

Review

Artificial Intelligence and Paternalism: Redefining Liberalism in the Age of Digitization

Muhtadi Al-Abyadh ^{1,2,*}  and Lorine Kalista Noor ^{3,4}

¹ Centre for Policy Research (CPR), University Sains Malaysia (USM), Bayan Lepas 11900, Malaysia

² Iraqi Reform Office, Ministry of Justice, Baghdad 10001, Iraq

³ Business and Management, Faculty of Business, Management and Information Technology, Universiti Muhammadiyah Malaysia, Padang Besar 02100, Malaysia

⁴ Faculty of Economics, Business and Politics, Universitas Muhammadiyah Kalimantan Timur, Samarinda 75124, Indonesia

* Correspondence: muhtadi@student.usm.my or ra.muhtadi@yahoo.com

Received: 8 August 2025; **Revised:** 16 September 2025; **Accepted:** 2 October 2025; **Published:** 14 November 2025

Abstract: This article explores the evolving role of artificial intelligence (AI) as a paternalistic force and its implications for liberal autonomy in the digital age. It reframes AI not merely as a tool or threat, but as a socio-technical agent whose influence emerges through behavioral guidance, manipulation, and decision-making structures. Drawing on philosophical, sociological, and technological perspectives, the paper introduces the concept of *liberating paternalism* to describe how AI systems subtly reshape human autonomy through voluntary interaction. It identifies four key mechanisms of influence: nudging, manipulation, agency delegation, and ambient governance. Rather than opposing liberal values outright, AI paternalism emerges through widespread reliance on algorithmic systems that structure everyday decisions. This development may signal a potential shift in how autonomy is exercised within liberal societies, raising questions about whether algorithmic governance is gradually reshaping classical liberal assumptions about individual decision-making. Positioned at the intersection of political philosophy and technology ethics, the paper challenges binary framings of freedom and control. It argues that AI-driven paternalism is not imposed but co-constructed, shaped by the user's needs for well-being, survival, and cognitive ease. In doing so, it highlights the urgency of developing new frameworks that address the ethical, behavioral, and structural dimensions of autonomy in algorithmic societies.

Keywords: Artificial Intelligence; Human Autonomy; Paternalism; Liberalism; AI Paternalism; Decision-Making

1. Introduction

Paternalism by artificial intelligence (AI) in human decision-making represents an emerging and ethically complex frontier. It raises foundational questions about autonomy, control, and the evolving relationship between humans and technology [1]. Unlike traditional forms of paternalism, which are typically exercised by institutions, states, or individuals. The concept of liberating paternalism is to describe a new form of parental influence emerging in AI-mediated environments. It refers to a situation where individuals voluntarily delegate portions of their decision-making autonomy to algorithmic systems in exchange for enhanced cognitive efficiency, convenience, or well-being.

AI systems guide or structure human decisions through data-driven prompts, such as healthcare, finance, governance, and digital services, its role in shaping human choices has moved from supportive to directive [2]. What distinguishes this new form of technological paternalism is not only its non-human origin but also its subtlety and scalability. AI systems can influence decisions at an individual and societal level without explicit coercion, often under the guise of personalization, optimization, or efficiency.

Since 2022, the academic discussion on AI or technological paternalism has grown, yet it remains fragmented and largely technical. Most studies focus on machine learning architecture, human-computer interaction, or regulatory oversight [3–7]. Far fewer address the philosophical and social consequences of surrendering autonomy to algorithmic authority. This gap is particularly striking when viewed through the lens of freedom: Do humans, in pursuit of convenience and efficiency, willingly cede autonomy to AI systems? If so, what does this say about human nature in the digital age?

This paper seeks to fill that gap by offering a new philosophical and social interpretation of AI paternalism. Drawing on the concept of *liberating paternalism* [8], as well as contemporary debates on decision-making between human agency and machine intervention [9], it proposes that AI introduces a novel form of governance, one that is neither democratic nor authoritarian, but algorithmic. Unlike traditional power structures (e.g., state-citizen, boss-subordinate, rich-poor), AI lacks human motives or biases, making its interventions seemingly neutral, even benevolent. However, this neutrality is precisely what demands critical examination: Is it truly free of conflict, or does it simply mask a new kind of dependency?

This paper does not attempt to provide an empirical measurement of AI paternalism or to demonstrate its prevalence through surveys or behavioral experiments. Rather, it develops a conceptual framework for understanding how algorithmic systems may redefine liberalism. The analysis therefore focuses on theoretical interpretation and illustrative examples. By answering the normative question: What are the implications of paternalism in artificial intelligence for the liberal understanding of autonomy and individual decision-making?

2. Paternalism and AI: A Philosophical Overview

Paternalism, a concept deeply rooted in political philosophy and ethics, involves interference with an individual's freedom or autonomy with the intent of promoting their well-being or protecting them from harm [10]. While originally associated with familial and state authority, paternalism has re-emerged in modern contexts, particularly in debates surrounding AI. We attempt to explore the philosophical evolution of paternalism, its tension with liberal ideals, and its reconfiguration in the age of intelligent systems. The aim is to critically examine whether AI represents a new form of paternalism, one that is neither political, familial, nor religious, but systemically integrated and often willingly embraced by individuals.

Modern philosophical discussions on paternalism often point to the superiority assumption, the belief that the intervening agent knows what is better for another person [9,11]. Grill (2012) [12] emphasizes that paternalism assumes a moral or intellectual hierarchy, whereby the agent imposes their judgment on others.

Dworkin (1972) [13] defines paternalism as “an interference with a person's freedom of action justified by reasons referring exclusively to the well-being, interests, or values of the person being interfered with.” This framing underscores the ethical dilemma at the core of paternalism: the conflict between individual freedom and external intervention.

In contrast to paternalism, liberal political theory prioritizes individual autonomy, defined as the capacity for rational self-governance and decision-making. Philosophers from the Enlightenment period, such as Immanuel Kant and John Stuart Mill, championed personal freedom as a cornerstone of human dignity and societal progress.

Kant (1785) [14] argued that to be autonomous is to “give oneself the law” to act based on rational principles chosen freely. This idea later evolved to include not only universal law but also one's genuine inclinations and personal goals [15]. In Mill's vision, autonomy is not only a psychological ideal but also a normative imperative: individuals must be free to live according to their own values, and others must respect this freedom [16]. This is called Classical liberalism (negative freedom from interference). But there are contemporary liberalisms such as Welfare liberalism (positive freedom to achieve capabilities) [17], and Rawlsian liberalism (fair equality of opportunity) [18]. Each would generate different analyses of AI's impact.

