When Security Meets Ethics: Balancing Control, Privacy, and Trust in Digital Systems[1]

Laercio Cruvinel[1,2][0000-0002-7672-3243], Valéria Pequeno[2,3][0000-0002-6424-0252]

[1] Universidade Autónoma de Lisboa Luís de Camões, 1169-023 Lisboa, Portugal
[2] Autónoma TechLab, 1169-023 Lisboa, Portugal
[3] Escola Superior Náutica Infante Dom Henrique, 2770-058 Paço de Arcos
lcruvinel@autonoma.pt✉, vpequeno@enautica.pt

Abstract. As digital systems become increasingly embedded in education, business, government, and everyday life, the boundary between technological security and ethical responsibility has grown more complex. This paper examines how the pursuit of digital security—through practices such as data collection, monitoring, automation, and AI-driven decision-making—often intersects with, and occasionally undermines, the ethical principles of privacy, autonomy, fairness, and transparency. Using ethical frameworks including deontology, consequentialism, and virtue ethics, the study analyzes key challenges such as algorithmic bias, digital surveillance, cybercrime, and digital burnout. It argues that achieving a sustainable balance between control and trust requires not only robust technical protections but also clear institutional policies and user-centered ethical guidelines. By synthesizing current literature and offering a structured set of best practices, the paper proposes an integrated approach to the governance of computing technologies—one that ensures security while safeguarding human dignity, agency, and well-being.

Keywords: digital ethics; cybersecurity; privacy; algorithmic bias; autonomy; digital surveillance; responsible AI; trust in digital systems.

1 Introduction

1.1 Purpose

The rapid expansion of digital technologies has transformed the way individuals learn, communicate, work, and interact with institutions. From data-driven decision-making in public administration to AI-enhanced platforms in education and business, digital systems now operate at the core of modern societal infrastructures. Yet with this transformation comes a growing tension between the need for security (protecting data,

[1] **This work was funded by national funds through FCT - Fundação para a Ciência e a Tecnologia - as part the Projeto Ci2.UAL - Refª UID/05567/2025**; https://doi.org/10.54499/UID/05567/2025

ensuring system reliability, and preventing cybercrime) and the ethical obligation to respect privacy, autonomy, and fairness.

Security mechanisms such as continuous monitoring, data aggregation, and automated profiling enhance protection but can simultaneously erode user trust, introduce opaque decision processes, and normalize forms of digital surveillance. Ethical concerns are therefore not peripheral but central to evaluating what "secure" digital systems ought to mean in practice. Balancing these competing demands requires more than technical safeguards; it demands principled reflection on how digital control is exercised, who benefits from it, and what values may be compromised along the way.

This paper addresses these challenges by examining how security practices intersect with core ethical principles, including privacy, autonomy, transparency, and justice. Drawing on deontological, consequentialist, and virtue-based ethical frameworks, it explores the value trade-offs embedded in contemporary digital environments, from AI-driven decision systems to everyday cybersecurity protocols. The discussion highlights the need for integrated strategies that align technical robustness with ethical responsibility, thereby ensuring that digital security strengthens, rather than undermines, public trust.

1.2 Methodology

This study follows a *conceptual literature review* methodology aimed at synthesizing current research on ethical and security challenges in computing and Internet technologies. The review adopts a qualitative, interpretative approach, drawing primarily from peer-reviewed publications, institutional reports, and policy documents published between 2015 and 2025. The search was conducted across major academic databases, including Scopus, Web of Science, and IEEE Xplore, using combinations of keywords such as "digital ethics", "cybersecurity", "AI governance", and "data privacy". Sources were selected based on their conceptual relevance and contribution to discussions on privacy, autonomy, fairness, transparency, and accountability in technological systems. Rather than providing a quantitative meta-analysis, this paper integrates and contrasts diverse scholarly perspectives to construct a comprehensive ethical framework for understanding the evolving relationship between computing technologies, security, and human values.

1.3 Structure

The remainder of this paper is structured as follows. Section 2 provides a review of the background on the theoretical foundations of the study, including ethics in computing and the related frameworks. Section 3 discusses the ethical and security challenges of computing technologies, including data privacy, algorithmic bias, autonomy, and cybersecurity. Section 4 suggests best practices for ethical technology use, including ethical implementation, AI literacy, human-AI collaboration, and responsible innovation. Finally, Section 5 offers conclusions and recaps the main insights of the work.

