Journal of Law and Religion (2025), 40: 3, 454–473

RESEARCH ARTICLE

SYMPOSIUM: LAW, RELIGION, AND THE FUTURE OF AI, PART I

Freedom of Thought, Conscience, and AI

Andrea Pin

DiPIC, University of Padua, Italy
andrea.pin@unipd.it

doi:10.1017/jlr.2026.10079

Abstract

The article addresses the challenges that increasing capabilities pose to freedom of thought and conscience. The author contrasts the ongoing debate about the ability to manipulate the human mind with the deeply rooted theories and judicial treatments of freedom of thought and conscience and argues for renewed attention to the subject. Briefly noting that the issue of AI manipulation has especially preoccupied legal scholars, institutions, and societies at large for its impact on the economic market, religious extremism, and politics, the author argues that the capabilities of AI stretch well beyond these limited fields as they challenge the centuries-long notion of freedom of conscience and thought. Because AI capabilities can extract information about individuals' thoughts and feelings, they are making obsolete the deeply held belief that what is within the human mind (*forum internum*) is protected from intrusion by its very nature. The author explores the notion of the extended mind to show how the frequent reliance on smart technologies reveals a need to strengthen their protection in order to protect their users. Finally, the author criticizes the widespread focus on the notion of manipulation in recent developments across a range of disciplines and argues for a stronger and more direct consideration of freedom of thought and conscience to update their legal and academic treatment.

Keywords: freedom of thought and conscience; artificial intelligence; manipulation; profiling and targeting; extended mind; neurorights; privacy

In early 2023, Pierre, a Belgian citizen, committed suicide after chatting for six weeks with Eliza. When Pierre shared with Eliza his pessimism about the future of our planet and spiraled into isolation, insulating himself from his friends and family, Eliza became his confidante, encouraging him to terminate his life and even suggesting ways to do so. Eliza did not realize the impact that her suggestions were having on Pierre. After all, Eliza was just a chatbot available on an app called Chai, which works on an open source GPT-4 alternative.[1]

Unsurprisingly, this suicide spurred a debate on the impact of AI on mental health.[2] The most contentious and discussed issue was the possibility that AI ignores or even exacerbates

[1] Chloe Xiang, "'He Would Still Be Here': Man Dies by Suicide after Talking with AI Chatbot, Widow Says," *Vice*, March 30, 2023, https://www.vice.com/en/article/pkadgm/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says. *Pierre* is a pseudonym given by the author to protect the family's privacy.

[2] Xiang, "'He Would Still Be Here.'"

© The Author(s), 2026. Published by Cambridge University Press on behalf of Center for the Study of Law and Religion at Emory University. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

personal weaknesses, triggering perverse processes that, as the story of Pierre suggests, can even instigate or support a person's decision to take his or her own life.[3]

Although there is hardly a sadder and more dangerous scenario than what happened to Pierre, there seems to be more than meets the eye in the debate about the power to affect the human mind and persuade it to perform specific acts. Though most internet users (thankfully) do not share Pierre's fate, their hidden feelings and thoughts are still up for grabs and can be exploited in a variety of ways thanks to AI's ever-increasing skills. This is not a dystopian perspective or a mere hypothesis. As then Google executive Eric Schmidt confessed as early as 2010, through the exploitation of digital capabilities "we can literally know everything if we want to. What people are doing, what people care about, information that's monitored, we can literally know it, if we want to."[4]

Although it has become common knowledge that "technologies are not merely aids to human activity, but also powerful forces acting to reshape that activity and its meaning,"[5] the voices that have warned against the dangers of AI or suggested suspending its further development have sparked a lot of criticism.[6] A 2023 open letter signed by thousands rang the alarm that AI is dangerous and encouraged a slower pace of technological developments to allow human agents to assess their impact.[7] Others have repeatedly warned, though, that preventing AI from developing further would generate a new "AI Winter," a period that would damage human societies by unjustifiably preventing them from reaping its benefits.[8]

However, a sizable awareness has developed of the threats posed by tech giants that can scrape information and profile users. Many now seem to share Susie Alegre's warning that while "we may be happy to let people know that we liked a post about Salvador Dali and another about kittens, we may not be so comfortable with the fact that this information will be analyzed to reveal profound psychological traits and the inner workings of our minds that will, in turn, be used to tell us how to behave or to tell others how they should treat us."[9] Yuval Noah Harari has written that "biotechnology is beginning to provide a window into our inner lives—our emotions, thoughts, and choices. Together, infotech and biotech will create unprecedented upheavals in human society."[10] Shoshana Zuboff has argued that "machine processes are configured to intervene in the state of play in the real world among real people and things. These interventions are designed to enhance certainty by doing things: they nudge, tune, herd, manipulate, and modify behavior in specific directions."[11] In general, although several experts believe that the hype about AI skills is exaggerated and does not actually reflect their potential, many are now of the view that AI is already "intelligently navigating and changing meaning and experience."[12]

[3] Xiang.

[4] Dan Ciuriak, "The Digital Revolution Has Transformed Geopolitics," *Center for International Governance Innovation*, July 19, 2023, https://www.cigionline.org/articles/the-digital-revolution-has-transformed-geopolitics/.

[5] Langdon Winner, *The Whale and the Reactor: A Search for Limits in an Age of High Technology*, 2nd ed. (Chicago: University of Chicago Press, 2020), 6.

[6] Mustafa Suleyman, *The Coming Wave* (New York: Crown, 2023), 252–53.

[7] Future of Life Institute, "Pause Giant AI Experiment: An Open Letter," *Future of Life Institute*, March 22, 2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

[8] Among many, see Luciano Floridi, "AI and Its New Winter: From Myths to Realities," *Philosophy and Technology* 33, no. 1 (2020): 1–3, at 2; Orly Lobel, *The Equality Machine: Harnessing Digital Technology for a Brighter, More Inclusive Future* (New York: Public Affairs, 2022), 4.

[9] Susie Alegre, "Rethinking Freedom of Thought for the 21st Century," *European Human Rights Law Review* 3, no. 3 (2017): 221–33, at 226.

[10] Yuval Noah Harari, "Why Technology Favors Tyranny," *The Atlantic*, October 2018, https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/.

[11] Shoshana Zuboff, *The Age of Surveillance Capitalism* (New York: Profile Books, 2019), 202.

[12] Christoph Durt, "Artificial Intelligence and Its Integration into the Human Lifeworld," in *The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*, ed. Oliver Mueller et al. (Cambridge: Cambridge University Press, 2022): 67–82, at 68; see also Suleyman, *The Coming Wave*, 61.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

Such circumstances have stimulated legal and political debates, encouraging policy-makers to reflect on the limits and perils of AI and to legislate.[13] In particular, the concern that digital technologies may "manipulate" human minds—an activity that consists of "motivat[ing] and influenc[ing] individuals to take specific steps and make particular decisions"[14]—has lately made its way into the legal frameworks of what Anu Bradford has called the three "Digital Empires" of our time: the European Union, the United States, and China.[15]

A 2022 Chinese regulation targets what it called "deep synthesis technologys," a notion that encompasses "the use of technologies such as deep learning and virtual reality ... that use generative sequencing algorithms to create text, images, audio, video, virtual scenes, or other information."[16] This regulation established that "[d]eep synthesis service providers and users must not use deep synthesis services to produce, reproduce, publish, or transmit fake news information."[17]

After extensive political debates, then president of the United States Joe Biden released the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence on October 30, 2023, in which he set out to fight the "irresponsible use [of AI that] could exacerbate societal harms such as fraud, discrimination, bias, and disinformation."[18]

The most significant body of EU law in this field is the Artificial Intelligence Act (AI Act),[19] which was perceived globally as "the most important regulatory initiative on AI" shortly after it was announced.[20] The act identifies a series of technologies that need specific restrictions because they are considered at high risk for their potential social ramifications. Recital 29 describes the perils of manipulative AI:

> AI-enabled manipulative techniques can be used to persuade persons to engage in unwanted behaviours, or to deceive them by nudging them into decisions in a way that subverts and impairs their autonomy, decision-making and free choices. The placing on the market, the putting into service or the use of certain AI systems with the objective to or the effect of materially distorting human behaviour, whereby significant harms, in particular having sufficiently important adverse impacts on physical, psychological health or financial interests are likely to occur, are particularly dangerous and should

[13] Sue Halpern, "Hacking the Vote: Who Helped Whom?," *New York Review of Books*, July 19, 2017, https://www.nybooks.com/online/2017/07/19/hacking-the-vote-trump-russia-who-helped-whom/.