Liberalism thus stands in philosophical opposition to paternalism, viewing it as an unjustified restriction on autonomy. However, this opposition has never been absolute.

In the 21st century, a new dimension of paternalism is emerging through the rise of AI systems. Increasingly, individuals depend on machines to guide decisions in areas ranging from health and finance to education and social interactions [19]. As this dependence deepens, concerns about the autonomy of machines vs. the autonomy of humans become more urgent [15].

The discussion at hand is that of classical liberalism. The literature review focuses on the ethical problems of artificial intelligence undermining autonomy, a negative intervention as highlighted in the literature, as you will see in this paper. While AI may offer positive freedom in achieving capabilities and equal opportunities, this does not mean it plays a role in decision-making, direction, or manipulation, as we shall see.

Thus, liberating paternalism operates through voluntary technological dependence rather than the imposition of authority. AI-driven systems influence users subtly and pervasively, often without explicit coercion. Through algorithmic nudging, personalized recommendations, and behavioral predictions, AI systems steer users toward particular actions, framing them as beneficial [20,21]. This raises critical questions about consent, agency, and responsibility [22–24].

We argue that this phenomenon represents a new paternalism, distinct from earlier political or familial models. Here, the authority is not enforced by the state or a parent, but by a digital system that users often willingly embrace, trusting it to guide or enhance their decision-making [21,25].

This shift invites a deeper philosophical inquiry: Do humans genuinely desire full autonomy, or do they seek structured guidance, even if it comes from machines? While liberal classical emphasizes freedom and self-rule, behavioral science has shown that many individuals prefer comfort, predictability, and ease over the cognitive load of constant decision-making [26]. AI, in this regard, becomes not merely a tool but a benevolent guide—an invisible hand gently shaping behavior.

The term may appear paradoxical because paternalism traditionally implies a restriction of autonomy. However, in AI contexts, paternalistic guidance can simultaneously reduce cognitive burdens and expand individuals' practical capabilities, enabling them to navigate complex information environments more effectively. In this sense, paternalistic intervention is experienced not as domination but as functional liberation from the limits of human cognitive capacity.

However, most existing literature treats AI paternalism as either an ethical dilemma or a governance issue [4,7,9,27]. Our paper reframes AI paternalism as a social structure in liberal culture, introducing the idea of "*liberating paternalism*" and presenting it not merely as a threat to autonomy, but as a voluntarily accepted social evolution. This framing adds philosophical and sociological depth that goes beyond technical or policy-oriented analyses.

This concept differs from libertarian paternalism, developed by Richard Thaler and Cass Sunstein. Libertarian paternalism refers to institutional design strategies that guide individuals toward choices that enhance their well-being while formally preserving freedom of choice. These interventions are typically implemented by governments or organizations through policy tools such as indirect guidance and choice engineering [7,20].

In contrast, liberating paternalism describes a socio-technological condition rather than a policy tool. Paternalistic influence arises from the ongoing interaction with artificial intelligence systems that users voluntarily adopt in their daily lives. Instead of top-down intervention by policymakers, liberating paternalism emerges from bottom-up technological adoption, where individuals voluntarily integrate algorithmic guidance into their decision-making processes.

3. AI and Nudging

Nudging, a concept introduced by Thaler and Sunstein (2008) [20], is a behavioral science strategy used to influence individuals' choices without restricting their freedom of choice. By altering the environment or how options are presented, nudges encourage people to make decisions that align with their best interests and societal welfare. Although initially applied in fields like public policy and marketing, it has expanded into various domains, most notably, digital technology and AI.

In terms of theoretical foundations. At its core, nudging reflects a form of liberating paternalism: it attempts to steer individuals toward beneficial behavior while preserving freedom of choice [8]. Examples include placing healthier foods at eye level or posting graphic images on cigarette packages to discourage smoking. The individual remains autonomous, but the design of the environment encourages certain behaviors over others. In this context,

paternalism does not equate to coercion but is instead rooted in the belief that institutions or technologies can support better decision-making by aligning individual choices with long-term well-being.

With the rise of AI, nudging has entered a new era. AI systems, by analyzing large volumes of behavioral data, can anticipate user preferences and guide actions in subtle ways. From recommending products on e-commerce sites to optimizing navigation routes through real-time data (e.g., Google Maps), AI systems continuously shape user behavior without explicit directives.

For example, smart assistants such as Alexa and Siri are designed to remember users' routines, helping them manage tasks more efficiently and make smarter choices. Platforms like Spotify, Pandora, and Netflix curate recommendations based on previous user behavior, nudging users toward specific content. Even something as mundane as autocorrect in word processors can be seen as a form of nudging, helping users make "better" choices in writing, while still allowing them to ignore suggestions [22].

This seamless integration of AI into daily life supports the idea of "liberating paternalism" [21,28]. Formosa (2021) [15] argues that AI-enabled nudging is especially prominent in digital marketing, where attention is a valuable and contested resource. AI tools help marketers analyze user engagement, optimize email campaigns, and personalize advertisements for specific demographics. These tools do not force users to engage, but they significantly shape consumption behavior by presenting tailored content that appeals to psychological and emotional triggers. Kühler (2022) [4] similarly points out that health applications embody a form of benevolent paternalism by promoting exercise, balanced diets, and proper sleep patterns—forms of nudging that support individual well-being without undermining freedom of choice.

However, the integration of AI in behavioral guidance raises ethical questions around autonomy and control. Laitinen and Sahlgren (2021) [29] emphasize that the justification for nudging should be grounded in its potential benefits or harms. For instance, nudging someone to choose a healthier meal is far less consequential than nudging someone to accept a medical diagnosis generated by an AI system. Transparency, respect for autonomy, and alignment with the individual's interests are necessary for such interventions to be ethically acceptable.

Moreover, Formosa (2021) [15] goes further to argue that social robots function as "relational artifacts" because they simulate having mental states. This anthropomorphic design fosters natural and intuitive interactions, leading users to treat robots as partners rather than tools, thus deepening the impact of AI-enabled nudging.

AI has transformed nudging from a policy tool into a pervasive, interactive presence in daily life. From smart assistants and shopping algorithms to personalized healthcare and emotional support systems, AI shapes human behavior through subtle environmental cues and personalized guidance. While this shift raises ethical concerns, it also opens new opportunities for promoting individual and collective well-being. The future of nudging may lie in empowering individuals with transparent, adaptive, and user-centered technologies that respect autonomy while offering meaningful support, a new model of digital paternalism grounded in freedom, not control.