2 Background

2.1 Ethics in Computing Technologies

The ethical implications of computing technologies have become increasingly significant as digital systems permeate nearly every aspect of daily life. Their impact extends well beyond technical performance, intersecting with core issues such as human rights, individual autonomy, and social equity [1]. This section introduces four central ethical principles—privacy, autonomy, fairness, and transparency—and explains their importance in the design and use of contemporary computing technologies:

- **Privacy**: Privacy in computing technologies has emerged as a paramount ethical issue, particularly in how user data is collected and managed across various platforms, including social media and learning management systems (LMS). The principle of informed consent is critical in these interactions; violations occur when data is gathered without explicit user agreement and when there is a lack of transparency regarding data handling practices. Ienca et al. illustrate that the ongoing ethical conversation predominantly focuses on privacy and data protection, urging for a broader understanding of ethical data practices beyond mere concerns for individual privacy [2]. This resonates with contemporary discussions that highlight healthcare settings, noting that many individuals do not examine privacy policies when using applications, consequently exposing themselves to privacy violations [3].
- **Autonomy**: Autonomy emerges as a foundational ethical principle in computing, asserting individuals' rights to make informed decisions free from external influences. In the context of computing technologies, autonomy can become compromised, particularly when automated systems, such as social media algorithms or adaptive learning platforms, make decisions without sufficient user input or understanding. Putica et al. highlight the importance of adhering to ethical frameworks that prioritize autonomy and informed consent in AI-assisted decision-making, especially in sensitive sectors like mental health [4]. This emphasis on ethical guidelines underscores the necessity of ensuring that automated systems do not override user autonomy through opaque decision-making processes.
- **Fairness**: Fairness in computing is an essential ethical principle that mandates equitable treatment of individuals and strives to eliminate discrimination in decision-making processes facilitated by algorithms. The challenge of ensuring fairness within algorithm design is increasingly critical as the reliance on artificial intelligence (AI) systems grows across various sectors. Bias can be introduced during the training phase of AI models, often due to unrepresentative datasets, leading to discriminatory outcomes, as evidenced by multiple studies Fazil et al. [5][6]. For instance, biased facial recognition technologies have been shown to disproportionately misidentify individuals from minority groups, raising significant ethical concerns about their deployment in sensitive areas like law enforcement [6]. Similarly, biased automated grading systems can perpetuate existing disparities in educational assessments, which may impact the academic outcomes of underrepresented groups [7].
- **Transparency**: Transparency in algorithmic systems denotes the extent to which a system’s operation, inputs, decision logic, and limitations can be observed,

interpreted, and contested by appropriate audiences, from end users to auditors [8]. It matters because opacity increases the risk of unfair or unaccountable outcomes and prevents meaningful oversight in high stakes domains such as public services and healthcare [8][9][10]. Transparency addresses both making internal mechanics interpretable (explainability) and enabling external inspection and testing (auditability) rather than a single unified output [11]. Transparent systems support ethical aims including informed consent, remediation of bias, and assignment of responsibility across development and deployment phases [8][12].

Privacy, autonomy, fairness, and transparency form an interdependent set of ethical principles in computing. Greater transparency can support autonomy by helping users make informed choices, and achieving fairness often depends on algorithms that are both understandable and respectful of privacy. Yet these principles may also come into tension—such as when increasing transparency risks revealing sensitive information. Navigating these trade-offs requires a nuanced, context-aware approach, making ethical reflection essential in the design and governance of digital technologies.

2.2 Theoretical Framework

Deontology, consequentialism and virtue ethics are ethical perspectivesthat provide complementary lenses for analyzing digital ethics.

- **Deontology**: Immanuel Kant's (1724–1804) ethical theory, deontology, is grounded in the idea that morality depends on duties defined by what is right or wrong. In this view, right actions are those we are obligated to carry out, whereas wrong actions are strictly prohibited, regardless of their consequences. Unlike outcome-oriented ethical theories, deontology holds that moral duties take priority over considerations of benefit or harm. Kant maintains that these duties stem from a single foundational principle—the categorical imperative [13]. In the context of computing technologies, deontological ethics supports approaches that prioritize user rights, informed consent, fairness, and the protection of autonomy and privacy. Several studies draw on deontological or rights-based frameworks to examine ethical issues in computing, particularly in relation to user protections and equitable treatment.
- **Consequentialism**: Consequentialism is an ethical theory that evaluates the moral worth of an action solely by its outcomes. Actions are deemed right when they produce the most favorable consequences and wrong when they result in harm. This perspective rests on the intuition that the future is what truly matters—since the past cannot be changed, ethical decision-making should prioritize generating the best possible results moving forward [14].
- **Virtue ethics**: virtue ethics is a normative framework that prioritizes the cultivation of moral character and virtues—such as kindness, honesty, and courage—over adherence to rules (as in deontology) or the pursuit of outcomes (as in consequentialism). Ethical conduct is understood through the quality of an individual's character. For instance, a virtue ethicist would interpret helping another person as an expression of compassion, demonstrating virtuous character [15].