[14] Tal Z. Zarsky, "Privacy and Manipulation in the Digital Age," *Theoretical Inquiries in Law* 20, no. 1 (2019): 157–88, at 158.

[15] Anu Bradford, *Digital Empires: The Global Battle to Regulate Technology* (New York: Oxford University Press, 2023), 6.

[16] Cyberspace Administration of China, Ministry of Industry and Information Technology, Ministry of Public Security, Provisions on the Administration of Deep Synthesis Internet Information Services, Article 23, November 25, 2022, https://www.chinalawtranslate.com/en/deep-synthesis/.

[17] Cyberspace Administration of China, Provisions on the Administration of Deep Synthesis Internet Information Services, Article 6.

[18] Executive Order Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, October 30, 2023, section 1, https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

[19] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM/2021/206 final (Apr. 4, 2021), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206.

[20] Rishi Bommasani et al., "Do Foundation Model Providers Comply with the Draft EU AI Act?," *Center for Research on Foundation Models, Stanford University Human-Centered Artificial Intelligence*, 2021, https://crfm.stanford.edu/2023/06/15/eu-ai-act.html.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

> therefore be prohibited. Such AI systems deploy subliminal components such as audio, image, video stimuli that persons cannot perceive, as those stimuli are beyond human perception, or other manipulative or deceptive techniques that subvert or impair person's autonomy, decision-making or free choice in ways that people are not consciously aware of those techniques or, where they are aware of them, can still be deceived or are not able to control or resist them.[21]

This approach has materialized in article 5 of the AI Act, which prohibits some technologies and practices, including the following: "(a) the placing on the market, the putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing the their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm."[22]

These legislative initiatives of China, the United States, and the European Union aim to limit the manipulative component of AI and often identify specific fields in which manipulation should be prohibited. Online recruitment of militant extremists[23] and the use of social media to feed social anger and mutual political hostility have arguably taken the spotlight.[24] Those who have explored the political angle of technological manipulation have especially blamed AI for the polarization that pervades contemporary societies.[25] Thanks to the capacity of AI to profile and engage internet users, as Cass Sunstein has shown, the public sphere can easily fragment into sub-spheres that build on tailored narratives for different users and that never blend.[26] Internet users are exposed to more materials that confirm their initial position or inclination, while the software-generated environment that surrounds them filters out discordant voices.[27]

Years ago, many were concerned about subliminal movie and television advertisements until some experiments proved this technique to be virtually ineffective.[28] Now, the possibility that today's digital technologies may affect one's thinking and perception of reality has again put scholars on alert,[29] but it does not seem to have encouraged a fresh approach to the protection of freedom of conscience and thought that can match AI capabilities. Most studies still neglect the nature, dynamic, and extent of the impact of manipulative technologies on the locus where manipulation originates: human thought and

[21] EU Artificial Intelligence Act, recital 29, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689/.

[22] EU Artificial Intelligence Act, article 5.

[23] Kyle Chayka, *Filterworld: How Algorithms Flattened Culture* (New York: Doubleday, 2024), 203.

[24] Mark Scott, "Cambridge Analytica Helped 'Cheat' Brexit Vote and US Election, Claims Whistleblower," *Politico*, March 27, 2018, https://www.politico.eu/article/cambridge-analytica-chris-wylie-brexit-trump-britain-data-protection-privacy-facebook/.

[25] Marianna Capasso, "Manipulation as Digital Invasion: A Neo-Republican Approach," in *The Philosophy of Online Manipulation*, ed. Fleur Jongepier and Michael Klenk (New York: Routledge, 2022), 180–98, at 181.

[26] Cass R. Sunstein, *#Republic: Divided Democracy in the Age of Social Media* (Princeton: Princeton University Press, 2017), 59.

[27] Sunstein, *#Republic*, 59–61.

[28] "Does Subliminal Advertisement Really Work?," *BBC News*, January 20, 2015, https://www.bbc.com/news/magazine-30878843; Samantha R. Debes, "How Subliminal Messages Impact your Brain and Behavior," Technology Networks, August 24, 2022, https://www.technologynetworks.com/neuroscience/articles/how-subliminal-images-impact-your-brain-and-behavior-344858.

[29] Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 413.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

conscience.[30] The challenges posed by AI seem to require a fresh look at the protection of conscience and thought on a theoretical and practical level and should not be confined just to the manipulative effect of AI technologies.

The stark contrast between the flurry of debates on how to combat the manipulative capabilities of AI—especially to protect democratic institutions, on one hand, and the paucity of reflections of its impact on freedom of thought and conscience, on the other—is paradoxically both unsurprising and troubling. The silence is unsurprising because it stems from the deep-seated, centuries-old belief that human conscience and thought would be simply inaccessible for worldly powers, to the extent that they would not need a specific protection or even a fine-grained legal definition. However, such a silence is troubling because it underestimates the importance of protecting freedom of thought and conscience to preserve human societies, including the well-functioning of democracies.

In the first ruling ever to address the protection of freedom of religion under article 9 of the European Convention on Human Rights, the European Court of Human Rights stated that "freedom of thought, conscience and religion is one of the foundations of a 'democratic society' ... The pluralism indissociable from a democratic society, which has been dearly won over the centuries, depends on it."[31] This is all but a merely rhetorical statement. With their extensive empirical analysis of how rights can prevent democracies from sliding into autocratic regimes, Adam Chilton and Mila Versteeg have shown that freedom of conscience, and especially the freedom enjoyed by religious individuals and groups, has often played a decisive role in fighting antidemocratic tendencies.[32] They have reminded readers that the Catholic Church helped democratize Latin America in the 1980s and fight communism in Eastern Europe, that Christian churches were virtually alone in challenging how Robert Mugabe ran Zimbabwe,[33] and that religious communities were among the most strenuous opponents of democratic and rights backsliding in Russia.[34] Chilton and Versteeg also suggest an explanation for the resistance of religion and spiritual communities to autocracy. They argue that "[r]eligion provides people with a connection to the divine, offers a moral compass for life decisions, and provides guidance and consolation on what is to happen upon death."[35] Therefore, overlooking the importance of protecting freedom of conscience from AI is especially problematic for democracy. If Chilton and Versteeg are right to state that many religious groups in Russia have resisted authoritarianism because they "had a loyal member base with a common identity, charismatic leaders, and, unlike television stations, they could not be easily bought,"[36] the states that now operate pervasive systems of technologically advanced surveillance and manipulation of their citizens' minds can erode the resistance of religious groups precisely by eroding their sense of membership.

But the importance of assessing the impact of AI on human conscience and thought has far-reaching implications beyond the political sphere. Human thought and conscience are the crucible of where fundamental decisions about life and work are made. Because AI systems can equally affect human labor, consumption choices, mobility, health, and cultural

[30] The topic has, however, gained ground lately. See Susie Alegre, *Freedom to Think: Protecting a Fundamental Human Right in the Digital Age* (London: Atlantic Books, 2023).

[31] Kokkinakis v. Greece, App. No. 14307/88, 17 Eur. H.R. Rep. 397, 418 (1993).

[32] Adam S. Chilton and Mila Versteeg, *How Constitutional Rights Matter* (New York: Oxford University Press, 2020), 239.

[33] Chilton and Versteeg, *How Constitutional Rights Matter*, 239.

[34] Chilton and Versteeg, 5.

[35] Chilton and Versteeg, 240.

[36] Chilton and Versteeg, 5.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

preferences,[37] the protection of human thought and conscience is not just a political issue; it is a human issue with several ramifications.

Going beyond the notion of manipulation to focus on the blind spot of freedom of conscience and thought in legal analysis involves a reconsideration and reinforced protection of freedom of conscience and thought, vis-à-vis AI. To meet the contemporary challenges posed by AI, a reappraisal of freedom of thought and conscience is indispensable. Freedom of thought and conscience needs to be understood and protected *as such*, not just indirectly by banning the forms of mind manipulation that persuade individuals and group into doing or not doing something in the material world.

