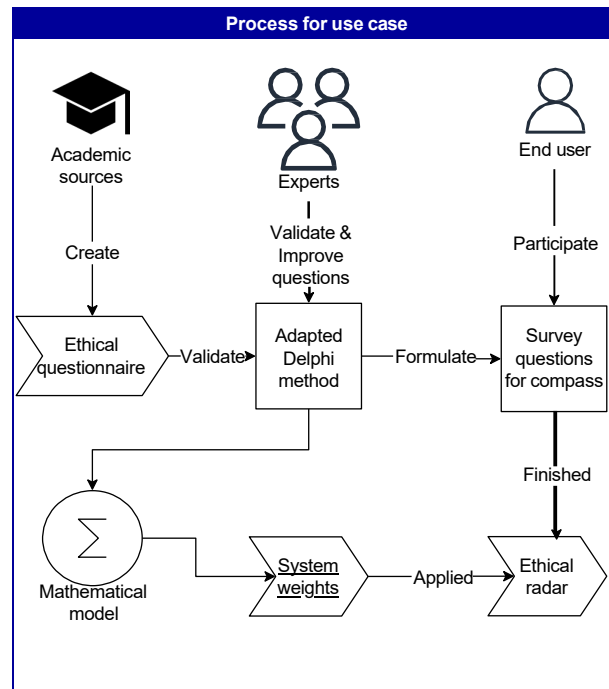
Fig. 5 – Overview of the framework's process for identifying and addressing ethical concerns in public digital services.

User Layer. In the user layer, a request and response system enables users, particularly internal Swiss Post employees, to complete surveys or seek clarification on ethical topics. The system then processes these inquiries and generates responses that assess ethical concerns and provide static recommendations. Additionally, users can interact with generative artificial intelligence (GenAI) to get extra information about the use case and digital ethics.
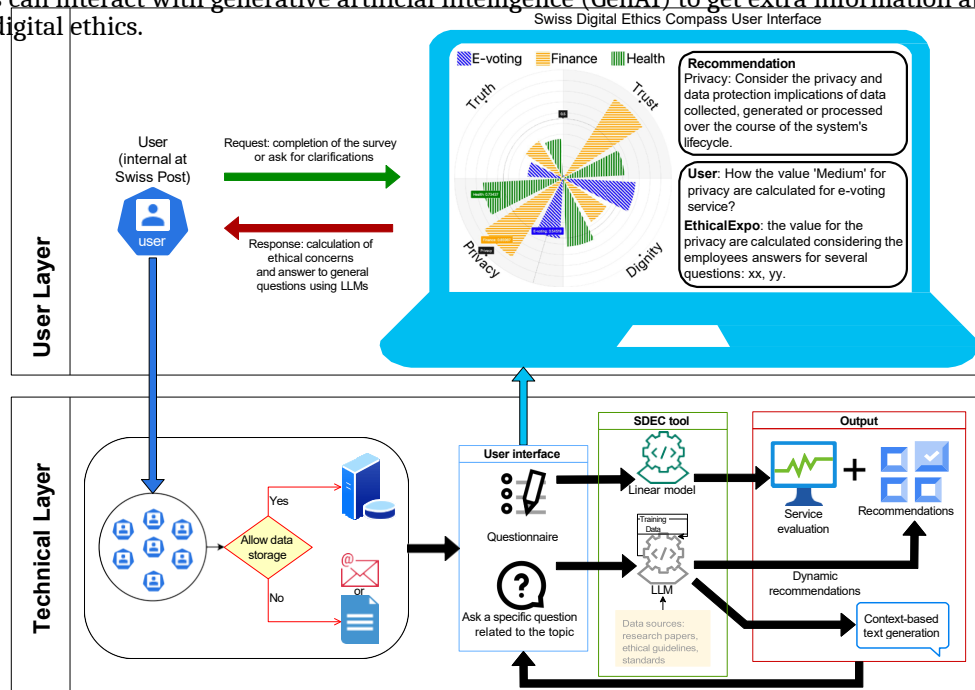
Fig. 6 – User and technical layers. Mock-up for the graphical interface of the smart radar, including static and dynamic recommendations.

Technical Layer. The technical layer begins with the user interface, where users can ask questions about ethical and compliance topics. The system employs a questionnaire format to capture the context of each query, prompting users to provide details relevant to their concerns. This information flows into an output module, which generates answers based on contextual text production, service evaluation, and dynamic recommendations. The data sources feeding into the GenAI modules include research papers, ethical guidelines, and standards, which were extracted from digital libraries such as ACM, IEEE, PhilPapers, and SprinerLink between 2010 and 2023 (Andreasyan et al., 2024; Teran et al., 2021). On top of LLM, retrieval-augmented generation (RAG) was built on academic literature identified within the development of this project. It interprets user questions and assists in ethical evaluations, while linear models are employed to refine response accuracy further. A recommendation component enhances this layer by giving users insights about privacy implications based on collected data. The framework is designed to support explainability and transparency, offering Swiss Post's internal users clear and contextualized responses to ethical questions related to data privacy and ethical standards. For instance, if a user would like to know more details on how the privacy rating of "medium" is determined for an e-voting service, the system provides a structured response. It includes referencing employee answers to relevant questions and offering transparency on how privacy values are computed.