Combining consequentialist assessments with virtue-based perspectives enables institutions and developers to navigate ethical choices around security measures by

weighing practical advantages against moral risks. The evidence suggests that integrating consequentialist reasoning with virtue-ethics principles can support a balanced approach, allowing organizations and developers to harness the benefits of computing technologies while remaining attentive to their ethical implications.

2.3 Applications

- Natural language processing, cybersecurity, systems development are examples of computing areas where deontological principles have been applied [16][17][18].
- Studies often adopt rights-based approaches that align with deontological ethics, emphasizing informed consent as a cornerstone for safeguarding personal autonomy and privacy [19][20][21].
- Framework proposals for algorithmic fairness [22] and digital human rights [21] illustrate how ethical design principles can mitigate inequities in information access and enhance security.
- Research grounded in consequentialist reasoning—frequently articulated through utilitarian ethics—advocates for risk–benefit evaluations and proportionality measures to weigh security advantages against ethical concerns, including privacy violations, surveillance practices, and profiling [23][19].
- Scholarship informed by virtue ethics underscores that fostering character traits such as courage, temperance, and open-mindedness can serve as a compass for ethical decision-making in cybersecurity. Nakato et al. [24], for instance, demonstrate how virtues shape cybersecurity practices in the banking sector, while Barford and Pieters together with Consoli argue that embedding virtues into design processes and providing ethical training are crucial for responsible technology development [25][26].

Together, these ethical perspectives—deontological, consequentialist, and virtue-based—provide complementary lenses for analyzing digital ethics. The following sections apply these frameworks to real-world issues such as data privacy, algorithmic bias, cybersecurity, and digital surveillance. By examining how each ethical theory interprets technological practices, the analysis reveals both convergences and tensions between moral duty, outcome-based reasoning, and character-centered approaches in the governance of computing systems.

3 Ethical Dilemmas in Computing Technologies

3.1 Data Privacy and Surveillance

Privacy remains one of the most critical ethical challenges in computing technologies [27]. As digital platforms become embedded in everyday life, the ongoing collection, storage, and processing of personal information raise concerns about autonomy, consent, and user control. Ethical responses to these challenges can be illuminated through three complementary perspectives:

- **Deontological Ethics (Rights-Based Approach):** From a deontological standpoint, protecting privacy is a moral duty grounded in respect for individual rights. Ethical practice requires honoring user autonomy through informed consent, transparent communication, and strict adherence to legal frameworks such as the GDPR. Violations of privacy are wrong in themselves, regardless of potential benefits.
- **Consequentialist Ethics (Utilitarian Approach):** Risk–benefit analysis and proportionality assessments are used to balance the security advantages of data collection against harms such as surveillance, profiling, or loss of trust. Ethical decisions aim to maximize overall well-being while minimizing negative consequences for individuals and society.
- **Virtue Ethics (Character-Based Approach):** Ethical data governance is not only about rules or outcomes but about fostering a culture of integrity. For example, designing systems with transparency and restraint reflects virtues that guide responsible technology development.

Safeguarding privacy in computing is not merely a technical challenge; it has become an ethical imperative. It requires a **deontological commitment to rights**, a **consequentialist evaluation of risks and benefits**, and a **virtue-based cultivation of integrity and responsibility**. Together, these perspectives provide a holistic framework for continuously adapting privacy practices to emerging threats and evolving user expectations [28].

3.2 Algorithmic Bias and Fairness

Artificial intelligence (AI) and machine learning have become integral across domains such as healthcare, finance, and education, delivering efficiency gains and personalized recommendations. Yet these systems remain vulnerable to algorithmic bias, which can generate inequitable outcomes, particularly for marginalized communities. Bias often arises from three primary sources: skewed training datasets, deficiencies in model design, and inappropriate deployment practices [29]. Confronting these ethical challenges demands a sophisticated grasp of fairness—one that can be examined through diverse ethical frameworks:

- **Deontological Ethics (Duty-Based Approach):** From a deontological perspective, fairness is a matter of respecting individual rights and ensuring equal treatment. Biased algorithms are unethical because they violate principles of justice and autonomy, regardless of efficiency gains. Ethical practice requires transparency, accountability, and strict adherence to non-discrimination norms.
- **Consequentialist Ethics (Outcome-Based Approach):** Consequentialism evaluates fairness by outcomes. Algorithms are considered ethical if they maximize overall benefits—such as improved healthcare or financial inclusion—while minimizing harms like exclusion, misrepresentation, or profiling. Risk–benefit analysis and proportionality assessments help balance innovation against potential social costs.
- **Virtue Ethics (Character-Based Approach):** Virtue ethics emphasizes the moral character of developers and institutions. Fairness is achieved not only through rules

or outcomes but through cultivating virtues such as honesty, empathy, and responsibility. Designing systems with humility and openness reflects a commitment to ethical character, guiding responsible technology development.