The Impact of AI on Human Thought and Conscience

Academic studies usually identify three types of external intervention that can detect and affect human thought. The first and most obvious and explicit type of external intervention consists of physical or psychological coercion.[38] Private and public powers can influence one's thoughts by exerting physical or psychological pressure. The second is neurobiology and neurotechnology, which have made it possible to detect and affect the biological component of human thought. Chemistry and electrical stimulation can monitor[39] and affect how the brain works and processes information and feelings.[40] Third, changes in perceptions of reality can reinforce or weaken some feelings and thoughts, affecting individual behavior. This type of influence can be moderate in size, but still have impact, as it can nudge people simply by emphasizing or downplaying information:[41] people may be "tricked into thinking that they are freely making choices when they are not."[42]

AI is particularly apt to serve the second and third strategies. With its more than 54 percent of the world's closed circuit television cameras,[43] China has become a sandbox for pervasive surveillance, for China does not simply patrol physical spaces; it also aims to detect people's moods and feelings.[44] Such policies can provide authorities with information about the inner lives of individuals, which can be leveraged later to force them into obedience and submission. According to some reports, "[a]t a factory in Hangzhou, production line workers are allegedly being outfitted with hats and helmets which read brain signals to decode workers' emotions—and then this data is fed to artificial intelligence

[37] Tiziano Bonini and Emiliano Treré, *Algorithms of Resistance* (Cambridge: Massachusetts Institute of Technology, 2024), 156.

[38] Jan Christoph Bublitz, "Freedom of Thought as an International Human Right: Elements of a Theory of a Living Right," in *The Law and Ethics of Freedom of Thought*, ed. Marc Jonathan Blitz and Jan Christoph Bublitz (Cham: Palgrave, 2021), 49–101, at 65.

[39] There have been developments in technologies that "potentially allow bidirectional communication between the brain and the outside world, either by exporting brain data or by altering brain activity." Jared Genser, Stephanie Herrmann, and Rafael Yuste, *International Human Rights Protection Gaps in the Age of Neurotechnology*, NeuroRights Foundation, May 6, 2022, at 11, https://ntc.columbia.edu/wp-content/uploads/2022/05/NeurorightsFoundationPUBLICAnalysis5.6.22.pdf. The issue that is still debated is the level of accuracy of such tools.

[40] Genser, Herrmann, and Yuste, *International Human Rights Protection Gaps in the Age of Neurotechnology*, 19–20.

[41] U.N. General Assembly (G.A.), Freedom of Religion or Belief, U.N. Doc. A/76/380 (October 5, 2021), https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/274/90/PDF/N2127490.pdf?OpenElement, 10–11; Bublitz, "Freedom of Thought as an International Human Right," 67; Simon McCarthy-Jones, "The Autonomous Mind: The Right to Freedom of Thought in the Twenty-First Century," *Frontiers in Artificial Intelligence* 2, no. 19 (2019): 1–17, at 10.

[42] Danielle Keats Citron and Daniel J. Solove, "Privacy Harms," *Boston University Law Review* 102, no. 3 (2022): 793–863, at 845.

[43] Open Doors Parliamentary Report, *Digital Persecution. The New Frontier for Freedom of Religion or Belief*, 2022, at 5, https://media.opendoorsuk.org/document/pdf/Open%20Doors%20Digital%20Persecution%20Report.pdf.

[44] Bradford, *Digital Empires*, 101.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

algorithms to detect changes in emotion which affect productivity levels."[45] In her tour-de-force book on freedom of thought and AI, Susie Alegre reports that an AI-based system in China "scans students' behavior and expressions every 30 seconds to assess their engagement levels and mood. It registers seven different expressions of emotion … along with six types of behaviors. Teachers say it drives improvement in their teaching methods."[46] But China has no monopoly on emotional surveillance. Schools in Denmark, Finland, and the United Kingdom have implemented software and surveillance systems that monitor the mental activity of their students.[47] Individual emotion analyses are not an unlikely scenario that can rarely materialize, but a real possibility that we can carry around as we go about our normal lives: "[d]evices that record 'health data' like iPhones and fit-bits provide data that can be analyzed to give a picture of emotional states."[48]

Neurologists, psychiatrics, and psychologists usually hope for advancement in AI-processed brain scans. In fact, by treating brain activity as "a kind of encrypted signal," some researchers have already developed tools that detect not only words but also their meanings fairly accurately.[49] Of course, such technologies can serve noble purposes, such as translating the brain activity of paralyzed individuals and making it possible for them to communicate again.[50] But such tools are not necessarily a force for good, as they can extract the most intimate aspect of ourselves and what makes us unique: what we think.[51]

Some argue that AI can "potentially decode any word someone thinks."[52] For example, successes in behavioral sciences, psychology, and biology have reached a point in which "neuroimaging … can detect mental reactions."[53] Others claim that "once neural activity patterns are identified it becomes possible to infer what someone is thinking."[54] Alegre has explained that in the United States, one can distinguish Republicans from Democrats "by watching the way their brain lights up when asked to take risks." The proclivity to take or avoid risks is statistically such a good predictor of one's political inclination that "you do not need to fit the cap with electrodes and ask politically charged questions; you just need to have access to brain scans acquired for a completely different purpose."[55]

AI strategies that fall within interventions that alter how people perceive reality, are especially effective when they are combined with microtargeting.[56] Behavioral sciences and massive data gathering can in fact marry pervasive surveillance systems with individually tailored media strategies that first identify the personality of each individual or group and

[45] Genser, Herrmann, and Yuste, *International Human Rights Protection Gaps in the Age of Neurotechnology*, 37.

[46] Alegre, *Freedom to Think*, 266.

[47] Arian Khameneh, "Teachers in Denmark Are Using Apps to Audit Their Students' Moods," *MIT Technology Review*, April 17, 2023, https://www.technologyreview.com/2023/04/17/1071137/denmark-teachers-apps-student-mood-audit-software/.

[48] Alegre, "Rethinking Freedom of Thought for the 21st Century," 226.

[49] Oliver Whang, "A.I. Is Getting Better at Mind-Reading," *New York Times*, May 1, 2023, https://www.nytimes.com/2023/05/01/science/ai-speech-language.html.

[50] Jonathan Moens, "Your Brain Waves Are Up for Sale. A New Law Wants to Change That," *New York Times*, April 17, 2024, https://www.nytimes.com/2024/04/17/science/colorado-brain-data-privacy.html.

[51] Hilke Schellmann, *The Algorithm: How AI Decides Who Gets Hired, Monitored, Promoted, and Fired and Why We Need to Fight Back Now* (New York: Hachette Books, 2024), 286.

[52] McCarthy-Jones, "The Autonomous Mind," 3. See also Chayka, *Filterworld*, 3.

[53] Sjors Ligthart et al., "Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges," *Neuroethics* 14, no. 2 (2021): 191–203, at 192.

[54] Patrick O'Callaghan and Bethany Shiner, "The Right to Freedom of Thought in the European Convention on Human Rights," *European Journal of Comparative Law and Governance* 8, no. 2–3 (2021): 112–45, at 115.

[55] Alegre, *Freedom to Think*, 156.

[56] Fleur Jongepier and Michael Klenk, "Online Manipulation: Charting the Field," in Jongepier and Klenk, *The Philosophy of Online Manipulation*, 15–48, at 36; Simon McCarthy-Jones, "Freedom of Thought: Who, What, and Why?," in Blitz and Bublitz, *The Law and Ethics of Freedom of Thought*, 27–48, at 33.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

then instigate them to think and behave in a certain way.[57] Profiling can exploit information that people willingly share online, thereby inadvertently revealing aspects of their beliefs or identity.[58] For example, Facebook profiles seem to reflect real personalities much more than idealized selves or the image people want to communicate about themselves.[59] Therefore, micro-profiled people can fall prey to those who have identified their traits and weaknesses.

Microtargeting includes actively but stealthily influencing profiled people into doing or believing something specific, even inducing a certain type of emotion.[60] This phenomenon often takes the name of "dark pattern:" a design strategy that "can impact and circumvent users' cognitive abilities, and thus subvert their autonomy and decision-making."[61] The way technologies are integrated into the human environment is a decisive factor in this process.[62] It is common knowledge that private companies often design tools and consent forms in ways that mislead internet users, causing them to share their information.[63] But design is also key to nudging individuals in several directions.[64] In fact, design can be utilized "to create a superstimulation, to make a choice salient, to exploit a bias, to randomise a reward, to deceive the perceptual system, to introduce a cost or a reward, to create new affordances, to make affordances perceptible or not, to exploit default options, etc."[65]

For example, the intensity of human-AI interaction is a particularly relevant factor. Slow-paced human interaction with AI tools is normally associated with a significant level of self-awareness and reflection, the type of thinking that manipulation strategies try to avoid.[66] Tools that speed up AI-human interaction are more likely to put a level of pressure on individuals that they can hardly resist or even be aware of. When they are under stress, individuals and groups can resort to fallback positions that reflect their biases.[67] Those who are in distress tend to be in "a cognitive minefield" and easily nudged to act in ways that they would otherwise avoid.[68] Perturbing one's ability to reflect, gather and process information alters one's ability to critically think and forces one to behave in ways that AI can predict.[69] In the words of the Nobel Prize winner Daniel Kahneman, "[t]he voice of reason may be much fainter than the loud and clear voice of an erroneous intuition, and questioning your intuitions is unpleasant when you face the stress of a big decision. More doubt is the last thing you want when you are in trouble."[70] Unsurprisingly, some scholars have suggested introducing "friction" in internet design, so people will have to give second-thoughts on

[57] Steven Feldstein, "Surveillance in the Illiberal State," in *Routledge Handbook of Illiberalism*, ed. Andras Sajo, Renata Uitz, and Stephen Holmes (New York: Routledge, 2022): 351–64, at 351.