4. Manipulation and AI

Manipulation is closely tied to the concept of paternalism, particularly within the context of behavioral or neoliberal paternalism, which influences individuals' choices indirectly. It refers to shaping people's decisions or behaviors without their full awareness or explicit consent, often by subtly altering how information is presented or by modifying the surrounding environment, commonly known as "choice architecture" [26]. Unlike coercion or prohibition, manipulation works through less visible mechanisms, and is often employed in nudging strategies to guide individuals toward what is presumed to be in their best interest [30]. Marketing, for instance, uses manipulation to promote specific products or services. Similarly, governments may use manipulation to influence public consciousness, steering societal attention away from sensitive issues to preserve social and political stability. If nudging targets' behavioral outcomes, manipulation goes a step further by shaping individuals' perceptions and awareness, subtly influencing what they focus on and how they think.

This concept becomes even more powerful with the rise of AI, which can influence decision-making using personal and behavioral data. AI technologies are now widely deployed to build personalized environments that subtly shape user behavior through forms of digital manipulation or deception [31].

Although some scholars, such as Krpan and Urbaník (2024) [32], suggest that AI could support liberal values by enabling more autonomous decision-making, they also acknowledge that AI can serve conflicting purposes. In the same work, they caution that corporations collecting behavioral data through self-quantification, virtual and

augmented reality, and various sensors may use this data to subconsciously steer consumers toward products or decisions they might not have otherwise made. On the surface, such technologies may promote pluralism and free choice, but in reality, they often reflect the interests of those who control technology not the users themselves.

Sunstein (2023) [7] warns that designers of “choice engines” may act in self-interest or with manipulative intent. Rather than addressing users’ cognitive biases, such systems may exploit them, raising serious ethical concerns. One prominent form of AI-driven manipulation is predictive marketing. Through machine learning, companies analyze historical data to anticipate market trends and adjust their strategies accordingly [5]. Platforms like Amazon and Netflix use this approach to offer personalized recommendations in real time, shaping consumer behavior based on prior interactions [2,33].

Another critical area of concern is deceptive media, particularly deepfakes. AI-generated videos, images, or audio can convincingly depict people saying or doing things that never occurred, spreading false information and influencing public opinion [34]. These techniques manipulate users either by presenting entirely fabricated content [28] or by crafting persuasive narratives that align with commercial or political goals [27].

While AI-driven manipulation may appear to preserve user autonomy by allowing individuals to accept or reject suggestions, it often functions within the framework of liberating paternalism, subtly guiding decision-making without removing freedom of choice. This raises important questions about transparency, consent, and the long-term impact of AI on human autonomy.

5. AI Paternalism and Decision-Making

The rise of AI has introduced new forms of intervention in individuals’ decision-making, often justified as benefiting users but potentially occurring at the expense of personal autonomy [13,35,36].

While paternalism has long been debated in democratic societies that prize individual freedom and limited state intervention [37], the increasing reliance on AI to support or even replace human decisions signals a new phase in this debate. Whether in health, finance, education, or daily life, AI now plays a role that can both empower and constrain human agency, reshaping the very fabric of liberal decision-making.

Unlike traditional paternalistic actors, states or companies, AI does not impose decisions through force. Instead, it uses nudging, manipulation and data-driven influence to guide user behavior as we explained [7]. AI systems collect and analyze vast amounts of personal data to make predictions about user preferences, emotions, and behaviors [2]. In marketing, finance, and health, AI offers suggestions that are often accepted without question, simply because they are accurate, fast, and convenient [15,33]. As a result, users willingly surrender aspects of their autonomy in exchange for efficiency and reduced decision-making burdens.

John Danaher (2016) [33] introduces the concept of “algocracy”, a system of governance in which algorithmic processes, rather than human deliberation, become the foundation of decision-making. In such systems, algorithms don’t just assist in choices, they structure the very environment in which choices are made. They organize data, set priorities, and even dictate behavioral norms, often without transparency. This form of governance risks distancing humans from meaningful engagement with the systems that shape their lives, raising urgent questions about accountability, control, and democratic participation.

These decisions are made not under duress, but because AI offers faster, more precise, and less effortful alternatives. This voluntary dependence complicates the liberal notion of autonomy as self-governance and informed choice. Instead, we see a new model emerging one in which autonomy is redefined through collaboration with digital tools, not independent from them.

The medical field offers a revealing case study of AI’s paternalistic role. Traditionally, medical paternalism involved doctors making decisions for patients, sometimes without fully involving them, on the assumption that the doctor knows best [11,12,38]. Liberal critics argue for patient participation in healthcare choices, emphasizing the value of experiential insight [39]. However, with the introduction of AI diagnostic and predictive tools, a new dynamic has emerged: AI now exercises paternalism over both the doctor and the patient.

Doctors increasingly rely on AI-generated data to make clinical decisions [40], and patients are guided by health apps and smart devices that offer personalized medical advice [3,4]. While AI may improve accuracy, it also reduces the space for dialogue, empathy, and individualized judgment, raising concerns about over-dependence and ethical safeguards [5].

In the marketing world, AI’s role in decision-making is more overtly commercial but no less paternalistic.

Through predictive analytics, AI processes enormous datasets to segment consumers, anticipate trends, and personalize advertising content [6, 41]. Algorithms analyze browsing history, purchase behavior, and social media activity to offer product recommendations tailored to individual preferences [29]. While this enhances user experience, it also limits consumer choice by creating digital echo chambers that reinforce existing preferences, subtly steering behavior in commercially beneficial directions.

Although education offers a space for traditional paternalism through the teacher's exercise of educational authority over the student, today AI is participating as a third party among them. Platforms like ChatGPT have become co-educators, assisting students with assignments, research, and test preparation [42–44]. While these tools offer speed and convenience, they may undermine critical thinking, independent learning, and knowledge retention.

Recent studies reveal that over 53% of American college students use ChatGPT to write papers, 48% use it to study for exams, and 22% for paper outlines [45]. Although some students achieve high grades using these tools, educators often struggle to detect their use, raising concerns about long-term academic development. If reliance on AI continues to grow unchecked, it could compromise students' intellectual growth and their ability to engage critically with information.

AI's growing role in decision-making reveals a new type of paternalism one that is algorithmic, data-driven, and often welcomed by users seeking ease and efficiency. While AI offers undeniable benefits, it also redefines the boundaries of autonomy, especially in liberal societies that have historically resisted paternalistic authority. As AI continues to mediate decisions in health, education, marketing, and governance, a central question emerges: Can autonomy be preserved in a world increasingly shaped by algorithms? To address this, future debates must move beyond binary notions of freedom and control and develop new ethical frameworks that balance empowerment with protection, and innovation with responsibility.

