Avances en Interacción Humano-Computadora, 2024, Volumen 9, No. 1, pp. 209-213
DOI: 10.47756/aihc.y9i1.169

Privacy-Aware Artificial Intelligence: A Review of Design Principles and Applications

Beatriz A. Álvarez Magallán, Ricardo Acosta-Díaz, Elba A. Morales-Vanegas*

Published: 30 November 2024

Abstract

Artificial intelligence has emerged as a transformative tool in managing personal data, presenting unprecedented opportunities and significant challenges. This review provides an overview of AI's ethical, technological, and legal dimensions in the context of personal data protection. A systematic literature review was conducted to identify key themes and gaps in these areas. Ethically, the findings highlight the importance of transparency, accountability, and privacy as guiding principles for the responsible use of AI. Technologically, advancements in AI offer innovative solutions for safeguarding data; however, challenges persist in ensuring their interoperability and adaptability across various applications. Legally, regulatory frameworks such as the General Data Protection Regulation (GDPR) and Mexico's General Law on Personal Data Protection Held by Obligated Subjects (LGPDPPSO) illustrate progress in safeguarding personal data. Yet, gaps in enforcement mechanisms and inconsistencies across jurisdictions highlight the need for further refinement. This review underscores the necessity of interdisciplinary collaboration to navigate the complexities of AI and personal data protection. By integrating ethical, technological, and legal perspectives, this study aims to contribute to developing AI systems that respect privacy and remarks on the importance of personal data protection-aware artificial intelligence applications while adapting to diverse regulatory environments.

Keywords

Data-Protection aware AI; Artificial Intelligence; Personal Data Protection; Ethical Principles in AI; Artificial Intelligence Regulation.

1 Introduction

In recent years, the use of technology has grown significantly across society. Technological tools have rapidly evolved, with artificial intelligence (AI) being a prominent field increasingly integrated into various domains to optimize processes, enhance the understanding of human learning, and simplify daily activities. The recent pandemic has dramatically accelerated this progress, which compelled individuals to rely on technology as an essential tool for everyday tasks. Consequently, there has been a remarkable increase in personal data collection and use, sparking a growing interest in data privacy and the pressing need to establish regulations and best practices to protect personal information in AI.

2 Background

2.1 Artificial Intelligence

One of the key characteristics of AI is its ability to enable computers to perform many human-like tasks, such as planning, controlling, and decision-making, with a level of autonomy miming human reasoning. AI pursues two main goals: the first, technological, focuses on developing computational systems to perform valuable tasks; the second, scientific, leverages AI models to address and solve questions about human behavior. [3]

Since its inception, AI has been accompanied by regulatory efforts that have evolved over time, aiming to ensure the technology's responsible and ethical development and application.

2.2 Personal data in AI

Personal data is defined as a legal category of information governed by specific rules, which must also be observed in the AI industry. These data are essential for operating various AI systems, as their development involves collecting, storing, analyzing, processing, and interpreting large volumes of data (commonly known as big data). This information is directly applied to generate outputs, actions, or behaviors in intelligent systems [5].

2.3 Personal data protection in AI

The advancement of AI poses significant challenges to personal data protection, as this technology inherently involves handling large amounts of data, including personal information. Currently, the trust level in managing such data remains low due to associated risks [2].

Protecting personal data in massive information use is essential, mainly because this activity is critical for operating technologies replicating or assisting in human activities. Large-scale data breaches can have severe consequences for data subjects, especially when sensitive information is involved [1].

Using personal information in AI systems must respect human rights and adhere to applicable legal frameworks. In this context, the right to personal data protection faces significant challenges,

Álvarez-Magallán, B. A., Acosta-Díaz, R., Morales-Vanegas E. A.
Universidad de Colima
Colima, Mexico.
Email: {balvarez3, acosta, abigail_morales}@ucol.mx

* Corresponding author

primarily the need to restore a humanistic perspective to this right. This approach fosters trust, provides an ethical digital environment, and ensures user certainty [4].

3 Methods

Systematic literature reviews compile and analyze information generated by prior research on a specific topic. Their primary objective in this document is to provide a comprehensive synthesis of multiple studies in a single text, employing rigorous methods that minimize bias and enhance the reliability of results.