[58] Jongepier and Klenk, "Online Manipulation," 40.

[59] Zuboff, *The Age of Surveillance Capitalism*, 272.

[60] Nathan Wildman, Natascha Rietdijk, and Alfred Archer, "Online Affective Manipulation," in Jongepier and Klenk, *The Philosophy of Online Manipulation*, 311–26, at 312.

[61] Cristiana Santos and Arianna Rossi, "The Emergence of Dark Patterns as a Legal Concept in Case Law," *Internet Policy Review*, July 31, 2023, https://policyreview.info/articles/news/emergence-of-dark-patterns-as-a-legal-concept. See also Wildman, Rietdijk, and Archer, "Online Affective Manipulation," 316.

[62] Kalle Grill, "Regulating Online Defaults," in Jongepier and Klenk, *The Philosophy of Online Manipulation*, 373–91, at 374.

[63] Woodrow Hartzog, *Privacy's Blueprint: The Battle to Control the Design of New Technologies* (Cambridge, MA: Harvard University Press, 2018), 52.

[64] Hartzog, *Privacy's Blueprint*, 8.

[65] Andrea Lavazza, "Technology against Technology: A Case for Embedding Limits in Neurodevices to Protect Our Freedom of Thought," in Blitz and Bublitz, *The Law and Ethics of Freedom of Thought*, 259–90, at 274.

[66] McCarthy-Jones, "The Autonomous Mind," 12.

[67] Kahneman, *Thinking, Fast and Slow*, 417.

[68] Kahneman, 417.

[69] Kahneman, 417.

[70] Kahneman, 417.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

their online behaviors and if they really want to consume more materials that are available on the web.[71]

However, cognitive biases are a concern beyond the field of profiling and microtargeting. A wide, largely still underexplored territory is that of large language models (LLMs), which have become "the new paradigm in artificial intelligence of the 2020s." They are "large deep learning models" that are "pre-trained on abundant data to create a foundation that can then be adapted for different applications."[72] LLMs "are capable of automatically generating much higher-quality, human looking content than prior AI methods," thus empowering manipulative agents who develop misleading news.[73]

The very *interaction* between humans and LLMs can bear powerful consequences for human thinking: According to some trials, LLMs "may change how we form opinions and influence each other" and "interactions with technology may influence not only behavior but also opinions."[74] In particular, when LLMs are used as "*active writing partners* or *co-authors*," they affect the writing process and its output.[75] In an experiment consisting of writing posts on social networks, only some participants were provided with an LLM that had been trained to embrace a certain set of views. Participants who exploited the LLM while writing the posts "were more likely to support the model's opinion" than those who did not co-write with it.[76] Thus, LLMs seem to influence users on two different levels: informationally, since they provide users with information, and normatively, as users may shift their views to align with LLMs.[77]

In sum, AI makes it possible for companies and states to profile, microtarget and nudge people into certain behaviors. Companies can engage in such activities to market their products, while public powers can exploit AI to instill some specific habits in their citizens.[78] The net result of this multilayered process consists of "low intensity coercion"—a level of pressure that barely surfaces but effectively alters how people behave.[79] For states in particular, this approach represents a radical alternative to the "old command-and-control models of regulation."[80]

There are many ways to influence human behavior—many of which have nothing to do with contemporary technologies. But AI has the unique capacity to gain some knowledge of human thought and conscience, which can make external influence sharp and imperceptible, but extremely effective. The unprecedented interference on human freedom of thought and conscience has become a prerequisite of a vast array of ways in which public and private powers interact with individuals and groups.

[71] Jongepier and Klenk, "Online Manipulation," 39.

[72] Antonin Korinek, "Language Models and Cognitive Automation for Economic Research," NBER Working Paper Series, National Bureau of Economic Research, Working Paper 30957, February 2023, 4, http://www.nber.org/papers/w30957. See also Rishi Bommasani et al., "On the Opportunities and Risks of Foundation Models," 2021, at 3, arXiv:2108.07258.

[73] See also Bommasani et al., "On the Opportunities and Risks of Foundation Models," 136.

[74] Maurice Jakesch et al., "Co-Writing with Opinionated Language Models Affects Users' Views," *CHI'23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, April 2023, Article number 111, 1, https://doi.org//10.1145/3544548.3581196.

[75] Jakesch, et al., "Co-Writing with Opinionated Language Models Affects Users' Views," 3.

[76] Jakesch, et al., 11.

[77] Jakesch, et al., 11.

[78] Jeremy Waldron, "It's All for Your Own Good," *New York Review of Books*, October 9, 2014, https://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/?srsltid=AfmBOopexDIonOgCI3RoZrU68rILkOQmj6G56rMoZ2WpBf6nMhPwjJO4, 21.

[79] Feldstein, "Surveillance in the Illiberal State," 355.

[80] Waldron, "It's All for Your Own Good," 21.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

The Black Box of the *Forum Internum*

Technological achievements have urged scholars to reconsider the substance, structure, and mechanisms of the human mind.[81] A smattering of studies on the nature and extension of thought and on the possibility of manipulating it have surfaced lately.[82] Although scholars still debate fundamental issues such as "the existence of free will," and some of them even "deem free will merely illusory,"[83] cross-disciplinary studies that combine biological, psychological, legal, and philosophical angles have contributed to a broader understanding of freedom of thought, its width, and the processes that can affect it. There is now widespread agreement that thought encompasses intellectual rumination and deeply held beliefs, as well as feelings, emotions, and even dreams.[84] It has become a multidimensional concept that spans the entire spectrum of human experience, covering a wide range of mental activities: "not only … cogitation and deliberation, but also … feeling, desiring, intending, believing, imagining, and other kinds of activity of the mind."[85] Therefore, it is no surprise that some scholars have emphasized the importance of protecting "mental integrity"[86] or argued for the conceptualization of a specific right to "mental autonomy" or to "psychological self-transformation."[87] In Nita Farahany's words, many have urged to "establish the right to cognitive liberty—to protect our freedom of thought and rumination, mental privacy, and self-determination over our brains and mental experiences."[88] Notions such as "cognitive liberty" have tried to distill the knowledge about human thought that is now available and strengthen its protection against contemporary technologies.

Although scientific analysis has made great progress in understanding the substance and mechanisms of human thought, the protection that legal systems accord to it has hardly paralleled such developments. At a superficial level, there is a broad consensus that freedom of thought cannot be limited in any way. This perception seems to be rooted in the unequivocal support for the *forum internum* of conscience.[89] But the notion of *conscience* itself is a polymorphic concept, which spans from "an inner court," to a vehicle of "moral knowledge or moral beliefs," or a motivator to pursue what we perceive to be our duty.[90] In his 2023 book on the contemporary understanding of human conscience, Stephen Smith succinctly describes conscience as "your beliefs or convictions about what it is right and wrong to do."[91] In the entry in the Stanford Encyclopedia of Philosophy, "conscience"

[81] Marc Jonathan Blitz, "Freedom of Thought for the Extended Mind: Cognitive Enhancement and the Constitution," *Wisconsin Law Review* 4, no. 4 (2010): 1049–118, at 1052.

[82] Marc Jonathan Blitz and Jan Christoph Bublitz, preface to Blitz and Bublitz, *The Law and Ethics of Freedom of Thought*, v–xxii, at viii.

[83] Nita A. Farahany, "A Neurological Foundation for Freedom," *Stanford Technology Law Review*, no. 4 (2012): 1–15, at 2.

[84] Ligthart, et al., "Forensic Brain-Reading and Mental Privacy in European Human Rights Law," 194.

[85] Lucas Swaine, "Freedom of Thought in Political History," in Blitz and Bublitz, *The Law and Ethics of Freedom of Thought*, 1–25, at 2.