In summary, addressing algorithmic bias benefits from a **deontological commitment to rights**, a **consequentialist evaluation of outcomes**, and a **virtue-based cultivation of integrity**. Together, these perspectives provide a holistic framework for embedding fairness into AI and machine learning systems.

3.3 Autonomy, Digital Ethics, and Work-Life Balance

AI-driven recommendations and automated decision systems increasingly shape personal and professional choices. While these tools can enhance efficiency and convenience, they also risk undermining user autonomy—particularly when they operate as "black boxes" with opaque decision-making processes [27]. A lack of transparency restricts individuals' ability to make informed choices, thereby weakening moral agency and raising ethical concerns about control and accountability.

- **Deontological Ethics (Rights-Based Approach):** From a deontological perspective, autonomy is a fundamental right. Users deserve to understand how AI systems influence their decisions and must retain meaningful control over their digital environments. Opaque algorithms are ethically problematic because they deny individuals the ability to exercise informed consent and self-determination.
- **Consequentialist Ethics (Utilitarian Approach):** Consequentialism evaluates automation in terms of outcomes. While efficiency gains and productivity improvements may justify AI-driven decision-making, these benefits must be weighed against potential harms, such as psychological stress, diminished user control, and erosion of work-life balance. Ethical deployment requires proportionality—ensuring that the social costs of reduced autonomy do not outweigh the practical advantages.
- **Virtue Ethics (Character-Based Approach):** Virtue ethics emphasizes the cultivation of responsible and empathetic practices among developers, institutions, and users. Designing transparent systems, fostering digital literacy, and encouraging mindful use of AI reflect virtues such as honesty, prudence, and respect. These character-driven commitments help ensure that technology supports, rather than undermines, human flourishing and balanced lifestyles.

Safeguarding autonomy in the digital age is not to be regarded as merely a technical challenge – it is an ethical imperative. It requires a **deontological respect for rights**, a **consequentialist balancing of benefits and harms**, and a **virtue-based commitment to integrity and responsibility**. Together, these perspectives provide a holistic framework for aligning AI with digital ethics and sustaining healthy work-life balance.

3.4 Cybersecurity and Digital Exhaustion

Cybersecurity is indispensable for safeguarding digital systems, networks, and data against unauthorized access and damage. Yet excessive or overly intrusive security measures can unintentionally undermine user well-being, eroding autonomy and

contributing to stress and fatigue [30]. Balancing protection with mental health and trust has therefore become a critical ethical challenge.

- **Deontological Ethics (Rights-Based Approach):** Organizations, particularly those managing sensitive data in education, healthcare, or finance, carry a moral obligation to protect user information. Measures such as encryption, secure authentication, access control, and regular vulnerability assessments are ethical duties. Failure to uphold these responsibilities compromises user rights and violates the principle of respect for autonomy [31].
- **Consequentialist Ethics (Utilitarian Approach):** While advanced monitoring and authentication systems may enhance protection, they can also generate negative consequences such as digital exhaustion, anxiety, and mistrust. Ethical decision-making requires proportionality: security benefits must be weighed against the psychological and social costs of reduced user control and well-being [32].
- **Virtue Ethics (Character-Based Approach):** Developers and institutions should embody virtues such as prudence, transparency, and respect, ensuring that protective measures do not devolve into excessive surveillance. A human-centric approach reflects integrity and care, fostering environments where security supports rather than undermines user flourishing.

This analysis show that the ethical tension between cybersecurity and digital exhaustion calls for a balanced framework: a **deontological commitment to rights**, a **consequentialist evaluation of risks and benefits**, and a **virtue-based cultivation of integrity and empathy**. Together, these perspectives guide organizations toward securing digital systems while respecting autonomy, trust, and mental health.

4 Strategies and Ethical Best Practices

Promoting the ethical and secure use of computing technologies demands a comprehensive approach that integrates technical safeguards, regulatory frameworks, and educational initiatives. Although best practices can foster responsible technology adoption, their implementation is often constrained by technical limitations, resource availability, and resistance from stakeholders.

4.1 Responsible AI Design and Governance

Developing AI responsibly requires adherence to best practices that embed ethical principles at the core of technological innovation. The following best practices are essential for ethical AI development and adherence to ethical principles that prioritize human rights, fairness, and accountability:

- **Transparency**: AI systems should communicate in clear and accessible ways how they operate, the basis for their decisions, and the data they rely on. Such openness fosters accountability and strengthens user trust. For instance, an AI-driven recruitment platform ought to explain the criteria it applies when evaluating candidate qualifications.

