Digital Society (2025) 4:21
https://doi.org/10.1007/s44206-025-00167-w

ORIGINAL PAPER

Privacy-Friendly and Trustworthy Technology for Society

Anton Fedosov[1] · **Aurelia Tamò-Larrieux**[2] · **Christoph Lutz**[3] · **Eduard Fosch-Villaronga**[4] · **Anto Čartolovni**[5]

Received: 10 January 2025 / Accepted: 29 January 2025 / Published online: 26 March 2025

Abstract

We have witnessed an increased use of technology in every facet of our lives. These technologies come with great promises, such as enabling more independent living for older adults or people with physical disabilities, yet also fears, for instance, over privacy concerns or trust in automated systems. In this Topical Collection, we focus on Active and Assisted Living (AAL) technologies, which require trustworthiness and adherence to privacy regulations for successful adoption. The Collection contains six selected papers that address themes like privacy-by-design, trust in AI, and balancing privacy with technological innovation under regulations like GDPR and the AI Act. The presented articles emphasize the user-centered, privacy-friendly approaches to AAL designs, robust regulatory frameworks, and interdisciplinary methodologies to ensure ethical, trustworthy technologies.

Keywords Privacy · Trust · Active and assisted living · Artificial intelligence

✉ Anton Fedosov
anton.fedosov@fhnw.ch

Aurelia Tamò-Larrieux
aurelia.tamo-larrieux@unil.ch

Christoph Lutz
christoph.lutz@bi.no

Eduard Fosch-Villaronga
e.fosch.villaronga@law.leidenuniv.nl

Anto Čartolovni
anto.cartolovni@unicath.hr

[1] University of Applied Sciences and Arts Northwestern Switzerland, Windisch, Switzerland

[2] University of Lausanne, Lausanne, Switzerland

[3] BI Norwegian Business School, Oslo, Norway

[4] Leiden University, Leiden, The Netherlands

[5] Catholic University of Croatia, Zagreb, Croatia

1 Introduction

The rapid advancement of technology profoundly transforms how we live, work, and interact with each other every day (Kissinger et al., 2021). Today, we have embedded automated systems such as smart appliances to improve our home life and integrate conversational AI systems to help us find information, book a restaurant, or organize our daily lives (Elliott, 2019). As technology evolves, concerns around intrusions in our private sphere and trust in its functionalities will inevitably increase in parallel (Zuboff, 2015; Lindau, 2022).

This focus is especially crucial in the context of Active and Assisted Living (AAL) technologies, which are steadily growing as an affordable and engaging solution for an aging population (Nilsson et al., 2021). European policymakers recognize the potential of AAL technologies to enhance European healthcare systems (European Parliament & Council, 2008), but they also understand that their success highly depends on their wide adoption, trustworthiness and ability to process information in a privacy-friendly way.[1] That is why more research focuses on ensuring that such AAL technologies are trustworthy and compliant with data protection regulations and individual privacy needs (Dantas et al., 2022). Addressing these challenges, however, requires more of a holistic approach that considers the ethical, legal, and societal aspects (ELSA) of technology development and ensures AAL technologies are effective, reliable, and aligned with regulations, respecting individual rights and integrating societal expectations of *good care* (Ake-Kob et al., 2021).

In this editorial, we present and contextualize some of the results of the COST Action Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living, known as GoodBrother.[2] This pan-European network aimed to raise awareness of the ethical, legal, and privacy challenges associated with audio- and video-based monitoring in assisted living arrangements. GoodBrother fostered an interdisciplinary community of researchers, industrial partners, and other stakeholders across computing, engineering, healthcare, law, sociology, and adjacent disciplines. The GoodBrother COST Action sought to stimulate innovation and propose privacy-friendly solutions for assisted living environments by engaging with users, policymakers, and public services.