LLM Framework Selection. Selecting an LLM framework was crucial for this project. Transparency and control were key drivers, leading to the adoption of an open-source LLM. This choice ensures visibility into the model's processes, allowing us to understand content generation and address ethical concerns. Such transparency aligns with our ethical AI practices and supports collaboration with the wider AI community. Data security and privacy were also vital. To avoid transmitting sensitive data to third-party providers, we chose a locally deployed solution. This minimizes the risk of data breaches and supports compliance with data protection regulations. Open-source LLMs

Tab. 1 – Licensing and adaptability of various LLMs (* - with extra conditions) (MindsDB, 2024; UbiOps, 2024).

LLM	License	Commercial?	Adaptable?
Mistral	Apache-2.0	Yes	Yes
GPT-4	Closed Source	No	No
GPT-3	Closed Source	No	No
GPT-2	MIT	Yes	Yes
LLaMa 2	Custom LLaMa 2	Yes	Yes*
BART	Apache-2.0	Yes	Yes
BERT	Apache-2.0	Yes	Yes
Falcon	Apache-2.0	Yes	Yes
Falcon-180B	Custom Falcon	Yes	Yes*

offer technical flexibility, enabling customization to meet our ethical standards, such as allowing users to verify AI-generated content and supporting human oversight. Additionally, this approach avoids reliance on third-party infrastructure, mitigating risks related to data sovereignty and provider-imposed changes. In summary, our choice of an open-source LLM was based on transparency, privacy, flexibility, and control (Kukreja et al., 2024; Manchanda et al., 2024). This allows us to build a secure, customizable, and ethically aligned system. For the SDEC implementation, we choose the Mistral LLM framework[4].

Mistral Technical Specifications. Mistral is a decoder-only, Apache-2.0 licensed model, making it commercially usable and highly adaptable. With parameter sizes of 7.24B and 46.7B, it balances performance and efficiency, competing with models like Falcon and LLaMa 2 while avoiding licensing constraints. Unlike closed-source models like GPT-3/4, Mistral enables customization and transparency, as detailed in Table 1. Ethically, Mistral's open license supports independent audits, bias detection, and safety improvements, in contrast

Tab. 2 – Types and parameter sizes of various LLMs (MindsDB, 2024; UbiOps, 2024).

LLM	Type	# of parameters
Mistral	Decoder-only	7.24B, 46.7B
GPT	Decoder-only	124M, 350M, 760M, 1.3B, 2.7B, 6.7B
LLaMa	Decoder-only	6.74B, 70B
BART	Encoder-decoder	139M, 406M
BERT	Encoder-only	110M, 336M
Falcon	Decoder-only	7B, 40B

to proprietary models. It avoids the restrictive terms found in models like LLaMa 2 and Falcon-180B, enabling ethical development and reducing risks of commercial lock-in. Mistral allows responsible fine-tuning to address bias and fairness. Its parameter efficiency supports sustainable AI development and makes it suitable

---

[4]Mistral LLM: https://mistral.ai/news/mistral-nemo

for deployment in resource-limited environments Table 2. In summary, Mistral's open-source nature, adaptability, and strong technical and ethical features position it as a superior alternative to closed or restricted models (Poszler et al., 2024).

3.2. Workshops with Experts

The experts have been chosen in the following way to provide diverse opinions. P2 and P5 was directly related to the e-voting team, but the rest were from different departments (see list of expert profiles in Table 3). We begin with expert consultations , structured around a customized version of the Delphi method, as illustrated in Figure 7. The primary objective is to engage specialists in fields relevant to the digital service under analysis (e.g.,

Tab. 3 – Background of experts in the survey.

ID	Profession
P1	Communication Specialist E-Government
P2	Team E-Voting Product Manager
P3	Accessibility and Corporate Social Responsibility Specialist
P4	Chief Information Security Officer
P5	Information Security Officer Trusted Interaction Services, E-Voting
P6	Professor on Soft Computing

e-voting) and foster an exchange of diverse perspectives. Particular attention was paid to ensure the quality and reliability of the consultations and to avoid potential pitfalls associated with expert interviews, as highlighted in (Döringer, 2021).