- **Accountability**: Developers should establish mechanisms to identify, mitigate, and prevent biases in AI models, employing practices such as routine bias audits and explainable AI (XAI) methods [33]. These safeguards ensure that algorithmic decisions—for example, in financial credit scoring—remain subject to review and correction when bias is detected.
- **Inclusivity**: AI systems ought to be developed using diverse and representative datasets that capture the breadth of demographic groups, thereby reducing the risk of reinforcing social inequalities. For example, facial recognition technologies should be evaluated across varied populations to safeguard against racial or gender bias.
- **Human-Centered Design**: AI systems should be developed with a focus on user well-being, incorporating features that enhance user control, provide clear and understandable explanations for recommendations, and reduce the risk of digital fatigue. For instance, educational AI platforms ought to enable students to personalize their learning trajectories instead of requiring strict adherence to algorithmic outputs.

Challenges for following and implementing these best practices include **Complexity and Cost**, **Resistance from Developers**, and **Bias in Training Data** [34][35][36].

4.2 Enhancing Cybersecurity Practices

To safeguard user data, preserve system integrity, and prevent unauthorized access, both organizations and individuals should adopt the following best practices:

- **Encryption and Strong Authentication**: Sensitive information must be secured through robust encryption techniques, while multi factor authentication (MFA) should be implemented to guarantee safe system access.
- **Routine Security Assessments**: Institutions should carry out regular vulnerability scans and penetration tests to detect and remediate potential security flaws.
- **Cybersecurity Awareness Programs**: Training initiatives should equip users with knowledge about prevalent threats, including phishing schemes, social engineering tactics, and malware infections.

Challenges for implementing these measures include **Evolving Threat Landscape**, **Resource Constraints**, and **Difficult User Compliance** [37][38].

4.3 Regulatory and Policy Frameworks

Governments and oversight bodies hold a central responsibility in promoting the ethical and secure use of computing technologies. Effective regulation should encompass several key areas:

- **Data Protection Legislation:** Authorities must enforce adherence to international data protection standards, such as the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) in the United States.

- **AI Ethics Standards:** Regulators should define clear ethical requirements for the design and deployment of AI systems, including provisions for transparency, accountability, and the mitigation of bias.
- **Safeguarding Consumer Rights:** Organizations must be obligated to uphold transparent data collection practices, secure informed user consent, and provide mechanisms that allow individuals to contest or erase their personal information.

Implementation, however, faces persistent challenges, including the **financial burden of compliance**, **conflicts across jurisdictions**, and **regulatory lag** caused by the rapid pace of technological innovation [39][40].

4.4 Digital Literacy and User Responsibility

Digital literacy is a cornerstone for enabling individuals to effectively address the ethical and security challenges of contemporary digital environments. Educational initiatives should emphasize the following dimensions:

- **Privacy Awareness:** Users need to be equipped with strategies to protect their personal information, including managing privacy settings, identifying phishing attempts, and controlling online permissions.
- **Critical Evaluation Skills:** Learners should cultivate the ability to assess digital content critically, detect misinformation, and question the reliability of algorithmic recommendations.
- **Ethical Engagement with Technology:** Training programs should foster responsible digital behavior, promoting the ethical use of AI, respectful online communication, and attentiveness to digital well-being.

Despite their importance, these initiatives face notable barriers, including issues of **accessibility**, the **rapid pace of technological change**, and **limited user engagement** [41][42][43].

4.5 Building a Comprehensive Ethical Framework

For best practices to be effectively realized, organizations must embrace a holistic ethical framework that integrates technical safeguards, regulatory oversight, and educational initiatives. Such a framework should encompass:

- **Ongoing Ethical Reviews:** Institutions should conduct systematic evaluations of their AI and cybersecurity practices, identifying potential ethical risks and opportunities for improvement.
- **Inclusive Governance:** Policy development should actively engage diverse stakeholders—including students, educators, developers, policymakers, and end-users—to ensure broad representation in ethical decision-making.
- **Dynamic Policy Adaptation:** Ethical guidelines must be continuously revised to respond to emerging challenges and rapid technological advancements.
- **Empowerment and Accessibility:** Digital literacy initiatives should emphasize inclusivity, ensuring that all users are equipped to benefit from ethical and responsible technology use.

By embedding these practices within a comprehensive framework, organizations can foster secure, transparent, and ethically accountable digital environments that uphold user rights while encouraging innovation.

5 Conclusions

The rapid evolution of computing technologies and the Internet has reshaped how we learn, communicate, conduct business, and govern. Yet these advancements also introduce significant ethical and security challenges, including data privacy concerns, algorithmic bias, cybersecurity vulnerabilities, digital exhaustion, and the complex implications of AI-driven decision-making. This paper has examined these issues through a critical ethical lens and outlined best practices to support more responsible, transparent, and user-centered technology use.