This study used the PRISMA Statement [5,7] as the methodological framework. The relevant articles were searched using the following scientific databases: **IEEE Xplore**, **Springer Link**, **Dialnet**, **Taylor and Francis**, and **ACM Digital Library**. Considering the topic's relatively recent nature, the articles' publication date range was limited to the past 10 years.

The keywords used in this search were Artificial Intelligence and personal data protection, combined with the terms privacy and ethics to ensure the comprehensiveness and effectiveness of the search.

Table 1. Search query strings.

Database	Search query	Results
IEEE Xplore	("All Metadata":artificial intelligence) AND ("All Metadata":personal data protection) AND ("All Metadata":privacy)	195
	("All Metadata":artificial intelligence) AND ("All Metadata":personal data protection) AND ("All Metadata":ethics)	9
Springer Link	(“artificial intelligence”) AND ("personal data protection") AND (privacy)	1,249
	(“artificial intelligence”) AND ("personal data protection") AND (ethics)	789
Dialnet (database with Spanish texts)	inteligencia artificial / protección de datos personales / privacidad protección de datos	37
	personales / inteligencia artificial / ética	27
Taylor and Francis	[All: "artificial intelligence"] AND [All: "personal data protection"] AND [All: privacy]	94
	[All: "artificial intelligence"] AND [All: "personal data protection"] AND [All: ethics]	68
ACM DL	[All: "artificial intelligence"] AND [All: "personal data protection"] AND [All: privacy]	49
	[All: "artificial intelligence"] AND [All: "personal data protection"] AND [All: ethics]	28

Additionally, to complete the study selection process, the following inclusion and exclusion criteria were established:

- Articles published in indexed journals.
- Articles addressing personal data protection in different countries.
- Articles presenting an ethical framework applicable to personal data protection.

The complete details of the search strategy are shown in Figure 2, based on the PRISMA methodology flowchart [5,7].

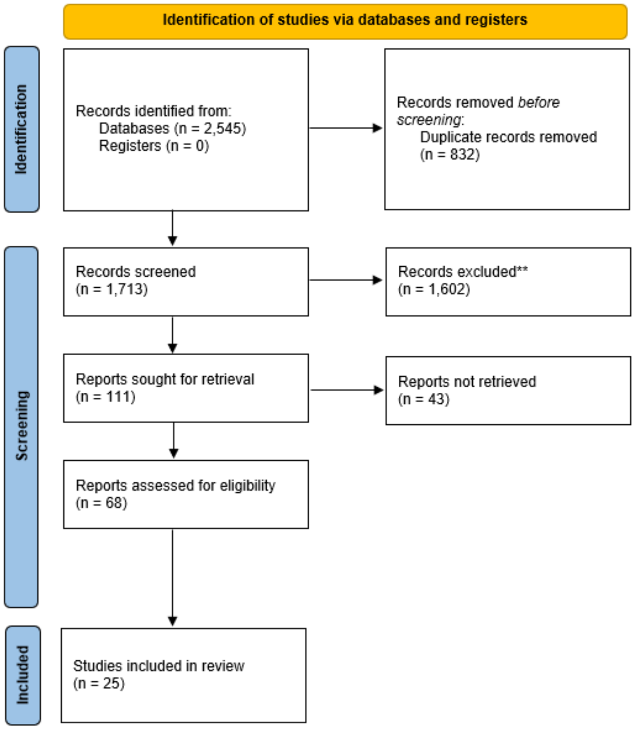

Figure 1. Search strategy flowchart.

4 Results

AI has garnered increasing research interest in personal data, and this systematic review reveals that related studies focus on three key perspectives: ethical, technological, and legal. Although AI development is relatively recent, the reviewed literature highlights significant efforts to understand its impact on personal data protection.

The findings indicate that:

- **Ethical perspective**: The analyzed studies emphasize the need to establish clear ethical principles to guide the development and use of AI, particularly concerning the responsible handling of personal data.
- **Technological perspective**: The reviewed research addresses technical solutions to mitigate privacy risks for users, such as anonymization, encryption, and differential privacy techniques.
- **Legal perspective**: The findings underscore the importance of creating and harmonizing regulatory frameworks to govern the use of personal data in AI systems tailored to the needs of different regions and contexts.