To accommodate the requirements of digital environments and mitigate time constraints, we adapted the Delphi method to be conducted partially online. This adaptation provides significant flexibility, allowing participants to contribute asynchronously from various geographic locations while maintaining the method's iterative and collaborative nature. The Delphi method offers several distinct advantages, especially when research is scarce, ethically, or logistically complex or where conflicting evidence exists. As noted in (Nasa et al., 2021) and their work on its application within the healthcare sector, the Delphi method has become a critical tool for developing best practice guidelines through collective intelligence over the past decades. While in (Nasa et al., 2021), the authors emphasize consensus-building, our adaptation shifts the focus to capture a broad range of expert opinions. After each round, participants are encouraged to revisit and refine their views, addressing limitations identified in the literature (Nasa et al., 2021) and ensuring a robust collection of insights.

Fig. 7 – Workflow of our customized Delphi method.

For instance, in analyzing ethical concerns for e-voting services, experts were invited to deliberate on critical issues such as data privacy, voter anonymity, and fraud risks. These discussions were further enriched with insights from academic literature, including the emphasis on transparency, as discussed in (Enguehard, 2014). Together, these expert contributions and literature-based considerations form the foundation for the subsequent stages of the framework.

3.3. Survey Creation and Implementation

Once the expert discussions are complete, the concerns raised are distilled into a comprehensive survey aimed at service providers. This survey evaluates how the digital service is implemented from an ethical perspective. The questions cover multiple ethical factors related to the service, including the following:

- How is data privacy managed in the service?

- What measures are in place to ensure transparency?
- How does the service protect against malicious misuse?

The questions are grounded in expert opinions and existing literature to ensure a well-rounded assessment. End users, such as Swiss Post employees, interact with the survey through a web application designed specifically for this purpose. This app provides real-time feedback as the user answers each question.

3.4. Analysis and Recommendations

After completing the survey, users receive an analysis of their service's ethical concerns. The analysis includes:

- **Visualised Data**: A chart that summarises the levels of concern across different ethical domains (e.g., privacy, transparency, security).
- **Static Recommendations**: A list of suggestions for mitigating the identified risks, accompanied by citations from the relevant literature.
- **AI-driven Insights**: Users can use a service-trained LLM to explore further ethical implications and ask questions about specific issues.

Figure 8 provides an overview of these components in the evaluation process. The technical backbone of this framework involves leveraging pre-assessment surveys and interactive web technologies to deliver a seamless user experience.

An illustrative example from the forty-question framework demonstrates the computation of values for the smart radar is presented in Figure 6. To illustrate, consider the following question with predefined response options: *"Which authentication method is utilized in the service?"* The available choices are:

1. One-Time Password (OTP), level 2 domination, level 3 privacy
2. Multi-Factor Authentication (MFA), level 4 domination, level 5 privacy
3. Blockchain Authentication, level 3 domination, level 3 privacy
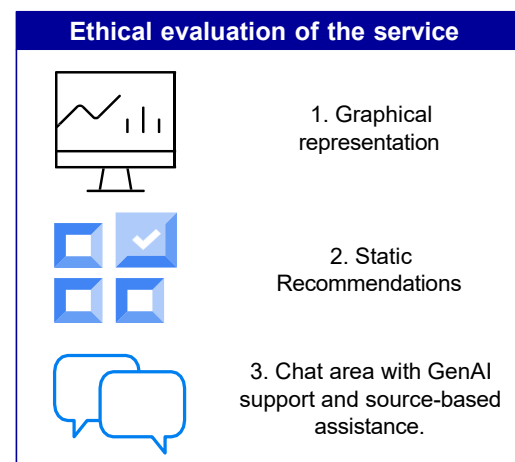4. Biometric Authentication, level 8 domination, level 7 privacy
5. None

Ethical evaluation of the service

1. Graphical representation

2. Static Recommendations

3. Chat area with GenAI support and source-based assistance.

Fig. 8 – Overview of the ethical evaluation process, including graphical representation, recommendations, and AI-supported insights.

Each authentication mechanism influences ethical concerns, such as privacy, trust, and domination, as substantiated by scholarly literature (Druliac et al., 2024; Lauer, 2004; Trechsel et al., 2016). This question is accompanied by predefined static recommendations designed to enhance ethical compliance. Users receive tailored recommendations if the cumulative calculation across all questions exceeds a defined threshold (currently set at 0.6 out of 1). For example:

1. **Enhancing Privacy Protections:** To mitigate privacy risks, ensure that any personal data collected through MFA processes, such as phone numbers or email addresses, undergoes encryption and secure storage. This measure bolsters voter confidentiality and strengthens trust in the system (Lauer, 2004).
2. **Implementing Secure Authentication Mechanisms:** Combine multi-factor authentication (MFA) with a lightweight token-based OTP system to enhance security. Additionally, incorporate privacy-preserving techniques, such as hashing user identifiers, to reduce privacy threats (Druliac et al., 2024).
3. **Ensuring Secure Data Storage:** All personal data gathered during authentication must be encrypted and securely stored, ensuring confidentiality and integrity (Trechsel et al., 2016).