The secure and ethical use of digital technologies is not merely a technical undertaking but a foundational responsibility that demands a careful balance between control, privacy, and trust. As digital systems become increasingly embedded in everyday life, principles such as transparency, fairness, and accountability must guide their design and governance. Users should be empowered through clear explanations of algorithmic processes, meaningful consent mechanisms, and safeguards that protect their autonomy and well-being.

At the same time, the pursuit of stronger security cannot come at the expense of fundamental rights. While robust cybersecurity measures are essential for protecting data and ensuring system reliability, they must be implemented in ways that avoid creeping surveillance, unjust profiling, or disproportionate control over users. Ethical AI development requires deliberate efforts to minimize bias, enhance fairness, and preserve the agency of individuals affected by automated decisions.

Ultimately, responsible technology use relies on cultivating a culture of ethical reflexivity—where institutions continuously reassess the implications of their digital practices, remain attentive to emerging risks, and adapt in ways that reinforce public trust. Such reflexivity strengthens not only compliance but also the broader resilience and legitimacy of digital systems within a democratic society.

While this paper proposes a comprehensive framework for addressing ethical and security challenges in computing technologies, it also recognizes certain methodological and epistemological limitations. Digital systems, despite their analytical power, cannot fully capture the moral, cultural, and contextual dimensions that shape ethical decision-making. Overreliance on automated or algorithmic assessments risks oversimplifying complex moral judgments, reducing ethics to compliance or technical control. Moreover, digitalization introduces its own epistemic constraints — privileging quantifiable data over qualitative insights and potentially marginalizing perspectives that resist computational formalization. These limitations underline the importance of maintaining a *critical human perspective* in the governance of digital systems, ensuring that ethical reflection remains grounded in dialogue, pluralism, and contextual understanding. Future research should continue exploring hybrid models that combine normative

reasoning, participatory governance, and technical safeguards, fostering an ethical digital ecosystem that balances innovation with human-centered responsibility.

References

[1] M. Ryan and B. C. Stahl, "Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications," *Journal of Information, Communication and Ethics in Society*, vol. 19, no. 1, pp. 61–86, Jun. 2020, doi: 10.1108/JICES-12-2019-0138.

[2] M. Ienca, A. Ferretti, S. Hurst, M. Puhan, C. Lovis, and E. Vayena, "Considerations for ethics review of big data health research: A scoping review," *PLoS One*, vol. 13, no. 10, Oct. 2018, doi: 10.1371/journal.pone.0204937.

[3] A. Prasuna and A. Rachh, "A STUDY ON CHALLENGES OF DATA SECURITY AND DATA PRIVACY IN THE HEALTHCARE SECTOR: SWOT ANALYSIS," in *Asia Pacific Journal of Health Management*, Mumbai: Australasian College of Health Service Management, Apr. 2023. doi: 10.24083/apjhm.v18i1.1675.

[4] A. Putica, R. Khanna, W. Bosl, S. Saraf, and J. Edgcomb, "Ethical decision-making for AI in mental health: The Integrated Ethical Approach for Computational Psychiatry (IEACP) framework," Jul. 24, 2025, *Cambridge University Press*. doi: 10.1017/S0033291725101311.

[5] Abdul Wajid Fazil, Musawer Hakimi, and Amir Kror Shahidzay, "A COMPREHENSIVE REVIEW OF BIAS IN AI ALGORITHMS," *Nusantara Hasana Journal*, vol. 3, no. 8, pp. 1–11, Jan. 2024, doi: 10.59003/nhj.v3i8.1052.

[6] Oyekunle Claudius Oyeniran, Adebunmi Okechukwu Adewusi, Adams Gbolahan Adeleke, Lucy Anthony Akwawa, and Chidimma Francisca Azubuko, "Ethical AI: Addressing bias in machine learning models and software applications," *Computer Science & IT Research Journal*, vol. 3, no. 3, pp. 115–126, Dec. 2022, doi: 10.51594/csitrj.v3i3.1559.

[7] E. Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," *Sci*, vol. 6, no. 1, p. 3, Dec. 2023, doi: 10.3390/sci6010003.

[8] P. Pappachan, M. Moslehpour, R. Bansal, and M. Rahaman, "Transparency and Accountability," 2024, pp. 178–211. doi: 10.4018/979-8-3693-3860-5.ch006.

[9] M. Veale, M. Van Kleek, and R. Binns, "Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, Apr. 2018, pp. 1–14. doi: 10.1145/3173574.3174014.