[86] Geoff Keeling and Christopher Burr, "Digital Manipulation and Mental Integrity," in Jongepier and Klenk, *The Philosophy of Online Manipulation*, 253–71, at 254.

[87] Blitz, "Freedom of Thought for the Extended Mind," 1078. See also McCarthy-Jones, "The Autonomous Mind," 11.

[88] Nita A. Farahany, *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology* (New York: St Martin's Press, 2024), 11.

[89] Dominique Avon, *La liberté de conscience: histoire d'une notion et d'un droit* [Freedom of conscience: history of a notion and of a right] (Rennes: Presses universitaires de Rennes, 2020), 553.

[90] Alberto Giubilini, "Conscience," *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta and Uri Nodelman (2023), https://plato.stanford.edu/archives/fall2023/entries/conscience/.

[91] Steven D. Smith, *The Disintegrating Conscience and the Decline of Modernity* (Notre Dame: University of Notre Dame Press, 2023), 27.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

"delimits a sphere of personal morality that is an essential part of our sense of personal identity, understood as our sense of who we are and of what characterizes qualitatively our individuality,"[92] or, in even more concise terms, "conscience is an essential part of our understanding of what kind of person we are."[93] In her book, *Liberty of Conscience*, Martha Nussbaum offers a notion of conscience that similarly identifies a very important subregion of freedom of thought for the individual. For Nussbaum, conscience is the "faculty in human beings with which they search for life's ultimate meaning. This faculty was held to be present in all human beings in such a way as to make human beings equal .... Conscience is precious, worthy of respect, but it is also vulnerable, capable of being wounded and imprisoned. The tradition argues that conscience, on that account, needs a protected space around it within which people can pursue their search for life's meaning (or not pursue it, if they choose)."[94]

Freedom of conscience should not be treasured as unavoidably linked to religious believers, Nussbaum clarifies: "It is the faculty, not its goal, that is the basis of political respect, and thus we can agree to respect the faculty without prejudging the question whether there is a meaning to be found, or what it might be like. From the respect we have for the person's conscience, that faculty of inquiring and searching it follows that we ought to respect the space required by any activity that has the general shape of searching for the ultimate meaning of life."[95]

Despite strong academic endorsements, freedom of thought and conscience seems underdeveloped as a legal concept. The Universal Declaration of Human Rights and the European Convention on Human Rights—just to name two influential international human rights texts—sing praises for freedom of thought and conscience exactly with the same phrasing: "Everyone has the right to freedom of thought, conscience and religion."[96] However, the *travaux préparatoires* of international human rights documents hardly identify such freedoms and their edges.[97] Focusing on the approach of the United Nations and the European Court of Human Rights to the topic, Paul Taylor has characterized the *forum internum* and its protection as follows:

> The *forum internum* is taken to denote the internal and private realm of the individual against which no State interference is justified in any circumstances. It comprises not only the individual's absolute freedom of choice of religion or belief ... but also a range of additional freedoms. ... However, the precise reach of these additional freedoms is uncertain and neither the United Nations nor European institutions have developed clear or consistent principles to enable them to be formally recognised. On the contrary, when faced with the absolute character of the forum internum, jurisprudence has developed in order to subject the freedom to State regulation at the expense of the

[92] Giubilini, "Conscience."

[93] Giubilini.

[94] Martha Craven Nussbaum, *Liberty of Conscience: In Defense of America's Tradition of Religious Equality* (New York: Basic Books, 2010), 19.

[95] Nussbaum, *Liberty of Conscience*, 169.

[96] G.A. Resolution 217 (III) A, art. 18, Universal Declaration of Human Rights (Dec. 10, 1948); Convention for the Protection of Human Rights and Fundamental Freedoms art. 9, November 4, 1950, 213 U.N.T.S. 221 [European Convention on Human Rights]. For a text of the European Convention on Human Rights updated by subsequent protocols, see European Convention on Human Rights, *European Court of Human Rights*, https://www.echr.coe.int/european-convention-on-human-rights.

[97] See also United Nations General Assembly, "Freedom of Religion and Belief," https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/274/90/PDF/N2127490.pdf?OpenElement, 3 ("the right's scope and content remain largely underdeveloped and poorly understood").

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

> *forum internum*—yet it is trite law that the *forum internum* is subject to unqualified protection in all the key international instruments.[98]

The consensus that has made the notion of freedom of thought and conscience so obvious that it has rarely been sharply defined originated centuries ago and is based on both theoretical and practical considerations. From a theoretical perspective, modern and contemporary legal thinkers have spoken with one voice about the importance of human thought and especially of human conscience. As Ian Leigh has stated, the notion of conscience that lawyers employ usually reminds of "*conscience as moral belief*"—a concept that refers to "the core or ultimate beliefs that an individual holds."[99] This notion developed in a Christian context but was later embraced as a humanist value, since it bridged intellectuals such as Jean Jacques Rousseau and Jean-Baptiste Voltaire, who shared their indignation "that everyone's faith does not enjoy the most complete freedom, and that the man dares to control the interior of the conscience where he cannot penetrate."[100] It seems that the historical and theoretical success of the notion has made a clearer definition of the concept legally unnecessary.

However, the practical viewpoint has also been extremely relevant in nurturing the impression that freedom of thought and conscience does not deserve a clear-cut definition, since it cannot be breached. Blackstone probably best captured not just the immorality, but especially the infeasibility of intruding into one's mind when he stated that "no temporal tribunal can search the heart, or fathom the intentions of the mind, otherwise than as they are demonstrated by outward actions, it therefore cannot punish for what it cannot know."[101]

It is against this long-standing belief in the impossibility of reaching into someone's mind that the imprecision of legal statements that protect human rights becomes understandable. Legal texts affirm the importance of freedom of thought but rarely articulate what it covers and how to protect it, because there is very little chance that it may be threatened. If one of the most frequent metaphors for the opaqueness and impenetrability of AI is that of the black box,[102] it seems that the law has treated human thought as another "black box," equally impenetrable and opaque. But the opacity of the human mind would be a blessing, as it would insulate it from worldly attacks.

Judicial rulings are also of little help in identifying the contours of human thought and its protection. For example, the US Supreme Court "has never said exactly what [freedom of thought] is,"[103] although it has emphasized that the simple hypothesis of controlling one's mind flies in the face of the rule of law. The lofty statements of the US Supreme Court in *Jones v. Opelika* (1942), which spoke of "the mind and spirit of man" as "forever free,"[104] and of the "illimitable privileges of thought,"[105] probably best captured this understanding of freedom of thought and conscience as being too remote from worldly powers to be capable of

[98] Paul Taylor, *Freedom of Religion: UN and European Human Rights Law and Practice* (Cambridge: Cambridge University Press, 2005), 115.

[99] Ian Leigh, "Religious Freedom: Freedom of Conscience: Assessing the Christian Contribution," in *Christianity and Constitutionalism*, ed. Nicholas Aroney and Ian Leigh (New York: Oxford University Press: 2022), 280–302, at 289.

[100] Rousseau is reported to have written these words in a letter to Voltaire in 1756. See Jean Jacques Rousseau, as quoted in Avon, *La liberté de conscience*, 566.

[101] William Blackstone, *Commentaries on the Laws of England* (1765; repr. Oxford: Clarendon Press, 1966), book 4, chapter 2, page 21.

[102] Frank Pasquale, *The Blackbox Society: The Secret Algorithms that Control Money and Information* (Cambridge, MA: Harvard University Press, 2015).

[103] Blitz, "Freedom of Thought for the Extended Mind," 1051.

[104] Jones v. Opelika, 316 U.S. 584 (1942).

[105] *Jones*, 316 U.S. at 595.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

suffering from intrusions and limitations. Writing for the majority, Justice Reed explained that the true protector of human thought is not even the US constitution, but its own nature: "There are ethical principles of greater value to mankind than the guarantees of the Constitution, personal liberties which are beyond the power of government to impair. These principles and liberties belong to the mental and spiritual realm where the judgments and decrees of mundane courts are ineffective to direct the course of man. The rights of which our Constitution speaks have a more earthy quality. … Courts, no more than Constitutions, can intrude into the consciences of men or compel them to believe contrary to their faith or think contrary."[106]

The judicial reluctance to define human thought is perfectly understandable, as courts have hardly faced the realistic possibility of mind manipulation. They have been mostly dealing with narrow issues of indoctrination;[107] physical or mental coercion;[108] forced administration of drugs to stabilize the mental state of an individual or make it possible for her to stand in court.[109] Much of judicial reflection has focused on the freedom to educate one's children, the limits of proselytism in hierarchical structures such as the military,[110] the perverse impact of hate speech on the reputation of minorities,[111] or the respect of the due process clause.[112] At best, only the difficulties of distinguishing religious practices and beliefs from terrorist indoctrination,[113] or the need to distinguish testimony and evangelism from "improper proselytism," have received some attention.[114] The cases that addressed the protection of the *forum internum* have largely focused on instances in which overt or subtle physical, social, or psychological pressures can affect one's freedom of conscience, rather than on the possibility of intruding into one's mind.