When users provide answers to questions, their selected responses determine concern levels associated with specific ethical factors. Based on expert opinions and literature, a concern score (1 to 10) is assigned to each answer for relevant ethical concerns. In this context, domination and privacy are two concerns commonly evaluated for such questions.

The scores are aggregated following a systematic procedure. The concern score contributes to a cumulative sum for each question for the respective ethical concern. The overall evaluation is calculated by summing the concern scores for each question and dividing by the total possible score. Assuming there are 20 questions, each with a maximum score of 10, the total maximum score for each concern is 200. If the cumulative score for privacy is 120 and for domination is 80, the evaluation is computed as follows:

$$\text{Evaluation (Privacy)} = \frac{120}{200} \times 100 = 60\% \tag{1}$$

$$\text{Evaluation (Domination)} = \frac{80}{200} \times 100 = 40\% \tag{2}$$

Based on the resulting percentage, the framework assigns fuzzy linguistic variables (e.g., *very high, high, neutral, low, very low*), which allow for more flexible and human-like reasoning by representing qualitative assessments rather than rigid numerical inputs. Figure 9 shows an example of the interface and the evaluation of ethical concern "exclusion" equal to "high risk." This is particularly useful in complex systems where expert judgments vary or when dealing with vague concepts such as ethics, trust, or risk.

By leveraging fuzzy logic, linguistic variables enhance interpretability and adaptability, leading to more realistic and context-aware evaluations (Zadeh, 1965, 2008). Furthermore, in Figure 9, a colored gradient is presented from *very low* to *very high* using a fuzzy index between (0:1). It indicates the overall ethnicity of the public digital services (see additionally Figure 6).

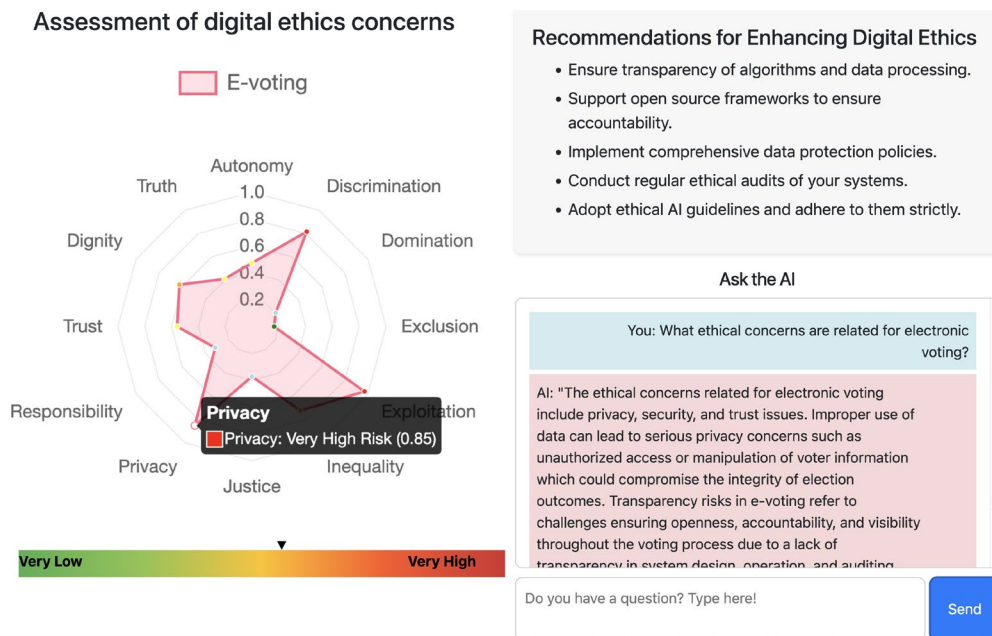**Fig. 9** – Swiss Digital Ethics Compass Interface.

Special Case Considerations: Not all questions adhere to the additive scoring model described in subsection 4.3. Certain questions are designed to assess the issue's magnitude rather than directly contribute to cumulative scores. For instance, consider the question: *"Specify the total number of eligible voters:"*

1. Less than 1k
2. 1k - 10k

3. 10k - 100k
4. 100k - 1M
5. More than 1M
6. None

This question evaluates the scale of the population affected and is used as a multiplier for related questions. By incorporating this proportionality, an issue's ethicality is scaled to reflect its broader societal impact.