[10] T. Thakur, S. S. Khurshid, and N. Sharma, "Enhancing Transparency in Healthcare Decision-Making: An Explainable AI Approach," in *2025 12th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP)*, IEEE, Aug. 2025, pp. 1–6. doi: 10.1109/ICETETSIP64213.2025.11156742.

[11] A. Springer and S. Whittaker, “Progressive disclosure,” in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, New York, NY, USA: ACM, Mar. 2019, pp. 107–120. doi: 10.1145/3301275.3302322.

[12] A. Farooq and K. Iqbal, “Towards Transparent Ethical AI: A Roadmap for Trustworthy Robotic Systems,” Aug. 2025, doi: 10.1145/3747393.374740.

[13] R. N. Johnson and A. Cureton, “Kant’s Moral Philosophy,” in *The Stanford Encyclopedia of Philosophy*, Winter 2025., E. Zalta and U. Nodelman, Eds., 2025.

[14] W. Sinnott-Armstrong, “Consequentialism,” in *The Stanford Encyclopedia of Philosophy*, Winter 2023., E. N. Zalta and U. Nodelman, Eds., 2023.

[15] R. Hursthouse and G. Pettigrove, “Virtue Ethics,” in *The Stanford Encyclopedia of Philosophy*, Fall 2023., E. N. Zalta and U. Nodelman, Eds., 2023.

[16] S. Prabhumoye, B. Boldt, R. Salakhutdinov, and A. W. Black, “Case Study: Deontological Ethics in NLP,” in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 3784–3798. doi: 10.18653/v1/2021.naacl-main.297.

[17] C. Posey and R. Folger, “An exploratory examination of organizational insiders’ descriptive and normative perceptions of cyber-relevant rights and responsibilities,” *Comput Secur*, vol. 99, p. 102038, Dec. 2020, doi: 10.1016/j.cose.2020.102038.

[18] H. Rahanu, E. Georgiadou, K. Siakas, and M. Ross, “Health, safety, and well-being through ethical behaviours in computer and information systems development and deployment,” *Journal of Software: Evolution and Process*, vol. 32, no. 3, Mar. 2020, doi: 10.1002/smr.2225.

[19] M. Loi and M. Christen, “Ethical Frameworks for Cybersecurity,” 2020, pp. 73–95. doi: 10.1007/978-3-030-29053-5_4.

[20] G. J. Walters, “Privacy and security,” *ACM SIGCAS Computers and Society*, vol. 31, no. 2, pp. 8–23, Jun. 2001, doi: 10.1145/503345.503347.

[21] T. Tsiakis, “The Role of Information Security and Cryptography in Digital Democracy,” in *Human Rights and Ethics*, IGI Global, 2015, pp. 1564–1580. doi: 10.4018/978-1-4666-6433-3.ch086.

[22] E. S. Udoh, X. Yuan, and A. Rorissa, “A Framework for Defining Algorithmic Fairness in the Context of Information Access,” *Proceedings of the Association for Information Science and Technology*, vol. 61, no. 1, pp. 667–672, Oct. 2024, doi: 10.1002/pra2.1077.

[23] A. D. Smith, “E-security issues and policy development in an information-sharing and networked environment,” *Aslib Proc*, vol. 56, no. 5, pp. 272–285, Oct. 2004, doi: 10.1108/00012530410560878.

[24] R. Nakato, M. G. Kituyi, and F. Kaggwa, “Establishing the Influences of Cardinal Virtues on Employees’ Cyber Security Ethical Behavior in the Banking Sector in Uganda,” *European Journal of Technology*, vol. 6, no. 1, pp. 1–13, Jan. 2022, doi: 10.47672/ejt.896.

[25] L. Barford, “Contemporary Virtue Ethics and the Engineers of Autonomous Systems,” in *2019 IEEE International Symposium on Technology and Society (ISTAS)*, IEEE, Nov. 2019, pp. 1–7. doi: 10.1109/ISTAS48451.2019.8937855.

[26] W. Pieters and L. Consoli, “Vulnerabilities and responsibilities: dealing with monsters in computer security,” *Journal of Information, Communication and Ethics in Society*, vol. 7, no. 4, pp. 243–257, Oct. 2009, doi: 10.1108/14779960911004499.

[27] M. Zostant and R. Chataut, “Privacy in computer ethics: Navigating the digital age,” *Computer Science and Information Technologies*, vol. 4, no. 2, pp. 183–190, Jul. 2023, doi: 10.11591/csit.v4i2.pp183-190.

[28] S. van Baalen, “‘Google wants to know your location’: The ethical challenges of fieldwork in the digital age,” *Res Ethics*, vol. 14, no. 4, pp. 1–17, Oct. 2018, doi: 10.1177/1747016117750312.