4.1 Ethical perspective

The systematic review highlights various approaches to the ethical perspective in AI, underscoring its importance in auditing, design, and the responsible use of this technology. The main findings from the reviewed articles are summarized below:

Minkkinen et al. [23] examine continuous auditing of AI systems (CAAI), combining AI-specific and constant auditing. Funded by Business Finland, this study evaluates existing frameworks and tools for CAAI. It emphasizes the need for human oversight and automation to mitigate ethical harms and ensure accountability in AI governance. Results classify frameworks based on suitability, identifying seven highly relevant works addressing specific challenges or providing future guidance in AI auditing.

Azam et al. [6] highlight the interplay between privacy, ethics, and cybersecurity in autonomous systems. They identify challenges such as GDPR compliance, transparency in data processing, and the need for innovative privacy-preserving techniques. Additionally, they stress the importance of addressing limitations in traditional threat modeling techniques like LINDDUN and STRIDE to respond to emerging risks.

Albornoz [1] questions the validity of consent as a basis for data processing in AI, arguing its insufficiency given algorithms' technical complexity and opacity. It advocates for global governance frameworks and proactive measures, such as privacy by design and ethical impact assessments, to strengthen user rights protection.

Castellanos [12] discusses the right to be forgotten, automated decision-making, and AI regulation, emphasizing the need to combine ethical and legal perspectives. It highlights the relevance of GDPR and the importance of safeguarding fundamental rights in applications such as education and healthcare.

Fernández-Aller et al. [16] examine how algorithmic biases undermine privacy and equity, proposing audits, impact assessments, and strict regulations to mitigate risks. It underscores the need for transparency, explainability, and regulatory compliance in high-risk AI systems.

Timmers [32] states adopting a global approach to AI ethics governance is essential. It suggests a shift towards an international regulatory framework promoting global public goods and solutions for emerging ethical challenges in cybersecurity and digital autonomy.

Ceross et al. [13] propose a framework for privacy engineering that incorporates privacy by design, impact assessments, and long-term risk management. This holistic approach addresses the complexity of privacy in AI systems, enhancing functionality and user trust.

These findings highlight the importance of ethical reflection in developing AI systems and the need for robust regulations to safeguard privacy and fundamental rights in an ever-evolving digital landscape.

4.2 Technological perspective

This systematic review highlights various technological approaches applied to AI concerning personal data protection and compliance with regulations such as the General Data Protection Regulation (GDPR). The key findings are presented below:

Amaral et al. [3] examine an AI-based approach to evaluate privacy policies under GDPR standards. This work includes the development of a conceptual model, integrity verification criteria, and automated support for classifying policy content. It employs natural language processing (NLP) techniques and machine learning to identify metadata, verify integrity, and improve accuracy in detecting violations. The results demonstrate promising performance, outperforming traditional keyword-based methods.

Lorè et al. [20] introduce the INTREPID framework to ensure GDPR compliance in Italian public documents. This framework employs NLP pipelines and classifiers to predict compliance, using techniques like Bag-of-Words (BoW) and Named Entity Recognition (NER). The article explores the GDPR's impact on public administration and AI's role in facilitating compliance.

Binjubeir et al. [7] discuss privacy-preserving data mining (PPDM) and its application in the data preprocessing phase. Techniques like k-anonymity, l-diversity, and t-closeness are highlighted to anonymize data, minimizing the disclosure of individual identities. This approach protects privacy during data collection and transmission while balancing data utility with security.

Olca and Can [25] present a domain-independent consent management model based on Turkey's Personal Data Protection Law and Semantic Web technologies. This model enables users to manage consent through ontologies such as the Personal Consent Ontology (PCO), defining the necessary metadata for consent management. Its applicability in healthcare is evaluated, demonstrating its effectiveness.

Mitropoulos et al. [24] propose PDGuard, a framework that gives users direct control over their personal data through applied cryptography, access control, and a specific API. This system allows users to define authorization rules, monitor activities, and protect data from internal and external attacks. Prototype results demonstrate an effective architecture for ensuring personal data security, transparency, and control.