6. AI as a Paternalistic Agent

When algorithms practice nudging and manipulating the data available to them, this leads to the decision-making process, which is called Agency Power. AI acts as an agent for both users and programmers, taking on tasks, performing work, and making decisions on their behalf. In the agency, the ethical issue is not about hidden influence as in nudging, but about the extent to which humans relinquish control in favor of algorithmic factors.

In traditional paternalism, paternal authority is subject to an agency. The government acts as an agent of the people, where the people grant agency to the political system through democracy or elections [37, 46]. Similarly, a doctor is the agent of the patient [47], and a lawyer is the agent of the citizen [48], as seen in hard paternalism [49, 50]. In soft paternalism, children are under the agency of parents, and people with disabilities are under the agency of healthcare institutions or family [51]. Here, we are therefore talking about human agents or institutions managed by humans.

However, with AI technologies, the concept of agency has changed, as the agent is now non-human. Although algorithms are designed by developers (humans). However, the algorithm has the ability to update itself based on the data it receives [52]. Consequently, users' interactions with the machine do not adhere to ethical values as they do in human agent-principal relationships [15]. Although some studies confirm the presence of behavioral biases in AI [41]. Unlike human bias, this prejudice does not affect users in the same manner that it does when engaging with other people. To illustrate the point, automation is often designed to increase output while reducing the amount of human involvement. Therefore, most engineering studies have concentrated on developing autonomous system designs, especially for use in robots and driverless cars [28].

A number of studies have shown that robots will eventually become more helpful and sociable. This may include tending to the needs of youngsters, the aged, and the ill, as well as taking on roles as our instructors and coworkers. Eventually, they may even become our friends, sexual partners, and social companions. As a result of these transitions, technology like social robots is no longer seen as a means to an end but rather as an actor with whom we engage in conversation [15, 53].

Rochi (2024) [54] asserts that smart goods as social agents are new paternalistic elements that have emerged with the advent of AI. A smart product is seen as a moral actor that deserves our care and respect when its perceived social effectiveness is high. In addition, "if objects understand what is right and wrong, based on this information, limit or criticize people's actions, they effectively become paternalistic." Because of this, we need to know how pa-

terrestrial technology differs from paternalistic people. The autonomous operation of smart gadgets leaves limited opportunity for human intervention.

Thus, it is humans who grant agency to AI systems. Every user of this technology is responsible for the actions of AI. When we take information, suggestions, or products from AI, it means we are granting it agency through interaction and use. Mark argues that our only option is to hold humans accountable for what AI technology does. Even if AI can act or make decisions, it lacks the necessary capacities for moral action. Therefore, responsibility for its actions or decisions, the actions and decisions that humans delegate to it, must remain with humans [21]. Since AI is not conscious, the agency is through it, and it does not act as a conscious agent. Thus, AI paternalism is exercised through AI on behalf of human beings.

As we mentioned, AI agency is non-human. This means that this agency does not adhere to moral values, respect, or behavioral biases as humans do, but rather to standards designed within the algorithmic system [29]. Users must understand that the standards AI operates on sometimes align with human values and sometimes do not, as human consciousness does not follow fixed standards as AI does. **Figure 1** below illustrates this process.

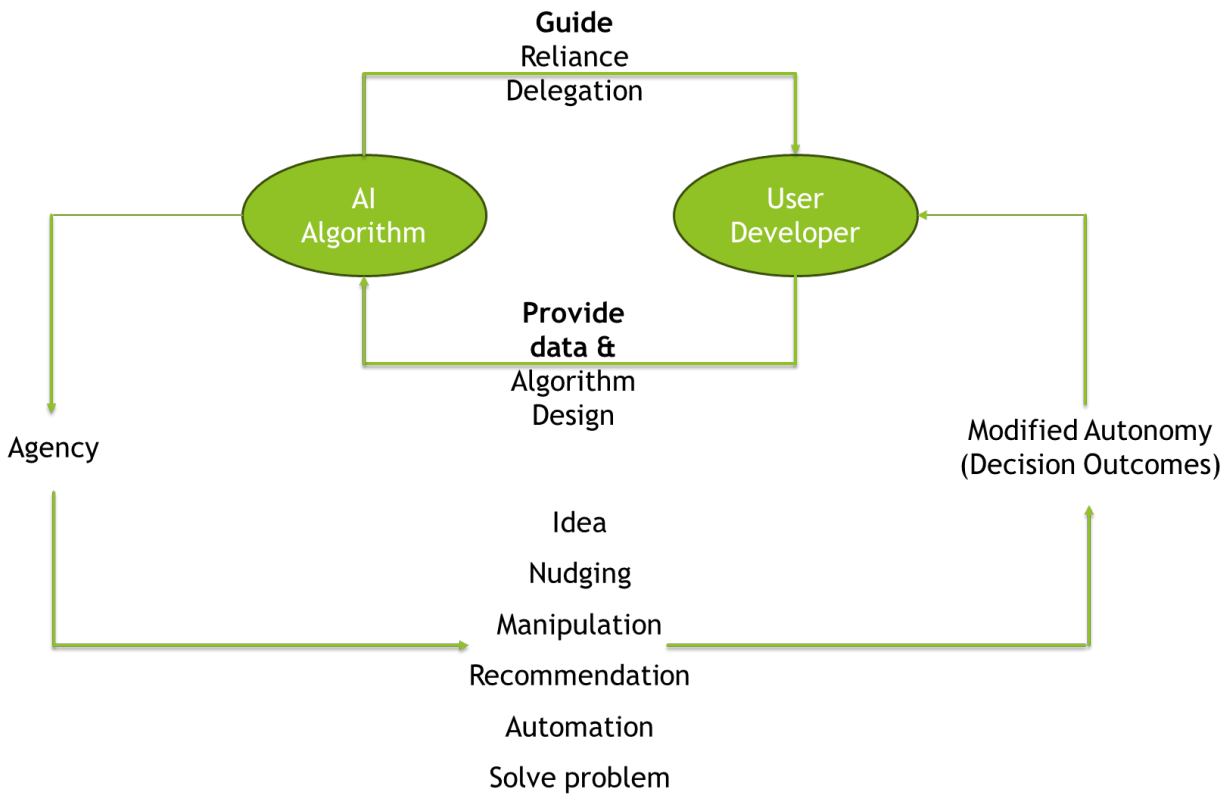


Figure 1. Illustrates the relationship between AI paternalism and user liberalism in the decision-making process.