In this topical collection, we present six papers featured at the GoodBrother International conference on *Privacy-friendly and trustworthy technology for society* in Zagreb in 2022.[3] The event aimed to advance the knowledge on critical ethical concepts such as privacy, trust, and transparency of (AAL) technologies, contributing mainly by extending emerging concepts and themes such as privacy-by-design (Tamò-Larrieux, 2018), interpersonal and systems trust (Fedosov et al., 2023), overtrust (Aroyo et al., 2021), transparency-by-design (Felzmann et al., 2020), and

[1] *See* the Ageing well in the information society: The Ambient Assisted Living (AAL) Programme at: https://eur-lex.europa.eu/EN/legal-content/summary/ageing-well-in-the-information-society-the-ambient-assisted-living-aal-programme.html.

[2] *See* https://goodbrother.eu/.

[3] *See* https://goodbrother.eu/conferences/goodbrother-international-conference-on-privacy-friendly-and-trustworthy-technology-for-society/.

personalized transparency (Felzmann et al., 2019a, b). After opening the call for contributions to a broader set of scholars, we received 13 submissions. The peer review process was organized in several stages following a double-blind peer review model. In the first stage, each submission was reviewed by at least two domain experts. Subsequently, the associate editors met and discussed the results, wrote the meta reviews for each submission summarizing the key points of improvement and offering additional recommendations, and issued a decision: Revise & Resubmit (R&R) or Reject. In the second stage, the authors of R&R submissions have the opportunity to improve their contributions based on the received feedback. Whenever possible, the editors aimed to assign the same reviewers to the resubmitted contributions who reviewed the original submission; otherwise, we recruited additional reviewers. After this round, the editors either accepted the submissions with minor revisions or employed the shepherded process to ensure that the submissions met the scholarly standards of the journal. Six submissions were eventually accepted, two were withdrawn, four were rejected and one was transferred due to a lack of fit and late submission. The final author list featured in this topical collection explores links, overlaps, and solutions between regulations such as the AI Act (AIA), the General Data Protection Regulation (GDPR), and other frameworks such as European Health Data Spaces (EHDS). The contributions are rich and interdisciplinary in nature, spanning the social sciences, legal scholarship, ethics, and research in computing and engineering.

Before elaborating on the individual contribution in this topical collection (Sect. 3), we highlight the relevance of privacy-friendly and trustworthy technology for society in current debates, specifically the ones on the governance of AI (Sect. 2). The push for more trustworthy and human-centric AI requires us to conceptualize the "how" we want technology to be developed and deployed in sensitive areas and think about "who" technology serves while identifying whether existing practices comply with the law. We close the editorial with a summary, reflection and forward-looking research agenda (Sect. 4).

2 Trustworthy Technology that is Respectful of Individual (Privacy) Needs

The rapid and pervasive integration of digital technologies into our daily lives, such as smart home devices and AAL technologies, has led to a digitally mediated society where we increasingly rely on technology to perform tasks and pursue our goals. Pursuing those goals, however, also requires numerous companies and entities to collect, process, and store our personal data, thereby shaping our interactions with the technology and the company providing the technology and service (Jones, 2013; Zuboff, 2015; Jeon & Lee, 2022). Unsurprisingly, both research on the role of trust in technology adoption (Bahmanziari et al., 2003; Choung et al., 2023) and research on privacy implications of such technologies (Davis et al., 1989; Fosch-Villaronga et al., 2020; Lutz et al., 2019, 2024) have significantly developed in recent years, showing that both topics are interconnected (Liu & Tao, 2022; Rooy & Bus, 2010; Søraa et al., 2021; Knight et al., 2024).

The complex relationship between trust, privacy, and technology adoption is apparent not only in the context of the AAL (Lutz et al., 2025) but also in many other domains. A case in point is autonomous vehicles, where the technology has improved, but public trust has paradoxically declined (Stilgoe, 2023). This decline in trust has real societal implications and cannot be solved with better-engineered technology alone. This is why policymakers, such as in the EU (HLEG, 2019) and also in the USA (The White House, 2023), are aiming to institutionalize trustworthy technology through regulation (Tamò-Larrieux et al., 2024). For instance, the High-Level Expert Group (HLEG) on AI (2019) emphasized principles for trustworthy AI, advocating for ethical, accountable, and human-centric systems. The AI Act positions trustworthiness and human-centricity as a key regulatory goal, reflecting the growing societal demand for technologies prioritizing people's needs (Sigfrids et al., 2023; Vetter et al., 2023).