[29] L. A. Celi *et al.*, “Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review,” *PLOS Digital Health*, vol. 1, no. 3, p. e0000022, Mar. 2022, doi: 10.1371/journal.pdig.0000022.

[30] C. Nobles, “Stress, Burnout, and Security Fatigue in Cybersecurity: A Human Factors Problem,” *HOLISTICA – Journal of Business and Public Administration*, vol. 13, no. 1, pp. 49–72, Jul. 2022, doi: 10.2478/hjbpa-2022-0003.

[31] S. Faltermaier, K. S. Strunk, M. Obermeier, and M. Fiedler, “Managing Organizational Cyber Security - The Distinct Role of Internalized Responsibility,” in *Hawaii International Conference on System Sciences*, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:256902068

[32] M. I. F. Pinto, “Devem as instituições de ensino superior estar preocupadas com o digital burnout? Uma perspetiva sobre a satisfação no trabalho,” 2023.

[33] E. Kyrimi *et al.*, “Explainable AI: definition and attributes of a good explanation for health AI,” *AI and Ethics*, vol. 5, no. 4, pp. 3883–3896, Aug. 2025, doi: 10.1007/s43681-025-00668-x.

[34] Y. De Lucchi, “Justicia digital y discapacidad, aprovechando la oportunidad,” *Revista Española de Discapacidad*, vol. 11, no. 1, pp. 51–72, Jun. 2023, doi: 10.5569/2340-5104.11.01.03.

[35] I. Inuwa-Dutse, “FATE in AI: Towards Algorithmic Inclusivity and Accessibility,” in *Equity and Access in Algorithms, Mechanisms, and Optimization*, New York, NY, USA: ACM, Oct. 2023, pp. 1–14. doi: 10.1145/3617694.3623233.

[36] A. N. I. Purwanto, “Human-Centered AI: Designing Intelligent Systems that Empower, Not Replace,” *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 3, pp. 3212–3221, Aug. 2025, doi: 10.31004/riggs.v4i3.2450.

[37] Z. H. Joy, S. Islam, M. A. Rahaman, and Md. N. Haque, “ADVANCED CYBERSECURITY PROTOCOLS FOR SECURING DATA MANAGEMENT SYSTEMS IN INDUSTRIAL AND HEALTHCARE ENVIRONMENTS,” *Global Mainstream Journal*, vol. 3, no. 4, pp. 25–38, May 2024, doi: 10.62304/jbedpm.v3i4.147.

[38] Daniel Ajiga, Patrick Azuka Okeleke, Samuel Olaoluwa Folorunsho, and Chinedu Ezeigweneme, “Designing Cybersecurity Measures for Enterprise Software Applications to Protect Data Integrity,” *Computer Science & IT Research Journal*, vol. 5, no. 8, pp. 1920–1941, Aug. 2024, doi: 10.51594/csitrj.v5i8.1451.

[39] Excel G Chukwurah, “Agile privacy in practice: Integrating CCPA and GDPR within agile frameworks in the U.S. tech scene,” *International Journal of Scientific Research Updates*, vol. 7, no. 2, pp. 024–036, Apr. 2024, doi: 10.53430/ijsru.2024.7.2.0035.

[40] Adebola Folorunso, Olufunbi Babalola, Chineme Edger Nwatu, and Urenna Ukonne, "Compliance and Governance issues in Cloud Computing and AI: USA and Africa," *Global Journal of Engineering and Technology Advances*, vol. 21, no. 2, pp. 127–138, Nov. 2024, doi: 10.30574/gjeta.2024.21.2.0213.

[41] C. A. Reilly, "Reading risk: Preparing students to develop critical digital literacies and advocate for privacy in digital spaces," *Comput Compos*, vol. 61, p. 102652, Sep. 2021, doi: 10.1016/j.compcom.2021.102652.

[42] Bambang Mudjiyanto, Launa Launa, M. Daffa Shilbi, and Afkar Khaibar Alfaruqi, "Digital Literacy Analysis in Increasing Students' Privacy Awareness as an Effort to Prevent Doxing Crimes in the Social Media Era," *Jurnal Yudistira : Publikasi Riset Ilmu Pendidikan dan Bahasa*, vol. 3, no. 4, pp. 110–123, Aug. 2025, doi: 10.61132/yudistira.v3i4.2372.

[43] A. Jordan, A. Julianto, and M. A. Firmansyah, "Integrating Digital Literacy into Curriculum Design a Framework for 21st Century Learning," *JOURNAL OF TECHNOLOGY, EDUCATION & TEACHING (J-TECH)*, vol. 1, no. 2, pp. 79–85, Jan. 2025, doi: 10.62734/jtech.v1i2.418.