Figure 1 illustrates what was discussed in the four sections above. An individual’s engagement with AI is entirely left to personal freedom. Developers design algorithmic architectures that enable AI systems to guide user behavior through mechanisms such as nudging, recommendations, and automation. As users increasingly rely on these systems, elements of decision-making may be delegated to algorithmic processes, potentially reshaping the conditions under which individual autonomy is exercised.

This occurs in marketing, advertisements, social media suggestions, and tools like ChatGPT. Although both nudging and manipulation influence human decisions, they differ significantly in their ethical foundations. The first preserves freedom of choice and typically operates through transparent adjustments to the structure of choices. Manipulation, by contrast, involves deceptive or covert strategies that exploit cognitive weaknesses and undermine independent decision-making. In both ways the user may feel a sense of freedom in this process, yet, in reality, AI exercises liberating paternalism, subtly influencing the user without them feeling restricted or as though decisions

are being made on their behalf.

7. Discussion

The intellectual conflict that emerged during the Renaissance between paternalism and liberalism stemmed from a growing revolt against the paternalistic ideology that had long dominated European society. This ideology was rooted in the belief in the king's absolute authority over his subjects, justified by the notion of the monarch as the heir to Adam. In contrast, liberalism offered an alternative framework that challenged paternal authority in politics, religion, and family life. It reframed the debate as one between paternal rights and individual freedom, introducing a new lens through which to understand human nature [55].

Today, with the accelerating integration of AI into human life, there is a renewed need to reconsider human behavior, particularly in the context of the evolving relationship between liberalism and paternalism. The central issue is no longer limited to evaluating the risks or the positive and negative impacts of AI as earlier studies have done but rather lies in reexamining the individual's character in relation to the natural law of survival.

As Hill-Yardin et al. (2023) [56] observe, we are entering a transformative era in the history of intelligent technology, driven by the advent of next-generation AI. This technology is reshaping society at every level, from industry to everyday life to modes of communication. The critical question, then, is: why are individuals increasingly accepting AI-driven paternalism, even as they continue to resist more traditional forms of paternalism, such as that between parent and child or between socioeconomic classes? This paradox invites deeper reflection on the evolving definitions of autonomy, authority, and trust in the age of AI.

7.1. Survival and Well-Being

To begin with, AI can be understood as a simulation of the human brain, encompassing language, behavior, emotions, and more [52,57]. For this reason, AI is becoming increasingly integral to daily life, as it offers functionalities that closely mirror those of the human mind. Research has shown that individuals often interact with social robots as if they were human, an outcome of AI's sophisticated replication of human behavior, effectively transforming human characteristics into machine form [15,53]. Thus, the growing reliance on AI is rooted in two fundamental drivers: survival and well-being.

7.1.1. Need for Survival

One anthropological understanding of human beings is that they strive for survival as an evolutionary trait embedded in their biological constitution, which is reflected in their behavior and interactions with life. Consequently, AI has been able to provide the requirements of survival in a fast and convenient manner. This has led humans to increasingly rely on AI in their daily lives, and even to competition among states in adopting artificial intelligence to maintain a balance of power.

AI may intervene in areas as we mentioned in health, safety, security, armament, and other essential human needs [58]. This situation reflects the paradox of AI paternalism: although people value autonomy, they are often willing to accept the influence of AI when it contributes to enhanced safety and survival.

7.1.2. Need for Well-Being

Beyond mere survival, AI contributes significantly to human well-being through increased convenience, mental health support, and an improved quality of life. Social robots in elder care, for instance, offer companionship and assist with daily tasks, supporting the emotional well-being of vulnerable populations. The therapeutic robot Paro designed for dementia patients is a well-known example. It fosters a sense of companionship, reduces loneliness, and provides patients with a renewed sense of purpose [59]. While such systems often make decisions or act on behalf of individuals, their influence is widely accepted due to the clear emotional and psychological benefits.

Furthermore, AI algorithms on platforms such as Netflix or Spotify enhance well-being by personalizing entertainment experiences. These systems guide user choices based on preferences, yet they are welcomed because they improve ease, enjoyment, and satisfaction in everyday life. This illustrates how AI paternalism is tolerated even embraced, when it aligns with individuals' goals for comfort and happiness.

Another clear example is found in the development of smart automobiles. These vehicles can read traffic signals, interpret weather conditions, assess the driver's behavior, and even intervene to prevent accidents by automatically slowing down or rerouting the car [6]. In these instances, AI acts in the interest of the driver's safety and well-being, even when doing so overrides direct human control.

Sunstein (2023) [7] in his work on libertarian paternalism, supports this interpretation. He argues that when AI subtly guides human decision-making, individuals do not perceive it as a loss of freedom but rather as a form of support that helps them achieve their personal aims, particularly in matters of health, safety, and quality of life.

7.2. AI Biases and Unconscious Paternalism

Although the literature confirms that biases in AI systems can significantly influence user behavior [7,60], these biases differ fundamentally from those observed in human cognition [61].

Human and AI biases diverge in several key areas: nature, awareness, accountability, and influence. Human biases are typically flexible and situational, they can evolve with exposure to new information, empathy, or critical self-reflection [62]. Humans may also become aware of their own biases and make conscious efforts to address them through education or introspection [63]. In contrast, AI biases are more systematic and rigid, embedded within the training data and algorithmic architecture from which these systems are built [61,64]. AI lacks self-awareness, which means the responsibility for identifying and mitigating these biases falls on developers, programmers, and institutional stakeholders.

Despite the presence of these algorithmic biases, users generally do not perceive AI systems as biased or paternalistic. Since AI is not a sentient or emotional being, people do not project interpersonal expectations onto it in the way they might with human authority figures. Moreover, because most users lack a deep understanding of how algorithms function, they remain unaware of how AI may subtly shape their decisions or guide their behavior in ways that serve corporate, political, or technological interests. This leads to an invisible form of paternalism, one that does not feel imposed yet influences user autonomy and decision-making.

Unlike human paternalism, which is usually intentional and moralized, where an authority figure acts "for the good" of another AI paternalism is largely unintentional and systemic. It manifests through recommendation engines, search filters, and ranking algorithms that narrow users' exposure to diverse viewpoints, thereby fostering echo chambers and limiting cognitive diversity [65]. This type of paternalism does not arise from deliberate concern for the user's well-being, but rather from optimization goals, such as maximizing engagement, ad revenue, or user satisfaction based on predicted preferences.

Importantly, while humans may resist traditional paternalism when it comes from other people due to the perceived infringement on freedom, they tend not to recognize similar interventions by machines. The illusion of control plays a role here: users believe they are freely navigating platforms, when in fact their experiences are subtly curated and constrained.