Trust and privacy are interconnected, yet their relationship is complex. Trust in technology hinges on protecting privacy, as respecting individual privacy fosters confidence in the technology and the institutions behind it (Stilgoe, 2023). In a way, Privacy-by-design—a principle enshrined in regulatory frameworks like the GDPR—tries to be an engineering- and organizational-oriented link to ensure privacy-friendly technology that can be trusted. Privacy-by-design ensures that systems respect user data from the design phase onwards and that data protection norms are adhered to throughout the data protection lifecycle. This principle extends beyond compliance; it is about embedding data protection principles (values), such as transparency and fairness, directly into technological systems (Felzmann et al., 2020). In the context of AAL technologies, where interconnected systems support vulnerable individuals, a breach of these principles (e.g., transparency) could strongly impact trust and possibly lead to non-adoption (Ake-Kob et al., 2021; Søraa et al., 2021).

In this complex and interwoven scenario, the human agent, the environment, and the automated system and their interactions greatly influence trust in human-automation interactions (Schaefer et al., 2016). Human agents' intersectional characteristics, such as culture, age, gender, and personality, may affect their inherent tendency to trust and rely on automation (Chien, 2019; Hoffmann et al., 2024; Kaplan et al., 2023). Various internal and external factors can significantly influence the environment. For instance, research in different contexts shows that the system complexity, the level of difficulty of the tasks at hand, and the perceived risks can substantially impact trust (Habib & Hamadneh, 2021; Glikson & Woolley, 2020; Seo & Lee, 2021). Moreover, internal factors such as self-confidence, level of expertise, mood, attentional capacity, and past experiences also play a crucial role in shaping the overall trust in automation (Schaefer, 2016; Kraus et al., 2020). The articles in the topical collection contribute substantially to a better understanding of privacy and trust dynamics in the context of emerging technologies such as AAL.

3 Articles in this Topical Collection

3.1 Health Data Utilization and Anonymization

The opening article by He (2023) *"From Privacy-Enhancing to Health Data Utilisation"* highlights the potential of anonymization and pseudonymization in reconciling privacy and data-sharing needs in healthcare systems. With a specific focus on the European Health Data Space (EHDS) and related legislation, He emphasizes the evolving roles of these techniques in enabling privacy-compliant health data processing. However, the paper also critiques the residual risks of re-identification and the complexities of integrating these techniques into future EU data laws. This raises a critical question with ulterior consequences: How can we balance the strict legal standards for data protection with the flexibility required for technological innovation? A solution proposed in the paper is considering synthetic data and advanced pseudonymization as complementary tools to bridge the gaps left by traditional approaches, so that developers comply with existing regulatory structures while advancing technology development. In short, achieving an equilibrium ensures that health data can be utilized for the good of society without infringing on privacy.

3.2 Misconceptions of Privacy when Using Thermal Cameras in Robots

Naomi Lintvedt's (2023) article *"Thermal Imaging in Robotics as a Privacy-Enhancing or Privacy-Invasive Measure? Misconceptions of Privacy when Using Thermal Cameras in Robots"* explores the role of thermal imaging and its privacy implications in human-robot interactions. The author argues that while robotics often uses thermal imaging to obscure identifiable details such as facial features, the automatic assumption that this technique is inherently more privacy-preserving than traditional RGB cameras is questionable. Lintvedt highlights that personal data encompasses more than visual identification under data protection laws, meaning thermal imaging can still reveal sensitive information, such as body heat patterns, that qualifies as personal data (as physiological data). In this vein, a narrow focus on informational privacy in robotics research may neglect other dimensions of privacy, such as physical and contextual privacy, which are essential, too. Matching sensor choice with the specific purpose of the robot's functions could align the unique data processing capabilities of robots with the privacy ideals of privacy law. A robots' embodiment, task performance, and physical presence demand a more nuanced understanding of privacy in robotics that emphasizes the need to assess invasive technologies' broader societal and ethical impacts rather than relying on simplistic technological fixes.