The final scores for the smart radar are computed using the mathematical formula detailed in subsection 4.3 (see Equation 3). This comprehensive approach ensures that qualitative and quantitative ethicality dimensions are captured accurately. The threshold, which is set at the moment 0.6 out of 1, then the user gets the following recommendation:

1. Enhancing Privacy Protections: To mitigate privacy risks, ensure that any personal data collected through MFA processes, such as phone numbers or email addresses, undergoes encryption and is securely stored. This measure reinforces voter confidentiality and strengthens trust in the system's privacy safeguards (Lauer, 2004).
2. Implementing Secure Authentication Mechanisms: Employ multi-factor authentication (MFA) with a lightweight token-based OTP system to fortify security. Additionally, integrate privacy-preserving techniques, such as hashing user identifiers (e.g., phone numbers), to minimize exposure to privacy threats (Druliac et al., 2024).
3. Ensuring Secure Data Storage: All personal data gathered during the authentication process should be encrypted and stored securely, fostering trust by guaranteeing the confidentiality and integrity of voter information (Trechsel et al., 2016).

Once the user answers the question, he/she provides the system with a fact. Based on this fact, the level of concern is extracted. In this case, we use a table that maps each answer to the relative concern level from 1 to 10 based on experts' opinions, literature research, and logic. The values for smart radar are calculated utilizing a mathematical formula (described in subsection 4.3, Equation 3).

All final questions and static recommendations for the Compass can be found in the GitHub repository: [5].

4. Mathematical Model Development and Application of the Delphi Method

This framework employs a structured methodology combining mathematical modeling and the Delphi method to evaluate and address ethical concerns in digital services, specifically on e-voting systems. By integrating expert input, literature reviews, and iterative refinement, we quantify ethical challenges across five general ethical factors: improper use of data, cybersecurity risks, lack of transparency, disparity of treatment, and sustainability. The resulting framework provides actionable insights for stakeholders to improve service integrity (see subsection 4.2).

4.1. Methodological Overview

Our approach involves the following key steps (additionally, see Figure 5):

1. **Literature Review**: A comprehensive review of academic and industry sources was conducted to identify and understand ethical concerns in digital services. It included broad research on ethics in digital contexts and a focused analysis of e-voting systems. Robinson et al. explored ethical issues in e-voting security analysis, emphasizing vulnerabilities in system architectures (Robinson and Halderman, 2011). The authors outlined fundamental principles and requirements for a secure e-voting system, providing a framework for ethical considerations in digital voting (Gritzalis, 2002).
2. **Expert Discussions Using the Delphi Method**: Partnering with Swiss Post, we engaged six experts with diverse backgrounds. Leveraging the Delphi method (Linstone, Turoff, et al., 1975), a structured approach to achieving consensus among experts, we facilitated rounds of discussion to identify and evaluate critical ethical issues. The method, known for its iterative process of refining expert opinions (Dalkey and Helmer, 1963), enabled the experts to rate concerns on a scale from 1 (no concern) to 9 (maximum concern). It provided a robust, quantitative foundation for the model.

---

[5]Additional materials for the project: https://github.com/SwissDigitalEthicsCompass

3. **Weighted Evaluation System**: Based on expert input, we developed a mathematical model that calculates concern levels using weighted averages. These weights reflect the severity and importance of ethical risks and were normalized between 0 and 1 for consistency.
4. **Identification of Risk Scenarios**: A subsequent literature review focused on the e-voting domain to identify potential risk scenarios. For example, in (Lauer, 2004), the authors discussed the inherent risks of e-voting, highlighting the potential for breaches in security and voter trust. Insights from this review informed the development of targeted survey questions, ensuring the survey addressed critical vulnerabilities and ethical considerations in e-voting systems.
5. **User Survey Development**: A survey was created with 40 questions designed to evaluate the ethical and practical implications of e-voting services. In developing the survey, we tried to adhere to literature-proven methodologies for survey design and evaluation (Fowler Jr and Cosenza, 2009), ensuring a rigorous approach. The survey aims to collect data from end users and service developers, providing a foundation for refining the evaluation and addressing critical concerns in e-voting systems.
6. **Scoring and Concern Levels**: Survey responses are analyzed against a predefined table, assigning concern levels to each answer. These levels are aggregated to provide an overall evaluation of the service.
7. **Results Presentation**: Results are displayed on a dedicated website, offering static recommendations based on survey outcomes and a dynamic interface powered by a language model trained on scientific literature. The findings are also visualized using radar charts to highlight key concerns.

4.2. General Ethical Factors

The general factors are defined as five macro areas that serve as the foundation of our investigation (Andreasyan et al., 2024). This categorization was chosen to facilitate understanding and to group ethical concerns into specific categories, enabling a more focused and contextualized study. These factors are sufficiently broad to encapsulate all the ethical issues associated with e-voting systems. However, as noted in the conclusion in section 6, this framework is part of an iterative process and is subject to ongoing refinement. Future iterations, informed by expert input, may expand the number of factors to better address specific use case requirements.

Improper Use of Data. This factor evaluates risks related to data autonomy, privacy, and potential misuse of personal information. Expert ratings were used to develop a scale measuring the severity of data misuse concerns.