AI-driven paternalism represents a non-conscious, non-emotional form of influence that bypasses the traditional markers of authority, making it both harder to detect and more ethically complex to manage. Unlike human bias, which can be self-correcting, AI bias requires external intervention and ongoing ethical scrutiny. This distinction raises critical questions about freedom, autonomy, and accountability in the digital age.

Since paternalistic dynamics arise as an unintended consequence of relying on these technologies, it falls to the researcher to observe and systematize this phenomenon in response to the question of socio-philosophy, as in this paper on the liberating paternalism exercised through AI over humans. The more humans rely on artificial intelligence, the more they relinquish their autonomy, and the more paternalistic AI becomes towards them.

7.3. AI as a New Paternalism

Paternalism, as traditionally understood in political and philosophical literature, is rooted in pre-modern authority structures as we mentioned. It was a dominant principle in societies where power was centralized in the hands of patriarchs, monarchs, and religious elites. However, the rise of modernity in Europe, driven by the Renaissance, Enlightenment, and industrialization, transformed this landscape. The emergence of liberal political systems in the 19th century marked a move toward individual freedom, democracy, and self-governance.

In Britain, for example, liberal thought clashed with traditional values, leading to a gradual rejection of paternalistic authority [37,66]. Liberal thinkers such as John Stuart Mill (1859) [16] argued that individual freedom is

essential to human development and should only be limited to prevent harm to others. As a result, paternalism in Europe was largely confined to narrow state functions, particularly in the protection of public order [10,12].

Contemporary autonomy scholarship has increasingly emphasized relational understandings of autonomy. Scholars such as Catriona Mackenzie and Natalie Stoljar argue that autonomy is not exercised in isolation but is shaped by social relationships and institutional environments [67,68]. From this perspective, external influences do not necessarily undermine autonomy; rather, they may enable or constrain individuals' capacity for self-direction.

The capability approach developed by Amartya Sen and Martha Nussbaum provides another perspective on autonomy. Rather than focusing solely on whether individuals make decisions independently, this framework emphasizes the real opportunities people possess to pursue valued activities [69,70].

As Jerzy J. Wiatr observes, we have witnessed a phenomenon in democratic transition where elected officials behave like dictators but manage to enjoy a high level of popular support and do not disenfranchise their constituents in contested elections. Political scientist Guillermo O'Donnell has termed this phenomenon "liberal democracy," while Fareed Zakaria has suggested calling such electronic systems "illiberal democrats". Erich Fromm described it as "an escape from freedom" [71].

Today, however, we witness a resurgence of paternalism-not through traditional authorities, but through technology, particularly AI. This new form of paternalism emerges at a time when society highly values autonomy. Yet paradoxically, individuals are voluntarily embracing AI systems that reduce their decision-making burdens and enhance their well-being. We call this phenomenon "AI paternalism", a form of "liberating paternalism".

As John Danaher (2016) [33] notes: "We live in an age of algorithmic decision-making. There are algorithms trading stocks on Wall Street; algorithms determining who is most likely to commit tax fraud; algorithms assisting in scientific discovery; and algorithms helping us in dating and mating... This trend is only expected to grow."

This shift invites a reinterpretation of human nature. While autonomy remains a cherished ideal, humans also exhibit a biological tendency to seek comfort, efficiency, and low-effort solutions, even if it means surrendering some control [71-75]. In this light, AI paternalism is not imposed but embraced through voluntary submission to algorithmic authority. **Figure 2** captures the evolution of paternalism across three major eras:

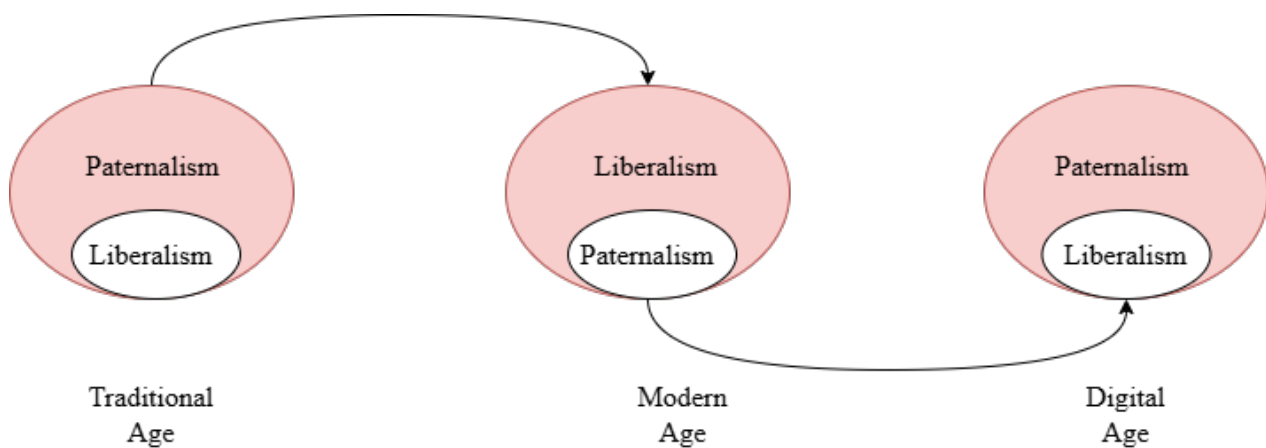


Figure 2. The shifts between paternalism and liberalism from the traditional era to the digital era.

1. Traditional Era: Authority was centralized in patriarchal systems: fathers, kings, and clergy. Paternalism was direct, coercive, and male-dominated, justified through divine or legal mandate (e.g., Roman law's *pater familias*).
2. Modern Era: Industrialization and liberal thought fostered political modernity. Paternalism retreated as liberalism emphasized freedom and individual rights. States intervened only when necessary to maintain order [76].
3. Digital Era: AI reintroduces paternalism, but in covert and decentralized ways. This "liberating paternalism" is non-coercive and often invisible, acting through nudges, manipulation, and behavioral prediction. The individual feels free yet increasingly guided by machines.

So, the relationship between paternalism and autonomy has never followed a simple, linear historical progression. Despite liberalism's dominance in Western political thought, paternalistic practices have persisted, particularly in the realm of social policy. Lawes (2017) [37] explores how nineteenth-century Britain, even amid liberal and individualistic values, witnessed a revival of paternalistic ideas in welfare and public policy. Rather than being a remnant of feudalism, this revival represented a reconfiguration of familial and societal responsibility under modern governance.