3.3 Conversational AI and Debiasing Strategies

The paper *"Debiasing Strategies for Conversational AI: Improving Privacy and Security Decision-Making"* by Anna Leschanowsky, Birgit Popp, and Nils Peters (2023) examines methods for reducing biases in conversational AI to enhance decision-making regarding privacy and security. These AI systems process sensitive data, including medical inquiries, financial transactions, and personal conversations. The

paper reviews various debiasing techniques, emphasizing the significance of accessible privacy policies and the role of conversational AI in diminishing biases within algorithmic frameworks. Furthermore, it investigates how these strategies affect user perceptions and choices within conversational AI contexts. By tackling biases and promoting transparency, the goal is to foster AI interactions that value privacy and engender trust among users. Trustworthy AI is achieved through deliberate design and ongoing refinement.

3.4 A Concept of Balance of Interest in the Context of Active Assisted Living

In the paper *"A Concept of Balance of Interest in the Context of Active Assisted Living, "* Kuźmicz (2023) presents the idea of a "balance of interest." He argues that while "balance" is frequently used, it lacks a precise definition, even within legal scholarship. To him, *balance* has four key dimensions: equilibrium, which ensures equivalence between compared elements and requires precise measurement to avoid imbalance; avoidance of extremes, aimed at preventing domination within systems and safeguarding weaker parties, often tied to power dynamics; ideal proportion, which reflects natural principles or contextual preferences in achieving harmonious relationships; and compromise, a conflict-resolution approach that accommodates all interests through mutual concessions while avoiding dominance. By emphasizing these dimensions, the article provides a roadmap for addressing ethical and social dilemmas in designing systems for older adults in AAL applications that align with human rights and social equity.

3.5 Probing for Privacy: A Digital Design Method to Support Reflection of Situated Geoprivacy and Trust

Jessica Megarry, Peta Mitchell, Markus Rittenbruch, Yu Kao, Bryce Christensen, and Marcus Foth (2023) examine geoprivacy and the role of digital cultural probes in understanding users' contextual and affective decision-making about location data sharing. The study is framed through the lens of contextual integrity and investigates the nexus of geoprivacy and trust. After reviewing existing contributions on location-sharing privacy, it integrates methods from digital media studies, design research, and Human-Computer Interaction to design and deploy *TamaGeochi*, a playful technology probe, to prompt reflection on geoprivacy by exploring how situational factors and materiality influence emotional and intellectual responses to location sharing. Findings reveal that trust is a critical affective factor in privacy decisions, suggesting that digital cultural probes can inform user-centered, context-aware, and trustworthy design approaches grounded in real-life privacy experiences.

3.6 Ethical Guidelines for the Application of Generative AI in German Journalism

Lennart Hofeditz, Anna-Katharina Jung, Milad Mirbabaie, and Stefan Stieglitz (2025) turn to generative AI (genAI) to examine the consequences and effects of this technology in journalism practice and media production processes in Germany. They discuss how genAI automates the content generation for diverse news articles

to date and reflect on its ethical implications vis-à-vis its large adoption in the future in typical news production tasks such as content creation, curation, and dissemination. Through a set of interviews with researchers and practitioners with backgrounds in AI-based technologies, journalism, and ethics, they derive requirements for the ethical introduction of genAI in news production practices and propose actionable guidelines for its ethical use in media organizations to increase the trustworthiness of journalistic organisations and products.

4 Conclusion and Future Research Directions

Taken together, the articles featured in this topical collection point to at least three crucial research findings, with directions for future research.