Finally, Asghar et al. [4] analyze visual surveillance from a technological and legal perspective, focusing on GDPR compliance. It explores techniques like pseudonymization and encryption alongside machine learning and computer vision tools. This work underscores the importance of designing surveillance systems that align with data protection regulations.

4.3 Legal perspective

The systematic review identifies significant challenges and developments in the legal regulation of artificial intelligence (AI) and its impact on personal data protection. Below are the key findings:

Ruschemeier [28] discusses the legal challenges of AI regulation, focusing on the European Union's Artificial Intelligence Act (AIA) proposal. It explores difficulties in defining AI and the implications of a broad definition for regulation. The material, personal, and territorial scope of the AIA is reviewed, highlighting limitations such as a lack of enforcement mechanisms for affected individuals and exclusions for military AI.

De Laat [14] examines efforts by technology companies like Google, Microsoft, and IBM to adopt responsible AI principles. It emphasizes technical and contextual challenges in ensuring transparency and explainability in AI systems. It highlights the need for independent audits and effective regulatory mechanisms, stressing collaboration between companies and regulators.

Enríquez [15] analyzes the legal challenges of protecting personal data in Mexico within the AI context. It identifies a lack of effective mechanisms to exercise rights, such as access and rectification and the influence of international treaties. A balance between technological benefits and fundamental rights is proposed.

Capdeferro [10] focuses on using and regulating AI in public administration, emphasizing the need for algorithmic transparency and accountability in AI-assisted decisions. It also identifies ethical and legal challenges related to data reuse and rights protection in the public sector.

Martínez Devia [21] examines Colombia’s legislation and its need for updates to address AI and big data challenges. Recommendations include adopting practices like data protection by design and minimizing data collection, emphasizing corporate self-regulation.

Van Ooijen and Vrabec [34] evaluated the GDPR’s impact on consumer empowerment, highlighting challenges such as information overload and process complexity. It reviews provisions like privacy by default and the right to explanation, identifying practical application limitations.

Simbeck [29] compares AI regulation in Schleswig-Holstein, Germany, with initiatives from the European Union and China. It discusses AI definitions, fairness, and transparency in regulation, noting areas of ambiguity and potential improvements.

Smuha [30] explores international regulatory convergence in AI, emphasizing the importance of developing common standards. It also addresses regulatory competition and its impact on fundamental rights.

Tikk [31] highlights the need for a multidisciplinary approach to protecting personal data from government surveillance and private sector data collection. It emphasizes an international coherent framework and active individual participation.

Bu [9] addresses the legal and ethical implications of automatic facial recognition (AFR), emphasizing the need for global governance, ethical design, and adequate regulation to protect fundamental rights. It identifies deficiencies in the current legal framework and proposes creating a new one.

These findings highlight growing concerns about the ethical, legal, and technical challenges artificial intelligence poses in managing personal data. From the need for more precise and adaptive regulations to implementing proactive measures such as transparency, explainability, and accountability in AI systems, the literature reflects a consensus on balancing technological innovation with protecting fundamental rights. This scenario underscores the urgency of adopting multidisciplinary and collaborative approaches to ensure that artificial intelligence is developed and used ethically, somewhat, and with respect for individual privacy.

5 Conclusions and discussion

This study highlights that the development and use of artificial intelligence in managing personal data presents significant challenges in three main dimensions: ethical, technological, and legal. These perspectives offer complementary approaches emphasizing balancing technological innovation with protecting fundamental rights.

The literature underscores the importance of establishing clear ethical principles to guide AI's responsible development and use. Continuous AI Auditing (CAAI) was highlighted to mitigate ethical risks through human oversight and automation. However, practical limitations remain, particularly in applying these tools to complex and autonomous systems.

Solutions such as Privacy-Preserving Data Mining (PPDM) and ontology-based consent management models were identified from a technological perspective. These innovations aim to protect personal data through anonymization, encryption, and differential privacy. However, challenges persist regarding the interoperability of these technologies across domains and their integration into diverse regulatory contexts.

From a legal standpoint, initiatives such as the European Union's Artificial Intelligence Act (AIA) propose comprehensive frameworks for AI regulation. However, gaps in practical implementation and the lack of robust enforcement mechanisms, especially regarding individual rights, remain significant. Additionally, harmonizing international regulations is critical to addressing the global impact of AI.