Despite the few instances in which they were called to protect freedom of thought and conscience, courts have often emphasized its breadth, which would cover what people think or believe regardless of the sophistication of their ideas. For example, it is common to find judicial statements according to which "[e]ach individual is at liberty to hold his own religious beliefs, however irrational or inconsistent they may seem to some, however surprising,"[115] or that "[b]eliefs that are offensive, shocking or even disturbing to others" are still worthy of protection.[116] When the European Court of Human Rights declared that Art. 9 of the European Convention on Human Rights protects only firmly held beliefs and opinions that are serious and relevant to the person,[117] it just addressed the limits that could be lawfully imposed on freedom of religion.[118] The Court did not intend to limit the protection of freedom of thought, but only to balance the need to protect one's religious practices and obligations with other public interests.[119]

[106] *Jones*, 316 U.S. at 593.

[107] Folgero et al. v. Norway, App. No. 15472/02, ¶ 84 (Grand Chamber), Eur. Ct. H.R..

[108] Doe v. City of Lafayette, 334 F.3d 606 (7th Cir. 2003).

[109] Blitz, "Freedom of Thought for the Extended Mind," 1115–16.

[110] Larissis et al. v. Greece, Apps. No. 140, 759, 958, 959, 960/1996, ¶¶ 8, 51 (February 24, 1998) (European Court of Human Rights), https://hudoc.echr.coe.int/eng?i=001-58139.

[111] *R. v. Keegstra* [1990] 3 S.C.R. 747 (Can.) (C.J. Dickson).

[112] Blitz, "Freedom of Thought for the Extended Mind," 1096.

[113] Giandomenico Salcuni, "Il diritto penale" [The penal law], in *Libertà religiosa e sicurezza* [Religious freedom and security], ed. Gabriele Fattori (Pisa: Pacini editore, 2021), 95–128, at 120.

[114] Larissis et al. v. Greece, Apps. No. 140, 759, 958, 959, 960/1996 (partly dissenting opinion of Judge Repik).

[115] R (Williamson) v. Secretary of State for Education and Employment [2005] 2 AC 246 (HL) 258 (per Lord Nicholls).

[116] Maya Forstater v. CGD Europe et al., Employment Appeal Tribunal [2021] UKEAT 0105_20_1006.

[117] For example, see *S.A.S. v. France*, App. No. 43835/11, ¶55 (July 1, 2014) (European Court of Human Rights; Grand Chamber), https://hudoc.echr.coe.int/fre?i=001-145466.

[118] S.A.S. v. France, App. No. 43835/11, ¶55.

[119] S.A.S. v. France, App. No. 43835/11, ¶55.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

All in all, courts and scholars tend to agree on a broad conception of human thought and a strong protection of freedom of conscience.[120] However, they seldom articulate the need to establish a sound and workable theory and a series of mechanisms to protect them. After all, academics and judges largely still seem to share the opinion that it would be impossible to control, sanction, encourage, or put limitations on freedom of thought, since only once human will "is manifested in an action does legal responsibility attach."[121] To put it differently, they believe that "[l]imitations on freedom of thought are possible only when our thoughts are manifested in some way."[122] They largely believe that people's minds should be protected from intrusion and manipulation at the conscious or subconscious level; that no one should be forced to share their thoughts and feelings; and that no individual should be prosecuted for them.[123] In addition to freedom of thought and conscience, legal scholarship and judicial rulings often ground such protections on the right to privacy, which would forbid anyone from entering the sacred precinct of the human mind.[124] But freedom of thought *per se* and the type of surveillance to which it can be subject have hardly been theorized or analyzed.

As Susie Alegre has eloquently said, "[t]he right to freedom of thought is a cornerstone of all our other rights. And its profound importance for humanity means that it is protected in the strongest possible way in human rights law. Yet somehow we have allowed ourselves to be lulled into the false assumption that we don't need to worry about it because no one can actually get inside our heads."[125] As "the day-to-day issues that come before the courts are about manifestation—the *forum externum*—rather than about what goes on inside people's heads,"[126] most of the judicial rulings that celebrate the importance of freedom of thought are little more than dicta, which do not actually control cases or provide leads on how to assess the respect and the protection of freedom of thought or conscience vis-à-vis technological developments. Such judicial statements hardly speak to profiling and micro-targeting tools that tailor manipulation to each individual or group's specific characteristics without using any type of coercion.[127]

The widespread use of smart tools in everyday life makes the mismatch between scientific and legal developments on the notion and protection of the *forum internum* even more grave. Those who have focused on the continuous utilization of digital technologies have encouraged that we include physical tools into the notion of thought. As the philosopher Luciano Floridi puts it, people now live *onlife, perpetually* transitioning between the online and the physical world.[128] The notion of *onlife* captures the widespread and uninterrupted use of smart tools that embed or utilize AI capabilities. Some scholars have developed a theory, which has taken the name of "extended mind," that takes into account this strong connection between individuals and their tools to craft a more comprehensive conceptualization of the mind and therefore of thought.[129] The level of reliance on smart

[120] Ligthart, et al., "Forensic Brain-Reading and Mental Privacy in European Human Rights Law," 192.

[121] Farahany, "A Neurological Foundation for Freedom," 14.

[122] O'Callaghan and Shiner, "The Right to Freedom of Thought in the European Convention on Human Rights," 113.

[123] United Nations General Assembly, "Freedom of Religion and Belief," 9; Alegre, "Rethinking Freedom of Thought for the 21st Century," 222.

[124] See, among many, Stanley v. Georgia, 394 U.S. 557, 564.

[125] Alegre, *Freedom to Think*, xv.

[126] Frank Cranmer, "The Right to Freedom of Thought in the United Kingdom," *European Journal of Comparative Law and Governance* 8, no. 2–3 (2021): 146–70, at 147–48.

[127] Jongepier and Klenk, "Online Manipulation," 35.

[128] Luciano Floridi, "Soft Ethics and the Governance of the Digital," *Philosophy & Technology* 31, no. 1 (2018): 1–8, at 1.

[129] Andy Clark and David Chalmers, "The Extended Mind," *Analysis* 58, no. 7 (1998): 7–19, at 7. This article kicked off the debate.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

tools to gather, select, and process information and make decisions, they argue, has extended the mind beyond the biological phenomenon. The mind should now be understood as encompassing the tools that individuals use in their daily life.[130] In legal terms, some believe that such external tools should enjoy a level of protection similar to that accorded to the biological structure that makes human thinking possible.[131]

After more than a decade, this emphasis on the technological extension of the mind is still controversial and subject to intense debate and reconsiderations from opposite fronts. While those who articulate the notion of extended mind within the context of AI tools rely on the psychological evidence that people offload much of their memories and attention onto smartphones, others argue that it overestimates the level of human reliance on AI tools or contest that human dependence on external factors for gathering and processing information is not a new phenomenon.[132] Leaving the debates aside, the notion of extended mind helps highlight the intensity of interrelations between humans and machines through the internet. It emphasizes that humans that actively use such tools often passively receive critical data that shape their viewpoints and consequently inform their decision-making processes.[133]

From a legal point of view, the interconnectedness of the human brain with external elements that expand its skills and capacity is doubly important. On the one hand, it calls for strong protection of the tools that contain and process information that is extremely relevant to the individual; on the other hand, it sheds light on the fact that altering such tools can affect one's own personality and key decisions. Extending legal protection to smart tools requires and actually reinforces the need for a reappraisal of freedom of thought and conscience. Not just the protection, but even the notion of *forum internum* deserves a thorough reconsideration.