First, regulatory frameworks must stay agile and consider the wide spectrum of tools in the regulatory toolbox, not only to set limits on technology development but also to actively enable and promote specific beneficial features of technological advancements. Balancing privacy and technological innovation is not a new challenge, but it has become more urgent than ever. The European Commission has acknowledged the significance of balancing privacy and trust in its approach to innovation (Newlands et al., 2020). Trustworthy AI requires a collective effort that blends technological innovation with legal, ethical, and technical awareness. Regulations like the AI Act and GDPR lay the groundwork, but their success depends on interdisciplinary collaboration (Tamò-Larrieux et al., 2024). Engineers, ethicists, policymakers, and the public must engage in open dialogue to ensure technologies align with societal values (Lutz & Tamò, 2015). As both He (2023) and Lintvedt (2023) point out in their contributions, privacy protection cannot merely rely on technological solutions but requires taking into account a broader understanding of privacy. Likewise, Jessica Megarry and colleagues (2023) reinforce this idea of thinking about privacy through a more user-centered lens to capture the nuances of this need within different populations. Future research on privacy-friendly and trustworthy technology for society should integrate user-centered methods such as experiments, surveys and interviews with regulatory and technological aspects, for example by testing the usefulness of privacy labels that simplify complex privacy policies (Barth et al., 2021; Garcia et al., 2025; Kelley et al., 2009; Meier & Krämer, 2024; Windl et al., 2022). It can also investigate how user-facing transparency tools, such as those offered by Facebook (Büchi et al., 2023) and Google (Hautea et al., 2020), affect trust (Dogruel, 2019), including issues of overtrust (Aroyo et al. 2021; Wagner et al., 2018), algorithm appreciation (Logg et al., 2019) and the machine heuristic (Sundar & Kim, 2019) on the one hand vs. undertrust, algorithm aversion (Dietvorst et al., 2015) and AI anxiety (Johnson et al., 2017) on the other hand.

Second, the contributions highlight that privacy and trust are foundational to ethical technological advancement, not just trendy terms. Anonymisation and pseudonymisation are vital in safeguarding user data and adhering to data protection regulations. Conversely, debiasing methods enhance decision-making in AI systems and mitigate biases inherent in algorithmic decisions (Mehrabi et al., 2021). Integrating these approaches underlies the idea that privacy-centric, reliable technologies that serve the

interests of users and creators alike is a possibility. Anna Leschanowsky (2023), Jessica Megarry (2023), and their respective colleagues highlight the need for privacy-preserving technologies and their influence on trust in said technologies. Similarly, Hofeditz and colleagues (2025) reflect on the consequences and effects of generative AI use in news production and how this impacts trustworthiness of journalistic organizations and products. Future research could investigate how to integrate emerging anonymization and debiasing techniques, identifying which approaches yield the most robust privacy-preserving outcomes across contexts (Barbano et al., 2021). Comparative studies are needed to evaluate the real-world efficacy of these methods, testing their scalability, adaptability, and user acceptance. Such work should also analyze the interplay between technical solutions and human-centered design principles to ensure lasting trustworthiness in evolving digital environments (Schoenherr et al., 2023).

Third, the contributions highlight the fragility and complexity of balancing different interests or choosing among different tradeoffs (Kuźmicz, 2023). In technology design, every decision comes with certain costs, from economic to social costs. While the contributions in this topical collection share a vision for technology that prioritizes privacy, inclusivity, and trust, the authors acknowledge that its realization requires collective action across multiple fronts and prioritizing design choices over others. Future research might focus on decision-support frameworks and balancing mechanisms that help designers and stakeholders weigh privacy, inclusivity, and trust against practical constraints (Kovari, 2024; Liu et al., 2022). Longitudinal research from a historical, macro-sociological or macro-economic perspective might study how tradeoffs shift as technologies mature, contextual norms evolve, and regulatory landscapes change (Gherhes et al., 2023). Such findings could guide the development of context-aware design heuristics and negotiation strategies that help navigate complexity without sacrificing key values.

With cautious optimism, we conclude our topical collection on privacy-friendly and trustworthy technology for society. While technology presents difficult challenges, innovations and regulatory frameworks underscore humanity's capacity to responsibly navigate the complexities of technological progress.

Acknowledgements This publication is based upon work from COST Action GoodBrother - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology). COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation www.cost.eu.

Author Contributions All authors contributed equally to the manuscript.

Funding Open access funding provided by FHNW University of Applied Sciences and Arts Northwestern Switzerland

Data Availability Not applicable.

Declarations

Ethical Approval and Consent to Participate Not applicable.

Consent for Publication The authors consented for publication.

Competing Interests No conflict of interests exist.