The findings underscore the urgency of adopting multidisciplinary and collaborative approaches to address the challenges posed by AI in managing personal data. While significant progress has been made across the three studied dimensions, the identified limitations reveal a need for further research to ensure the development of ethical, secure, and data-protection-aware AI systems.

6 Acknowledgments

We would like to express our gratitude to the Mexican National Council for Humanities, Science, and Technology (CONAHCyT) for the financial support provided through the Beatriz A. Álvarez Magallán scholarship (CVU 1242722). We are also grateful to Professor Pedro Santana for his leadership and guidance throughout this project, as well as for his support in improving this paper.

7 References

[1] Albornoz, M. M. (2021). El titular de datos personales, parte débil en tiempos de auge de la Inteligencia Artificial. ¿Cómo fortalecer su posición? Revista IUS, 15(48). https://doi.org/10.35487/rius.v15i48.2021.715

[2] Alexin, Z. (2017). Hungary’s unorthodox approach to personal privacy. Health and Technology, 7(4), 423–440. https://doi.org/10.1007/s12553-017-0181-7

[3] Amaral, O., Abualhaija, S., Torre, D., Sabetzadeh, M., & Briand, L. C. (2021). AI-enabled automation for completeness checking of privacy policies. https://doi.org/10.48550/ARXIV.2106.05688

[4] Asghar, M. N., Kanwal, N., Lee, B., Fleury, M., Herbst, M., & Qiao, Y. (2019). Visual surveillance within the EU general data protection regulation: A technology perspective. IEEE Access: Practical Innovations, Open Solutions, 7, 111709–111726. https://doi.org/10.1109/access.2019.2934226

[5] Ayling, J., & Chapman, A. (2022). Putting AI ethics to work: are the tools fit for purpose? AI and Ethics, 2(3), 405–429. https://doi.org/10.1007/s43681-021-00084-x

[6] Azam, N., Michala, L., Ansari, S., & Truong, N. B. (2023). Data privacy threat modelling for autonomous systems: A survey from the GDPR’s perspective. IEEE Transactions on Big Data, 9(2), 388–414. https://doi.org/10.1109/tbdata.2022.3227336

[7] Binjubeir, M., Ahmed, A. A., Ismail, M. A. B., Sadiq, A. S., & Khurram Khan, M. (2020). Comprehensive survey on big data privacy protection. IEEE Access: Practical Innovations, Open Solutions, 8, 20067–20079. https://doi.org/10.1109/access.2019.2962368

[8] Boden, M. (2017). Inteligencia Artificial. Turner publicaciones.

[9] Bu, Q. (2021). The global governance on automated facial recognition (AFR): ethical and legal opportunities and privacy challenges. International Cybersecurity Law Review, 2(1), 113–145. https://doi.org/10.1365/s43439-021-00022-x

[10] Capdeferro, O. (2020). La inteligencia artificial del sector público: desarrollo y regulación de la actuación administrativa inteligente en la cuarta revolución industrial.

IDP Revista de Internet Derecho y Política, 30. https://doi.org/10.7238/idp.v0i30.3219

[11] Casanovas, P., De Koker, L., Mendelson, D., & Watts, D. (2017). Regulation of Big Data: Perspectives on strategy, policy, law and privacy. Health and Technology, 7(4), 335–349. https://doi.org/10.1007/s12553-017-0190-6

[12] Castellanos Claramunt, J. (2020). La gestión de la información en el paradigma algorítmico: inteligencia artificial y protección de datos. Métodos de Informacion, 11(21), 059–082. https://doi.org/10.5557/iimei11-n21-059082

[13] Ceross, A., & Simpson, A. (2018). Rethinking the proposition of privacy engineering. Proceedings of the New Security Paradigms Workshop.