Cybersecurity Risks. Cybersecurity was assessed based on its role in ensuring privacy, trust, and fraud prevention. Weighted scores highlight the importance of robust authentication mechanisms and organizational accountability.

Lack of Transparency. Transparent communication and system comprehensibility were identified as critical factors for fostering trust and preventing exploitation. Numerical assessments informed recommendations for enhancing public trust in e-voting systems.

Disparity of Treatment. This factor addresses ethical concerns from unequal treatment, such as the digital divide and systemic exclusion. The model integrates expert evaluations to promote fairness and inclusivity.

Sustainability. Experts assessed environmental and economic considerations, including material reuse and cost efficiency. Sustainability scores were incorporated to balance ethical and practical priorities in e-voting systems.

4.3. Scoring and Explanation

The framework's scoring system for the end-user survey uses a formula designed to aggregate the concern levels derived from user responses. This method ensures that the results are normalized and accurately represent ethical risks. The calculation is defined as follows.

$$S = \frac{\sum_{i=1}^{n} C_i}{\sum_{i=1}^{n} M_i}. \tag{3}$$

where S is the normalized concern score. C_i represents the concern level for the ith question, derived from the user's response. M_i is the maximum possible concern level for the ith question. n is the total number of questions answered by the user.

Explanation. The numerator ($\sum_{i=1}^{n} C_i$) sums up the concern levels indicated by the user's responses across all survey questions. These levels are derived from predefined scales (e.g., 1 to 9) that measure the severity of ethical risks. The denominator ($\sum_{i=1}^{n} M_i$) represents the sum of the maximum possible concern levels for the same set of questions. The formula normalizes the score by dividing the actual concern levels by the maximum potential concern levels, ensuring it falls within a range of 0 to 1. A higher score indicates more significant ethical concerns across the evaluated dimensions. This normalized score provides a concise and comparable measure of ethical risks, enabling stakeholders to prioritize areas requiring intervention or improvement.

5. Results

We employed a customized Delphi method to explore ethical concerns in public digital services, focusing on an e-voting use case and engaging expert opinions through iterative discussions. This approach was designed to capture a broad range of insights rather than reaching a consensus, prioritizing diverse perspectives that could shape a nuanced ethical analysis. Experts evaluated ethical general factors such as improper use of data, cybersecurity risks, lack of transparency, disparity, and sustainability with ratings assigned to each ethical concern based on their perceived significance within the context of Swiss Post's e-voting service (see subsection 4.2). This method allowed experts to highlight key aspects, such as privacy, fairness, and transparency, regarding their impact on public trust and data security.

Further, after implementing the Delphi method, the findings were integrated into a structured survey, specifically assessing digital services on ethical concerns. Experts rated various concerns, from minimal to high significance, identifying risks like inequality, discrimination, and environmental sustainability related to e-voting, such as unequal access to digital infrastructure and the environmental footprint of on-site voting. These ratings contributed weighted factors that were essential in building a final model, a straightforward linear framework assigning weights to each question, enabling the classification of ethical risks. The resulting model, a functional tool for ethical analysis, offers quantifiable measures of ethical compliance, allowing Swiss Post to prioritize critical ethical factors within digital voting services. By systematically processing expert insights into weighted survey responses, we developed a robust framework highlighting ethical concerns, providing a scalable and adaptable assessment tool for evaluating ethical risks across other sectors and digital applications.

6. Conclusion

Our framework is currently in the research phase, with initial testing focused on Swiss Post employees as our target users. We are continuously refining the expert discussion process to ensure it yields actionable insights, and we plan to enhance the survey tool to allow for more robust data analysis. In the future, we aim to expand the framework to a broader audience, allowing for assessing ethical concerns in various digital services. **Iterative Refinement and Broader Applications**. While already effective for e-voting use cases, the current framework is designed for iterative improvement. Future iterations will include the following.

- **Expand Expert Participation**: Increasing the number and diversity of experts to reduce bias and enhance the model's robustness.
- **Refine Weights and Survey Questions**: Conduct additional Delphi rounds and review survey weighting mechanisms for greater precision.
- **Broaden Contextual Applicability**: Adapt the framework to other digital services, leveraging insights from diverse sectors.

This iterative methodology ensures the framework's continuous evolution, enabling it to address emerging ethical challenges in digital services with precision and adaptability in different sectors. The prototype in Figure 9 is under development. It is expected to complete its implementation considering ethical concerns related to digital platforms and GenAI. Afterward, user-centric evaluation with employees has to be conducted to evaluate ethical problems and interface usability. In conclusion, our framework provides a scalable and

adaptable solution to evaluate ethical concerns in digital services. Combining expert discussions, surveys, and automated analysis enables service providers to identify and mitigate risks proactively. Over the coming months, we aim to transition from research to full implementation, targeting a service release within the next few years. Furthermore, we will compare and evaluate different LLM frameworks and versions in our compass for different use cases. Moreover, it is planed to implement linguistic summaries based on Fuzzy Logic using static recomendations and smart radar values to enhance explainability and interpretability by converting numerical patterns into natural language insights (Kacprzyk and Yager, 2001; Wrede et al., 2022). The implementation will focus on the development of an adaptive summarization model that generates context-aware linguistic interpretations.