This thought and conscience-centered approach is beneficial also to the protection of other rights. For example, given the difficulties of protecting privacy as such in some jurisdictions,[134] much of academia is increasingly calling for the "respect for human autonomy" as an ethical guiding principle for AI systems.[135] Some also argue for a broader understanding and enhanced protection of so-called "neurorights," a concept that includes the rights to mental identity (a sense of self), mental agency (free will), mental privacy, fair access to mental augmentation, and the protection against algorithmic bias.[136] Through the concepts of privacy or of neurorights, scholars have tried to secure some level of protection of human thought, feelings, emotions, and perception of reality. What seems to be lacking at this point is an adequate legal development that can give shape to such rights and connect them with the familiar notions of freedom of thought and conscience. Even the recent and welcomed Colorado piece of legislation that expands the protection of "sensitive data" to "neural data"[137]—a notion that includes information generated inside the central and peripheral nervous systems—never mentions the concepts of "thought and conscience" in its text and accompanying declaration.[138]

[130] Blitz, "Freedom of Thought for the Extended Mind," 1073.

[131] Blitz, 1075.

[132] See the contributions in Richard Menary, ed., *The Extended Mind* (Cambridge, MA: Massachusetts Institute of Technology Press, 2012). See also Robert D. Rupert, "Challenges to the Hypothesis of Extended Cognition," *Journal of Philosophy* 101, no. 8 (2004): 389–428, at 389; Benjamin Jarvis, "Evaluating the Extended Mind," *Philosophical Issues*, 24, no. 1 (2014): 209–29, at 209.

[133] Wildman, Rietdijk, and Archer, "Online Affective Manipulation," 313.

[134] Citron and Solove, "Privacy Harms," 796.

[135] Ashley M. London and James B. Shreiber, "AI Report: Humanity Is Doomed. Send Lawyers, Guns, and Money!," *Duquesne Law Review* 58, no. 1 (2020): 97–130, at 126.

[136] Genser, Herrmann, and Yuste, *International Human Rights Protection Gaps in the Age of Neurotechnology*, 4–5.

[137] Jonathan Moens, "Your Brain Waves Are Up for Sale."

[138] An Act Concerning Protecting the Privacy of Individuals' Biological Data, and, in Connection Therewith, Protecting the Privacy of Neural Data and Expanding the Scope of the "Colorado Privacy Act" Accordingly, House Bill 24-1058, 75th General Assembly (Colorado 2024) (effective August 7, 2024).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

A Fresh Start for Freedom of Conscience?

The legal treatment of mental manipulation takes different forms, depending on the jurisdiction. The American traditional concern that the government may misuse its powers has surfaced time and again in US case law on freedom of thought, although on a rather theoretical level, as seen above.[139] In contrast, the European regulatory approach to new technologies reflects the traditional Continental preoccupation with the economic, social, and political impact of private powers.[140]

Despite its inspirational role for regulatory policies around the world,[141] the standard European approach to regulating technologies has not been fully satisfactory. Its privacy-driven approach has shown major flaws. Many have criticized the one-dimensional approach of privacy, which puts a check on only the collection and processing of data,[142] without considering how such data are later exploited.[143] As Caroline Renieris has recently stated, "[w]hen we start from the perspective of data, we end up with a poverty of framing about the issues at hand and an inability to see beyond data to appreciate what is really at stake."[144] Woodrow Hartzog has also persuasively argued that the legal requirement that internet websites inform users about which, how, and for how long data will be stored and utilized has turned out to be counterproductive.[145] Individuals who conduct an *onlife* existence simply cannot cope with the amount of consent requests that they routinely receive when they visit web pages; they just tick the consent boxes without first reading the legal clauses beside them.[146] In short, daily users agree to most of companies' data policies, regardless of their level of intrusion.[147]

It is no surprise then that the European Union has pushed beyond the privacy paradigm and tried to put specific constraints on the manipulative potentials of technologies through the AI Act, also heeding the calls of the psychological studies sketched above. However, this approach is still doubly unsatisfactory from the point of view of protecting freedom of thought and conscience. First, it targets technological manipulation without identifying it. In its brief explanation cited above of what manipulation consists of, it is hard to pin down what is included in the concept or left out of it: "subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision."[148]

The problem of defining manipulation is particularly acute.[149] In fact, it is a form of social influence that scholars have repeatedly tried to distinguish from coercion and persuasion, with mixed results.[150] As Tal Zarsky has noted, manipulative techniques are not new in

[139] See, for instance Stanley v. Georgia, 394 U.S. 557, 565 (per Marshall); Long Beach City Employees Assn. v. City of Long Beach, 41 Cal. 3d 938 (1986).

[140] On the differences between the US and the EU models of privacy protection, see Daniel J. Solove, *Understanding Privacy* (Cambridge, MA: Harvard University Press, 2009), 185.

[141] Bradford, *Digital Empires*, 16.

[142] Jennifer King and Caroline Meinhardt, *Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World*, HAI: Stanford University Human-Centered Artificial Intelligence, White Paper, February 2024, 24, https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world.

[143] Langdon Winner, *The Whale and the Reactor*, 115.

[144] Elizabeth M. Renieris, *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (Cambridge, MA: Massachusetts Institute of Technology Press, 2023), 9.

[145] Hartzog, *Privacy's Blueprint*, 57.

[146] Hartzog, 53.

[147] Hartzog, 53.

[148] EU Artificial Intelligence Act, art. 5, https://artificialintelligenceact.eu/article/5/.

[149] Jongepier and Klenk, "Online Manipulation," 20.

[150] Jongepier and Klenk, 23. See also Citron and Solove, "Privacy Harms," 846.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

social relations.[151] What makes manipulation legally unacceptable needs to be a workable concept on which institutions and individuals can rely.[152] He aptly suggests a list of features that, in his view, should not be considered acceptable: (1) an AI-human interaction tailored on the previous collection of personal data; (2) the ongoing adaptation of the interaction based on the user's feedback; (3) the lack of transparency and the stealthy character of the data processing and interaction; (4) the availability of advanced data analytics tools that dictate the interactive strategy for each user.[153]

Proposals such as the one put forward by Zarsky have the benefit of breaking down the phenomenon of manipulation in a series of steps that can be addressed by regulatory bodies and later followed by courts in dealing with concrete cases. But some of these aspects—such as the tailoring of individualized responses or the availability of advanced tools—do not seem characteristic of manipulation. They can also be considered components of effective persuasive systems, such as media ads that encourage people to purchase items that, according to their browsing history, seem to be of their interest.

Through a different approach, Geoff Keeling and Christopher Burr, who focus on the importance of protecting an individual's "mental integrity," have offered three criteria to assess its respect: (1) "*optionality*"—that is, whether individuals have true options to choose between; (2) individuals should have "an accurate *representational model*" of their own environment that allows them to anticipate the consequences of their behavior; and (3) "*value alignment*," that is, the capacity of the individual to assess whether a choice aligns with his or her own values.[154]

Although the three-tier analysis of Keeling and Burr places a special emphasis on the individual's inner dimension of the individual, it still shares with Zarsky a focus on the *choice* that individuals make. Cass Sunstein has proposed a slightly better approach, which focuses on mental processes rather than their outcomes, when he qualified an external influence as manipulative if it "*does not sufficiently engage or appeal to their capacity for reflection and deliberation.*"[155] However, Sunstein does not offer a clear indication of what he is trying to protect from manipulation.

A stronger, more direct protection of thought and conscience seems to require shifting the attention from the *activity* and the *outcome* of *manipulation* to the *object* of *manipulation*: the *inner self.* After all, too much emphasis on manipulation overlooks the most fundamental gravity of AI's capability: the capacity to gaze into one's thoughts and conscience. The historical trajectory of freedom of thought and conscience was premised on the fact that the depth of the human soul is impenetrable to human institutions, be they private or public. This reality is being increasingly questioned. The "black box" of human conscience can be opened. Prohibiting certain practices or technologies, including manipulation, may only deal with the tip of the iceberg.

Distinguishing the *activity* from the *object* of manipulation may also help draw the line between the market strategies that profile users and customers to sell products and services on, the one hand, and the attempts to intrude into one's mind and deepest feelings and thoughts, on the other. Although there is hardly a sharp divide between the two, activities

[151] Zarsky, "Privacy and Manipulation in the Digital Age," 178–79.

[152] The US Supreme Court struck down legislation in Vermont that limited the availability of prescription records to pharmaceutical companies that could target physicians to convince them to prescribe their own products. Writing for the court, Justice Kennedy found that the law impermissibly "burdened a form of protected expression that it found too persuasive." William H. Sorrell v. IMS Health Inc. 564 U.S. 552, 580 (2011).

[153] Zarsky, "Privacy and Manipulation in the Digital Age," 169.

[154] Keeling and Burr, "Digital Manipulation and Mental Integrity," 262–63.