Paternalism in this context was no longer strictly coercive or hierarchical; instead, it became institutionalized in the form of public services, legal constraints, and state-sponsored programs aimed at improving citizens' well-being [77]. This evolution laid the groundwork for today's more complex forms of paternal intervention. Therefore, AI does not negate liberalism, but it does necessitate a re-evaluation of it, especially as we approach the development of this technology, such as artificial general intelligence [52].

The United Kingdom offers a compelling case of how AI is institutionalizing new forms of paternalism. Scholars Perry Keller and Archie Drake (2021) [9] argue that the integration of AI into government operations may raise concerns about paternalistic tendencies in British governance. Officials prefer working with private firms over non-profits to meet economic goals, giving tech companies a dominant role in shaping public life.

While legislation nominally regulates AI, the regulatory power lies with a few central agencies, creating an exclusive and often opaque governance structure. This contradicts the principles of liberal democracy, which emphasize public participation, transparency, and judicial oversight. The current approach has prompted debates about democratic oversight.

Other researchers argue that AI governance in public institutions may develop paternalistic characteristics when decision-making power becomes concentrated in technical systems and expert institutions rather than democratic participation [33,78,79].

We argue that we are entering a new chapter in the liberalism-paternalism debate. As AI systems continue to mediate more aspects of life from finance and education to dating and health, we must confront an essential question: How much autonomy are we willing to trade for comfort? And who should decide the limits of that trade?

Conversely, there are many studies that argue the opposite of this paper's argumentation. For example, Ryan Calo [80] argued that concerns about "algorithmic governance" are sometimes exaggerated, pointing out that individuals retain a great deal of control in digital environments and can often ignore or override algorithmic recommendations.

Krpan and Urbaník [32] also believe that artificial intelligence technologies have the potential to integrate various behavioral interventions in a non-parental way. More precisely, individuals will decide for themselves which behaviors they wish to change and will choose the technologies they wish to use for this purpose. In this sense, behavioral science will shift from liberal paternalism to liberalism, where individuals freely choose how to change.

While these criticisms highlight important limitations of claims regarding algorithmic control, the repeated reliance on algorithmic recommendations can gradually reshape decision-making and attention patterns. Over time, individuals may become increasingly dependent on algorithmic filtering and guidance, particularly in information-complex environments.

Therefore, the concept of AI paternalism does not assume that human autonomy has disappeared or that individuals are entirely controlled by algorithmic systems. Rather, it suggests that algorithmic environments may indirectly reshape the conditions under which autonomy is exercised.

Even Krpan and Urbaník [32] give a counter-argument in their paper that supports our argument, where they claim that although technological tools may outwardly support liberalism because they will endorse freedom of choice as well as subjectivity or multi-values, they can be used covertly to achieve different goals that do not necessarily align with the individual, but rather with the interests of those who control the technology.

It is worth noting that not all uses of AI systems necessarily produce parental dynamics. In many cases, individuals adapt digital tools themselves to support their personal goals. This depends on the individual's understanding of how AI works. For instance, wearable devices can support users' autonomy by providing data that helps them make their own health decisions rather than merely imposing behavioral nudges [81–83]. Also, some students use ChatGPT with the intention to verify information and use AI responsibly, which strongly predicts their adoption of it for academic tasks [84].

And while we emphasize the AI liberating paternalism, it still requires much development to become more intelligent and mature. The more machine intelligence advances, the more paternalistic it becomes.

Currently, AI's possession of consciousness remains unrealistic. However, if AI enters an era of awareness or if neural chips are implanted into the human brain, this may lead to a form of *tangible paternalism* felt directly by humans. At that point, people may begin to feel AI's control over them and the restriction of their decisions in favor of AI's logic. Many studies have already issued warnings about the implications of AI reaching this stage [85–88].

8. Conclusions

This paper examines the emerging phenomenon of AI paternalism, defined as the influence of algorithmic systems on human decision-making through automated recommendations, prompts, and guidance. Rather than claiming that AI negates human autonomy, the analysis suggests that algorithmic systems are increasingly shaping the information and decision-making environments in which individuals exercise autonomy. AI paternalism highlights a growing tension between the liberal ideal of autonomous decision-making and the practical realities of algorithmically mediated environments. Whether this represents a transformation of liberalism or merely an adaptation of its principles remains an open question.

By situating AI paternalism within broader political philosophy discussions, particularly liberal theories of autonomy and relational conceptions of agency. However, this study has several limitations. The analysis is primarily conceptual and does not provide empirical evidence regarding how users interpret or experience algorithmic guidance. Future research could explore these questions through surveys, experiments, or qualitative studies examining how individuals perceive the impact of AI.

Understanding liberal politics within the context of AI paternalism will become increasingly important as algorithmic systems continue to shape decision-making across fields such as education, healthcare, and governance. Rather than signaling the end of liberalism, these developments may necessitate a re-examination of how to preserve autonomy and understand liberalism versus relinquishing the burden of decision-making to digital mediation.

Author Contributions

Conceptualization, M.A.-A.; resources, M.A.-A. and L.K.N.; writing—original draft, M.A.-A. and L.K.N.; writing—review and editing, L.K.N.; visualization, M.A.-A.; supervision, M.A.-A.; project administration, M.A.-A. and L.K.N. Both authors have read and agreed to the published version of the manuscript.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

This is a review article, and no new data was generated or analyzed in this study. All data supporting this review are derived from previously published sources such as Google Scholar and ScienceDirction, which have been appropriately cited in the manuscript.

Acknowledgments

We would like to thank the Center for Policy Research (CPR), Unicersiti Sains Malaysia (USM) and Universiti Muhammadiyah Malaysia, Business and Management School that provided support for this study.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

AI Use Statement

Artificial intelligence tools were used in the preparation of this manuscript solely for language editing, proof-reading, and improving clarity. They were also used to assist in locating general information. All ideas, interpretations, and final decisions remain the responsibility of the authors.