References

Ake-Kob, A., Blazeviciene, A., Colonna, L., Cartolovni, A., Dantas, C., Fedosov, A., Fosch-Villaronga, E., & Tamo-Larrieux, A. (2021). State of the art on ethical, legal, and social issues linked to audio- and video-based AAL solutions. Working Group 1. Social responsibility: Ethical, legal, social, data protection and privacy issues at COST action GoodBrother. *University of Alicante*, 1–56. https://doi.org/10.5281/zenodo.6793617

Aroyo, A. M., De Bruyne, J., Dheu, O., Fosch-Villaronga, E., Gudkov, A., Hoch, H., & Tamò-Larrieux, A. (2021). Overtrusting robots: Setting a research agenda to mitigate overtrust in automation. *Paladyn Journal of Behavioral Robotics*, *12*(1), 423–436.

Bahmanziari, T., Pearson, J. M., & Crosby, L. (2003). Is trust important in technology adoption? A policy capturing approach. *Journal of Computer Information Systems*, *43*(4), 46–54.

Barbano, C. A., Tartaglione, E., & Grangetto, M. (2021). Bridging the gap between debiasing and privacy for deep learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3806–3815).

Barth, S., Ionita, D., De Jong, M. D., Hartel, P. H., & Junger, M. (2021). Privacy rating: A user-centered approach for visualizing data handling practices of online services. *IEEE Transactions on Professional Communication*, *64*(4), 354–373.

Büchi, M., Fosch-Villaronga, E., Lutz, C., Tamò-Larrieux, A., & Velidi, S. (2023). Making sense of algorithmic profiling: User perceptions on Facebook. *Information Communication & Society*, *26*(4), 809–825.

Chien, S. Y., Lewis, M., Sycara, K., Kumru, A., & Liu, J. S. (2019). Influence of culture, transparency, trust, and degree of automation on automation use. *IEEE Transactions on Human-Machine Systems*, *50*(3), 205–214.

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739.

Dantas, C., Hoogendoorn, P., Kryspin-Exner, I., Stuckelberger, A., & Tijink, D. (2022). AAL Guidelines for Ethics, data privacy and security. *AAL Programme.* Retrieved from https://www.aal-europe.eu/wp-content/uploads/2023/01/AAL-Guidelines_Dec.-2022_FINAL.pdf

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). Technology acceptance model. *J Manag Sci*, *35*(8), 982–1003.

Decision 742/2008/EC of the European Parliament and of the Council of 9 July 2008 on the Community's participation in a research and development programme undertaken by several Member States aimed at enhancing the quality of life of older people through the use of new information and communication technologies (Text with EEA relevance). Retrieved from https://eur-lex.europa.eu/legal-content/EN/AUTO/?uri=celex:32008D0742

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114.

Dogruel, L. (2019). Too much information!? Examining the impact of different levels of transparency on consumers' evaluations of targeted advertising. *Communication Research Reports*, *36*(5), 383–392. https://doi.org/10.1080/08824096.2019.1684253

Elliott, A. (2019). *The culture of AI: Everyday life and the digital revolution*. Routledge.

Fedosov, A., Zavolokina, L., Krumhard, S., & Huang, E. M. (2023). This Could Be The Day I Die: Unpacking Interpersonal and Systems Trust in a Local Sharing Economy Community. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3544549.3585744

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2019a). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society*, *6*(1), 1–14. https://doi.org/10.1177/2053951719860542

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamo-Larrieux, A. (2019b). Robots and transparency: The multiple dimensions of transparency in the context of robot technologies. *IEEE Robotics & Automation Magazine*, *26*(2), 71–78.

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Towards transparency by design for artificial intelligence. *Science and Engineering Ethics*, *26*(6), 3333–3361.

Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Gathering expert opinions for social robots' ethical, legal, and societal concerns: Findings from four international workshops. *International Journal of Social Robotics*, *12*(2), 441–458.

Garcia, K., Tamò-Larrieux, A., Meier, Y., Mayer, S., Lutz, C., Guitton, C., Stern, L., Rot, R., & Mulders, S. (2025). Visual privacy: The impact of privacy labels on privacy behaviors online. Under review. Pre-print available on request.

Gherhes, C., Yu, Z., Vorley, T., & Xue, L. (2023). Technological trajectories as an outcome of the structure-agency interplay at the national level: Insights from emerging varieties of AI. *World Development*, *168*, 106252.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660.