Acknowledgment

References

Andreasyan, N., Buson, D., Rüst, J., Portmann, E., & Terán, L. (2024). Theoretical framework of digital ethics concerns for public services: Electronic voting use case. *2024 Tenth International Conference on eDemocracy & eGovernment (ICEDEG)*, 1–9.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': The us, eu, and uk approach. *Science and engineering ethics*, *24*. DOI: https : / / doi . org/ 10. 1007/ s11948-017-9901-7.

Dalkey, N., & Helmer, O. (1963). An experimental application of the delphi method to the use of experts. *Management Science*, *9*(3), 458–467.

Döringer, S. (2021). 'the problem-centred expert interview'. combining qualitative interviewing approaches for investigating implicit expert knowledge. *International journal of social research methodology*, *24*(3), 265–278.

Druliac, H., Bardsley, M., Riches, C., Dunn, C., Harrison, L., Roy, B., & Hao, F. (2024). On the feasibility of e2e verifiable online voting–a case study from durga puja trial. *Journal of Information Security and Applications*, *81*, 103719.

Dubai, S. (2023). Ai system ethics self-assessment tool [Last accessed: 2025-01-17]. https : / / www . digitaldubai.ae/self-assessment

Enguehard, C. (2014). Ethics and electronic voting. *ETHICOMP 2014-Liberty and Security in an Age of ICTs*.

Fowler Jr, F. J., & Cosenza, C. (2009). Design and evaluation of survey questions. *The SAGE handbook of applied social research methods*, *2*, 375–412.

Gritzalis, D. A. (2002). Principles and requirements for a secure e-voting system. *Computers & Security*, *21*(6), 539–556.

Hakimi, L., Eynon, R., & Murphy, V. A. (2021). The ethics of using digital trace data in education: A thematic review of the research landscape. *Review of Educational Research*, *91*(5), 671–717.

Hardebolle, C., Macko, V., Ramachandran, V., Holzer, A., & Jermann, P. (2023). Digital ethics canvas: A guide for ethical risk assessment and mitigation in the digital domain. *European Society for Engineering Education (SEFI)*. DOI: https://doi.org/10.21427/9wa5-zy95.

Hart, R. (2024). *Clearview ai fined over 30 million usd for facial recognition database* [Accessed: 2025-04-10]. Forbes. https://www.forbes.com/sites/roberthart/2024/09/03/clearview-ai-controversial-facial-recognition-firm-fined-33-million-for-illegal-database

Huang, P.-h., Kim, K.-h., & Schermer, M. (2022). Ethical issues of digital twins for personalized health care service: Preliminary mapping study. *Journal of Medical Internet Research*, *24*(1), e33081.

Joisten, K., Thiemer, N., Renner, T., Janssen, A., & Scheffler, A. (2022). Focusing on the ethical challenges of data breaches and applications. *2022 IEEE International Conference on Assured Autonomy (ICAA)*, 74–82. DOI: https://doi.org/10.1109/ICAA52185.2022.00018.

Kacprzyk, J., & Yager, R. R. (2001). Linguistic summaries of data using fuzzy logic. *International Journal of General System*, *30*(2), 133–154.

Kukreja, S., Kumar, T., Purohit, A., Dasgupta, A., & Guha, D. (2024). A literature survey on open source large language models. *Proceedings of the 2024 7th International Conference on Computers in Management and Business*, 133–143. DOI: https://doi.org/10.1145/3647782.3647803.

Lauer, T. W. (2004). The risk of e-voting. *Electronic Journal of E-government*, *2*(3), pp169–178.

Linstone, H. A., Turoff, M., et al. (1975). *The delphi method*. Addison-Wesley Reading, MA.

Lukianets, N., Nekrutenko, V., & Pavaloiu, A. (2021). Openethicsai/canvas: The open ethics canvas v1.0.1 [Accessed: 2025-01-16]. DOI: https://doi.org/10.5281/zenodo.5211845.

Manchanda, J., Boettcher, L., Westphalen, M., & Jasser, J. (2024). The open source advantage in large language models (llms). *arXiv preprint arXiv:2412.12004*. https://arxiv.org/abs/2412.12004

MindsDB. (2024). Which llm to choose: 12 key aspects to consider when building ai solutions [Accessed: 2025-01-23]. https://mindsdb.com/blog/which-llm-to-choose-12-key-aspects-to-consider-when-building-ai-solutions

Nabbosa, V., & Kaar, C. (2020). Societal and ethical issues of digitalization. *Proceedings of the 2020 international conference on Big Data in Management*, 118–124.

Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to decide its appropriateness. *World journal of methodology*, *11*(4), 116.

Olszewska, J. I., Systems, Committee, S. E. S., et al. (2022). Ieee/iso/iec international standard–systems and software engineering–life cycle management–part 7000: Standard model process for addressing ethical concerns during system design. *ISO/ IEC/IEEE 24748-7000 First edition 2022-11*, 1–86. DOI: https://doi.org/10.1109/IEEESTD.2022.9967807.

on AI, E. C. H.-L. E. G. (2020). The assessment list for trustworthy artificial intelligence (altai) [Last accessed: 2025-01-17]. https : / / digital - strategy . ec . europa . eu / en / library / assessment - list - trustworthy - artificial-intelligence-altai-self-assessment

Ortega, E., Tran, M., & Bandeen, G. (2023). Ai digital tool product lifecycle governance framework through ethics and compliance by design†. *2023 IEEE Conference on Artificial Intelligence (CAI)*, 353–356. DOI: https://doi.org/10.1109/CAI54212.2023.00155.

Osterwalder, A., & Pigneur, Y. (2010). Business model generation: A handbook for visionaries, game changers, and challengers.

Pakhnenko, O., Kuan, Z., et al. (2023). Ethics of digital innovation in public administration. *Business Ethics and Leadership*, *7*(1), 113–121.

Poszler, F., Portmann, E., & Lütge, C. (2024). Formalizing ethical principles within ai systems: Experts' opinions on why (not) and how to do it. *AI and Ethics*, 1–29.

Robinson, D. G., & Halderman, J. A. (2011). Ethical issues in e-voting security analysis. *International Conference on Financial Cryptography and Data Security*, 119–130.

Schoentgen, A., & Wilkinson, L. (2021). Ethical issues in digital technologies [conference session]. *23rd Biennial Conference of the International Telecommunications Society (ITS), Gothenburg, Sweden. https://www. econstor. eu/bitstream/10419/238052/1/Schoentgen-Wilkinson. pdf.*

Teran, L., Pincay, J., Wallimann-Helmer, I., & Portmann, E. (2021). A literature review on digital ethics from a humanistic and sustainable perspective. *14th International Conference on Theory and Practice of Electronic Governance*, 57–64. DOI: https://doi.org/10.1145/3494193.3494295.

The Guardian. (2010). *Google admits collecting wi-fi data through street view cars* [Accessed: 2025-04-10]. The Guardian. https://www.theguardian.com/technology/2010/may/15/google-admits-storing-private-data

The Guardian. (2018). *Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach* [Accessed: 2025-04-10]. The Guardian. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

The Guardian. (2022). *Clearview ai fined in uk over facial recognition data collection* [Accessed: 2025-04-10]. The Guardian. https://www.theguardian.com/technology/2022/may/25/techscape-clearview-ai-facial-recognition-fine

Trechsel, A. H., Kucherenko, V. V., & da Silva, F. F. (2016). *Potential and challenges of e-voting in the european union* (tech. rep. No. 2016/11) (Last accessed: 2025-01-17). European Union Democracy Observatory (EUDO). https://hdl.handle.net/1814/44926

UbiOps. (2024). Which llm to choose for your use case [Accessed: 2025-01-23]. https://ubiops.com/which-llm-to-choose-for-your-use-case

Wallimann-Helmer, I., Terán, L., Portmann, E., Schübel, H., & Pincay, J. (2021). An integrated framework for ethical and sustainable digitalization [ISSN: 2573-1998]. *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, 156–162. DOI: https : / / doi . org / 10. 1109/ ICEDEG52154 . 2021.9530972.

Warner, L. M. (2023). Ethics in public service. In *Global encyclopedia of public administration, public policy, and governance* (pp. 4394–4398). Springer.

Wired. (2012). *Google engineer told colleagues street view wi-fi data collection was not a mistake* [Accessed: 2025-04-10]. Wired. https://www.wired.com/2012/05/google-wifi-fcc-investigation/

Wrede, C., Winands, M., & Wilbik, A. Linguistic summaries as explanation mechanism for classification problems [The 34th Benelux Conference on Artificial Intelligence and the 31th Belgian Dutch Conference on Machine Learning, BNAIC 2022 ; Conference date: 07-11-2022 Through 09-11-2022]. English. In: In *The 34th benelux conference on artificial intelligence and the 31th belgian dutch conference on machine learning*. The 34th Benelux Conference on Artificial Intelligence and the 31th Belgian Dutch Conference on Machine Learning, BNAIC 2022 ; Conference date: 07-11-2022 Through 09-11-2022. 2022, November.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338–353.

Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences*, *178*(13), 2751–2779. DOI: https://doi.org/https://doi.org/10.1016/j.ins.2008.02.012.