[14] de Laat, P. B. (2021). Companies committed to responsible AI: From principles towards implementation and regulation? Philosophy & Technology, 34(4), 1135–1193. https://doi.org/10.1007/s13347-021-00474-3

[15] Enríquez, O. A. M. (2021). El derecho de protección de datos personales en los sistemas de inteligencia artificial. Revista IUS, 15(48). https://doi.org/10.35487/rius.v15i48.2021.743

[16] Fernández-Aller, C., & Serrano Pérez, M. M. (2022). ¿Es posible una Inteligencia artificial respetuosa con la protección de datos? Doxa, 45, 307. https://doi.org/10.14198/doxa2022.45.11

[17] Gill, K. S. (2020). Ethics of engagement. AI & Society, 35(4), 783–793. https://doi.org/10.1007/s00146-020-01079-8

[18] Ibarra Cadena Blanca Lilia, Acuña Llamas Francisco Javier, Alcalá MéndezAdrián, Del Río Venegas Norma Julieta, & Román Vergara Josefina. (2022). Recomendaciones para el tratamiento de datos personales derivado del uso de la inteligencia artificial.

[19] Ishii, K. (2019). Comparative legal study on privacy and personal data protection for robots equipped with artificial intelligence: looking at functional and technological aspects. AI & Society, 34(3), 509–533. https://doi.org/10.1007/s00146-017-0758-8

[20] Lorè, F., Basile, P., Appice, A., de Gemmis, M., Malerba, D., & Semeraro, G. (2023). An AI framework to support decisions on GDPR compliance. Journal of Intelligent Information Systems, 61(2), 541–568. https://doi.org/10.1007/s10844-023-00782-4

[21] Martínez Devia, A. (2019). La inteligencia artificial, el big data y la era digital: ¿una amenaza para los datos personales? Revista La Propiedad Inmaterial, 27, 5–23. https://doi.org/10.18601/16571959.n27.01

[22] Mendoza, O. (2022). El derecho de protección de datos personales en los sistemas de inteligencia artificial. 15(48), 179–207. https://doi.org/10.35487/rius.v15i48.2021.743

[23] Minkkinen, M., Laine, J., & Mäntymäki, M. (2022). Continuous auditing of artificial intelligence: A conceptualization and assessment of tools and frameworks. Digital Society: Ethics, Socio-Legal and Governance of Digital Technology, 1(3). https://doi.org/10.1007/s44206-022-00022-2

[24] Mitropoulos, D., Sotiropoulos, T., Koutsovasilis, N., & Spinellis, D. (2020). PDGuard: an architecture for the control and secure processing of personal data. International Journal of Information Security, 19(4), 479–498. https://doi.org/10.1007/s10207-019-00468-5

[25] Olca, E., & Can, O. (2022). DICON: A domain-independent consent management for personal data protection. IEEE Access: Practical Innovations, Open Solutions, 10, 95479–95497. https://doi.org/10.1109/access.2022.3204970

[26] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical Research Ed.), n71. https://doi.org/10.1136/bmj.n71

[27] Román, J. (2022). La Inteligencia Artificial y la Protección de Datos Personales. 3, 39-41.

[28] Ruschemeier, H. (2023). AI as a challenge for legal regulation – the scope of application of the artificial intelligence act proposal. ERA Forum, 23(3), 361–376. https://doi.org/10.1007/s12027-022-00725-6

[29] Simbeck, K. (2022). FAccT-check on AI regulation: Systematic evaluation of AI regulation on the example of the legislation on the use of AI in the public sector in the German federal state of Schleswig-Holstein. 2022 ACM Conference on Fairness, Accountability, and Transparency.

[30] Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. Law, Innovation and Technology, 13(1), 57–84. https://doi.org/10.1080/17579961.2021.1898300

[31] Tikk, E. (2017). Privacy online: up, close and personal. Health and Technology, 7(4), 489–499. https://doi.org/10.1007/s12553-017-0197-z

[32] Timmers, P. (2019). Ethics of AI and Cybersecurity When Sovereignty is at Stake. Minds and Machines, 29(4), 635–645. https://doi.org/10.1007/s11023-019-09508-4

[33] Urrútia, G., & Bonfill, X. (2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. Medicina clinica, 135(11), 507–511. https://doi.org/10.1016/j.medcli.2010.01.015 van

[34] Ooijen, I., & Vrabec, H. U. (2019). Does the GDPR enhance consumers' control over personal data? An analysis from a behavioural perspective. Journal of Consumer Policy, 42(1), 91–107. https://doi.org/10.1007/s10603-018-9399-7