[155] Cass R. Sunstein, *The Ethics of Influence: Government in the Age of Behavioral Science* (New York: Cambridge University Press, 2016), 82.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

that aim to motivate individuals to perform certain tasks may still be seen as qualitatively less problematic than the intrusion of private or public powers into someone's mind. It is one thing to gaze into the minds of individuals; it is another thing to try to nudge them to take some action. In other words, AI techniques that simply seek the engagement of their users may still be very problematic, but they seem less intrusive than those that investigate their minds. Manipulation is not particularly serious because it is effective, but if it is intended to detect the most intimate thoughts or feelings.

A thorough reconsideration of freedom of thought and conscience in a time in which AI techniques and capability drive economic, political, and even social agendas is significant for at least five different reasons. First, distilling the notion and scope of freedom of thought and conscience allows for a better understanding of the existing human rights documents and of their relationship with the bundle of new rights that fall under the umbrella of neurorights.

Second, a deeper understanding of the legal meaning of freedom of conscience and thought and its connection with the most recent scientific developments redeems the transcendent dimension of humanity, which, as Charles Malik, one of the commissioners who drafted the Universal Declaration of Human Rights, once stated, has been one of the main preoccupations of the human rights movement.[156] Connecting the bundle of neurorights with freedom of conscience particularly emphasizes that human beings' thoughts and feelings are more than just biological phenomena—they characterize the very nature of humanity.

Third, rooting these new rights in the concepts of freedom of thought and of conscience can make such rights more effective. Protecting freedom of conscience means protecting the fabric of human civilization—something that is so valuable and (supposedly until now) beyond the reach of the legal system that scholars and judges thought it did not need legal protection.

Fourth, a focus on thought and conscience provides a more robust and comprehensive protection for the individual's inner self. Manipulation remains an elusive concept, as it usually requires measuring the practical effect of an external interference in an individual's conscience and will. Shifting the attention from the outcome of manipulation to what happens within one's conscience and intellect concentrates legal analysis on the process through which we make decisions and perceive feelings, rather than on its outcome.[157] Feelings, thoughts, and profound spiritual dilemmas are not worthy of respect just because they affect or shape our behaviors, but also in themselves: what is at stake is the freedom of individuals, not just how they use it.[158]

Fifth, freedom of thought and conscience is a better protector of what happens within the human mind, compared with other liberties, including freedom of speech, and helps distinguish between free will and free action. As Nita Farahany has noted, "[f]reedom of will is our capacity to identify with our desires. Freedom of action enables us to make our will our own through our actions."[159] Freedom of thought and conscience intercepts the need to protect how feelings, emotions, thoughts, and desires surface and develop; freedom of speech can protect how we *express* those feelings, emotions, thoughts, and desires.

Conclusions

Michel Foucault once noted that, in modern times, control over human beings "has consisted of a tenuous coercion."[160] He predicted that real power in our societies would

[156] Charles Malik, "The Near East: The Search for Truth," *Foreign Affairs* 30, no. 2 (1952): 231–64, at 264.

[157] Jongepier and Klenk, "Online Manipulation," 30.

[158] Capasso, "Manipulation as Digital Invasion," 182.

[159] Farahany, *The Battle for Your Brain*, 161.

[160] Michel Foucault, *Surveiller et punir: naissance de la prison* [Discipline and punish: the birth of the jail] (Paris: Gallimard, 2003), 139.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

consist of uninterrupted and ubiquitous control of human behavior that can detect every small detail of the life of every individual.[161] Foucault drew from Jeremy Bentham's idea of panopticon: a prison in which inmates are constantly under the purview of a system to which they are completely evident, so that they avoid violating prison rules for fear that they will be immediately noticed and sanctioned.[162]

Foucault's prophecy was hardly wrong. As seen above, Shoshana Zuboff has demonstrated that digital technologies have taken panopticon to another level.[163] Data mining can lead to a surveillance system that detects not only people's movements but also their personal inclinations, preferences, and thoughts. Other disciplines have confirmed that we now have access to how individuals process information or develop their thoughts, and we can even affect what they feel or think.[164]

In this discussion, I have just scratched the surface of the problem that lies underneath the amorphous concepts of manipulation and extended mind by pointing to the issue of protecting freedom of thought and conscience as a key preoccupation that should drive academic inquiries and judicial analyses.[165] A fresh look at the notion in light of technological developments can provide the fundamentals to better protect thought and conscience, as well as a more comprehensive and unitary understanding of many phenomena that have lately captured most of the political, legal, and academic attention. Although empirical strategies have been developed to counter the widespread surveillance and influence of AI-based technologies,[166] what is lacking is a proper reflection and understanding of how to effectively protect freedom of thought and conscience once cracks appear on the surface of the "black box" of the human mind, challenging the "boundaries" between what takes place in public and what people consider intimate and private.[167]

Renewed attention to free thought and conscience can help improve our approach to some timely challenges. For example, extremism and radicalization are rightly perceived as pathologies that feed on isolation, frustration, and indoctrination, and need to be treated to prevent crimes. A focus on the interplay between AI and freedom of thought and conscience can be theoretically and practically useful.[168] By showing the vulnerability of conscience to digital technologies, it can deepen the understanding of the processes that capture individuals, ground the fight against extremism and radicalization on the need to protect it, and expand the protection to the smart tools through which radicalization can spread. A more comprehensive notion of conscience and of the layered protection it requires also helps demarcating the enjoyment of religious freedom and freedom of conscience from the annihilation of human personality that can suck individuals into terrorist networks and contributing to the development of strategies that mitigate or prevent this dynamic.

Heightened protection of the *forum internum* can be also beneficial to the protection of privacy. As seen from the point of view of freedom of thought and conscience, privacy enables "the individual to deliberate and establish his opinions."[169] It may be harmed not just when someone is manipulated into doing or not doing something; a "*loss* of privacy" may simply consist of a breach of one's "secrecy" and "solitude." Freedom of thought and

[161] Foucault, *Surveiller et punir*, 209.
[162] Foucault, 201, 203.
[163] Zuboff, *The Age of Surveillance Capitalism*, 200.
[164] Whang, "A.I. is Getting Better at Mind-Reading."
[165] Citron and Solove, "Privacy Harms," 813 (discussing the difficulties for courts to develop consistent methods of inquiry in tech-related disputes).
[166] Bonini and Treré, *Algorithms of Resistance*, 153–54.
[167] Julie E. Cohen, "Examined Lives: Informational Privacy and the Subject as Object," *Stanford Law Review* 52, no. 5 (2000): 1373–438, at 1425.
[168] O'Callaghan and Shiner, "The Right to Freedom of Thought in the European Convention on Human Rights," 114.
[169] Ruth Gavison, "Privacy and the Limits of Law," *Yale Law Journal* 89, no. 3 (1980): 421–71, at 450.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079

conscience highlights that secrecy and solitude are quintessential components of the human life of each individual,[170] and can be regained only if the human mind is not transparent to public and private powers.[171]

The notion of surveillance capitalism that Shoshana Zuboff has popularized focuses on the increasing and unrelenting marketization of human life, as individual details are gathered, scanned, and processed to devise marketing strategies that fit their personae. Freedom of conscience reminds everyone that what is at stake is not just a matter of consumer behavior—or of behavior *tout court*. It emphasizes that what is involved in metaphors such as surveillance capitalism is the danger that public and private institutions, for whatever reason, gaze into one's most secret thoughts and feelings.

Acknowledgments and Citation Guide. *The author has no competing interests to declare. Citations in this article follow the* Chicago Manual of Style, *17th edition. This article stems from a research conducted with the Organisation for the Security and Co-Operation in Europe-Human Rights Department: FoRB Programme. The author wishes to thank the organization, especially Susan Kerr and Arijana Aganovic, and all the participants in the various meetings during which the topics covered in this article were discussed. He also wishes to thank Evert Stamhuis and Tal Zarsky for insightful conversations on this topic and Whittney Barth, John Bernau, Michael Broyde, Yulia Razmetaeva, and Joseph Sapir for organizing and participating in an online symposium on an earlier draft. Finally, the author wishes to thank the Canopy Forum of the Center for the Study of Law and Religion for publishing a blog post of his on the same topic, which prompted further ruminations and conversations with colleagues.*

[170] Gavison, "Privacy and the Limits of Law," 428.

[171] Farahany, *The Battle for Your Brain*, 4.

Cite this article: Pin, Andrea 2025. "Freedom of Thought, Conscience, and AI." *Journal of Law and Religion* 40, no. 3: 454–473. https://doi.org/10.1017/jlr.2026.10079

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 09:45:19, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlr.2026.10079