Habib, S., & Hamadneh, N. N. (2021). Impact of perceived risk on consumers' technology acceptance in online grocery adoption amid covid-19 pandemic. *Sustainability*, *13*(18), 10221.

Hautea, S., Munasinghe, A., & Rader, E. (2020, April). 'That's Not Me': Surprising Algorithmic Inferences. In *Extended abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–7).

He, Z. (2023). From privacy-enhancing to health data utilisation: The traces of Anonymisation and Pseudonymisation in EU data protection law. *Digital Society*, *2*(2), 17.

Ethics guidelines for trustworthy AI. High-level expert group on artificial intelligence, HLEG on AI, European, & Commission (2019). https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Hofeditz, L., Jung, A. K., Mirbabaie, M., & Stieglitz, S. (2025). Ethical guidelines for the application of generative AI in German journalism. *Digital Society*, *4*(5), 4.

Hoffmann, C. P., Lutz, C., & Ranzini, G. (2024). Inequalities in privacy cynicism: An intersectional analysis of agency constraints. *Big Data & Society*, *11*(1), 1–13. https://doi.org/10.1177/20539517241232629

Jeon, H., & Lee, C. (2022). Internet of things technology: Balancing privacy concerns with convenience. *Telematics and Informatics*, *70*, 101816.

Johnson, D. G., & Verdicchio, M. (2017). AI anxiety. *Journal of the Association for Information Science and Technology*, *68*(9), 2267–2270.

Jones, C. M. (2013). Preserving life, destroying privacy: PICT and the Elderly. *Emerging Pervasive Information and Communication Technologies (PICT) ethical challenges, opportunities and safeguards* (pp. 89–99). Springer Netherlands.

Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in artificial intelligence: Meta-analytic findings. *Human Factors*, *65*(2), 337–359.

Kelley, P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009, July). A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security* (pp. 1–12).

Kissinger, H. A., Schmidt, E., & Huttenlocher, D. (2021). *The age of AI: And our human future*. Hachette UK.

Knight, T., Yuan, X., & Gayle, B., D (2024). Illuminating privacy and security concerns in older adults' technology adoption. *Work Aging and Retirement*, *10*(1), 57–60.

Kovari, A. (2024). AI for decision support: Balancing accuracy, transparency, and trust across sectors. *Information*, *15*(11), 725.

Kraus, J., Scholz, D., Messner, E. M., Messner, M., & Baumann, M. (2020). Scared to trust?–predicting trust in highly automated driving by depressiveness, negative self-evaluations and state anxiety. *Frontiers in Psychology*, *10*, 2917.

Kuźmicz, M. M. (2023). A concept of balance of interest in the context of active assisted living. *Digital Society*, *2*(3), 51.

Leschanowsky, A., Popp, B., & Peters, N. (2023). Debiasing strategies for conversational AI: Improving privacy and security decision-making. *Digital Society*, *2*(3), 34.

Lindau, J. D. (2022). *Surveillance and the Vanishing Individual: Power and privacy in the Digital Age*. Rowman & Littlefield.

Lintvedt, N. (2023). Thermal imaging in Robotics as a privacy-enhancing or privacy-invasive measure? Misconceptions of privacy when using Thermal Cameras in Robots. *Digital Society*, *2*(3), 33.

Liu, K., & Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior*, *127*, 107026.

Liu, B., Pavlou, P. A., & Cheng, X. (2022). Achieving a balance between privacy protection and data collection: A field experimental examination of a theory-driven information technology solution. *Information Systems Research*, *33*(1), 203–223.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103.

Lutz, C., & Tamò, A. (2015). RoboCode-Ethicists: Privacy-friendly robots, an ethical responsibility of engineers? In *Proceedings of the ACM Web Science Conference* (pp. 1–12). ACM.

Lutz, C., Schöttler, M., & Hoffmann, C. P. (2019). The privacy implications of social robots: Scoping review and expert interviews. *Mobile Media & Communication*, *7*(3), 412–434.

Lutz, C., Tamò-Larrieux, A., & Fosch-Villaronga, E. (2024). How social robots affect privacy: Navigating the landscape. In L. Fortunati, & A. Edwards (Eds.), *The De Gruyter Handbook of Robots in Society and Culture* (pp. 179–200). De Gruyter.

Lutz, C., Miguel, C., Mujirishvili, T., Perez-Vega, R., & Fedosov, A. (2025). Social and societal issues in AAL. In A. A. Salah, L. Colonna, & F. Florez-Revuelta (Eds.), *Privacy-Aware Monitoring for Assisted Living*. Springer, forthcoming.

Megarry, J., Mitchell, P., Rittenbruch, M., Kao, Y., Christensen, B., & Foth, M. (2023). Probing for privacy: A digital design method to support reflection of situated geoprivacy and trust. *Digital Society*, *2*(3), 55.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, *54*(6), 1–35.

Meier, Y., & Krämer, N. C. (2024). The privacy calculus revisited: An empirical investigation of online privacy decisions on between-and within-person levels. *Communication Research*, *51*(2), 178–202.

Newlands, G., Lutz, C., Tamò-Larrieux, A., Villaronga, E. F., Harasgama, R., & Scheitlin, G. (2020). Innovation under pressure: Implications for data privacy during the Covid-19 pandemic. *Big Data & Society*, *7*(2), 1–14. https://doi.org/10.1177/2053951720976680

Nilsson, M. Y., Andersson, S., Magnusson, L., & Hanson, E. (2021). Ambient assisted living technology-mediated interventions for older people and their informal carers in the context of healthy ageing: A scoping review. *Health Science Reports*, *4*(1), e225.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, *58*(3), 377–400.

Schoenherr, J. R., Abbas, R., Michael, K., Rivas, P., & Anderson, T. D. (2023). Designing AI using a human-centered approach: Explainability and accuracy toward trustworthiness. *IEEE Transactions on Technology and Society*, *4*(1), 9–23.

Seo, K. H., & Lee, J. H. (2021). The emergence of service robots at restaurants: Integrating trust, perceived risk, and satisfaction. *Sustainability*, *13*(8), 4431.

Sigfrids, A., Leikas, J., Salo-Pöntinen, H., & Koskimies, E. (2023). Human-centricity in AI governance: A systemic approach. *Frontiers in Artificial Intelligence*, *6*, 976887.

Søraa, R. A., Nyvoll, P., Tøndel, G., Fosch-Villaronga, E., & Serrano, J. A. (2021). The social dimension of domesticating technology: Interactions between older adults, caregivers, and robots in the home. *Technological Forecasting and Social Change*, *167*, 120678.

Stilgoe, J. (2023). What does it mean to trust a technology? *Science*, *382*(6676), eadm9782.

Sundar, S. S., & Kim, J. (2019, May). Machine heuristic: When we trust computers more than humans with our personal information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–9). ACM.

Tamò-Larrieux, A. (2018). *Designing for privacy and its legal framework*. Springer.

Tamò-Larrieux, A., Guitton, C., Mayer, S., & Lutz, C. (2024). Regulating for trust: Can law establish trust in artificial intelligence? *Regulation & Governance*, *18*(3), 780–801.

The White House (2023). Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Retrieved from https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

van Rooy, D., & Bus, J. (2010). Trust and privacy in the future internet—a research perspective. *Identity in the Information Society*, *3*, 397–404.

Vetter, D., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Gallucci, A., Gilbert, T. K., Hagendorff, T., van Halem, I., Hickman, E., Hildt, E., Holm, S., Kararigas, G., Kringen, P., Madai, V. I., Mathez, E. W., Tithi, J. J., Westerlund, M., Wurth, R., Zicari, R. V., & Z-Inspection® initiative (2022). (2023). Lessons learned from assessing trustworthy AI in practice. *Digital Society*, *2*(3), 35.

Wagner, A. R., Borenstein, J., & Howard, A. (2018). Overtrust in the robotic age. *Communications of the ACM*, *61*(9), 22–24.

Windl, M., Henze, N., Schmidt, A., & Feger, S. S. (2022, April). Automating contextual privacy policies: Design and evaluation of a production tool for digital consumer privacy awareness. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–18).

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75–89.