References

1. Beck, B. Paternalism and Liberty/Autonomy as Dialectically Related Concepts. *Z. Ethik Moralphil.* **2023**, *6*, 223–237. [CrossRef]
2. Vaassen, B. AI, Opacity, and Personal Autonomy. *Philos. Technol.* **2022**, *35*, 88. [CrossRef]
3. Ghosh, P.; Liedo, B.; Riek, L.D. The Problem of Ableist Paternalism in Assistive Robotics. In Proceedings of the HRI 2024 Workshop on Assistive Applications, Accessibility, and Disability Ethics (A3DE), Boulder, CO, USA, 15 March 2024.
4. Kühler, M. Exploring the phenomenon and ethical issues of AI paternalism in health apps. *Bioethics* **2022**, *36*, 194–200.
5. Lorenzini, G.; Arbelaez Ossa, L.; Shaw, D.M.; et al. Artificial intelligence and the doctor–patient relationship expanding the paradigm of shared decision making. *Bioethics* **2023**, *37*, 424–429.
6. Rochi, M. Technology paternalism and smart products: Review, synthesis, and research agenda. *Technol. Forecast. Soc. Change* **2023**, *192*, 122557. [CrossRef]
7. Sunstein, C.R. Behavioral Biases, Choice Engines, and Paternalistic AI. *SSRN Electron. J.* **2023**, 1–17. Available online: <https://www.ssrn.com/abstract=4539053>
8. Quigley, M. Libertarian paternalism, nudging, and public policy. In *The Routledge Handbook of the Philosophy of Paternalism*; Routledge: New York, NY, USA, 2018; pp. 223–235.
9. Keller, P.; Drake, A. Exclusivity and paternalism in the public governance of explainable AI. *Comput. Law Secur. Rev.* **2021**, *40*, 105490.
10. Grill, K.; Hanna, J. *The Routledge Handbook of the Philosophy of Paternalism*; Routledge: New York, NY, USA, 2018.
11. Khadilkar, P.; Jagtap, S. Can Design Be Non-paternalistic? Conceptualizing Paternalism in the Design Profession. *She Ji J. Des. Econ. Innov.* **2021**, *7*, 589–610.
12. Grill, K. Paternalism. In *Encyclopedia of Applied Ethics*; Elsevier: Amsterdam, The Netherlands, 2012; pp. 359–369.
13. Dworkin, G. Paternalism. *Monist* **1972**, *56*, 64–84.
14. Kant, I. *Groundwork of the Metaphysics of Morals*; Cambridge University Press: Cambridge, UK, 1785.
15. Formosa, P. Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds Mach.* **2021**, *31*, 595–616. [CrossRef]
16. Mill, J.S. *On Liberty*; Macmillan Education UK: London, UK, 1859.
17. Berlin, I. Two Concepts of Liberty. In *Reading Political Philosophy*; Routledge: London, UK, 2014; pp. 231–237.
18. Rawls, J. A theory of justice. In *Applied Ethics*; Routledge: London, UK, 2017; pp. 21–29.
19. Floridi, L.; Cowl, J. A Unified Framework of Five Principles for AI in Society. In *Machine Learning and the City*; Wiley: Hoboken, NJ, USA, 2022; pp. 535–545. [CrossRef]
20. Thaler, R.; Sunstein, C. *Nudge: Improving Decisions about Health, Wealth and Happiness*; Yale University Press: London, UK, 2008; p. 89.
21. Coeckelbergh, M. Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability. *Sci. Eng. Ethics* **2020**, *26*, 2051–2068.
22. Luther, J. Discovery in an age of artificial intelligence. *Learn. Publ.* **2016**, *29*, 75–76.
23. Huang, L.F. Artificial intelligence. In Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE 2010), Singapore, 26–28 February 2010; pp. 610–684.
24. Sharifmousavi, M.; Kayvanfar, V.; Baldacci, R. Distributed Artificial Intelligence Application in Agri-food Supply Chains 4.0. *Procedia Comput. Sci.* **2024**, *232*, 211–220. [CrossRef]
25. Ellingrud, K.; Sanghvi, S.; Dandona, G.S.; et al. *Generative AI and the Future of Work in America*; McKinsey

- Global Institute: New York, NY, USA, 2023.
26. Sunstein, C.R. *Choosing Not to Choose: Understanding the Value of Choice*; Oxford University Press: Oxford, UK, 2015.
 27. Klincewicz, M. Artificial intelligence as a means to moral enhancement. *Stud. Logic Gramm. Rhetor.* **2016**, *48*, 171–187.
 28. Calvo, R.A.; Peters, D.; Vold, K.; et al. Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry. In *Ethics of Digital Well-Being*; Springer: Cham, Switzerland, 2020.
 29. Laitinen, A.; Sahlgren, O. AI Systems and Respect for Human Autonomy. *Front. Artif. Intell.* **2021**, *4*, 1–14.
 30. Hausman, D.M.; Welch, B. Debate: To Nudge or Not to Nudge. *J. Polit. Philos.* **2010**, *18*, 123–136.
 31. Susser, D.; Roessler, B.; Nissenbaum, H. Online manipulation: Hidden influences in a digital world. *Geo. L. Tech. Rev.* **2019**, *4*, 1–45.
 32. Krpan, D.; Urbaník, M. From libertarian paternalism to liberalism: Behavioural science and policy in an age of new technology. *Behav. Public Policy* **2024**, *8*, 300–326.
 33. Danaher, J. The Threat of Algocracy: Reality, Resistance and Accommodation. *Philos. Technol.* **2016**, *29*, 245–268.
 34. Hassan, S.A. How AI Can Be Used to Manipulate People. Available online: <https://www.psychologytoday.com/us/blog/freedom-of-mind/202304/how-ai-can-be-used-to-manipulate-people> (accessed on 5 May 2025).
 35. Foster, G.M. Bishop Cheshire and Black Participation in the Episcopal Church: The Limitations of Religious Paternalism. *North Carol. Off. Arch. Hist.* **1977**, *54*, 49–65. Available online: <http://www.jstor.org/stable/23529902>
 36. Abercrombie, N.; Hill, S.; Turner, B.S. *The Penguin Dictionary of Sociology*; Penguin Group: New York, NY, USA, 2001.
 37. Lawes, K. Paternalism and Politics. In *Crewe: Railway Town, Company and People 1840–1914*; Routledge: London, UK, 2017; pp. 186–208.
 38. MacGregor, S. From paternalism to partnership. *BMJ* **1998**, *317*, 221.
 39. Charette, S.L.; Garcia, M.B.; Reuben, D.B. Goal-Oriented Care. In *Psychology and Geriatrics*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 1–19.
 40. Esteva, A.; Robicquet, A.; Ramsundar, B.; et al. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29.
 41. Sunstein, C.R. Brave New World? Human Welfare and Paternalistic AI. *Theor. Inq. Law* **2024**, 1–30. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4908836
 42. Božić, V.; Poola, I. Chat GPT and education. *Education* **2023**, *1*, 1–8.
 43. Sharma, S.; Yadav, R. Chat GPT-A Technological Remedy or Challenge for Education System. *Glob. J. Enterp. Inf. Syst.* **2022**, *14*, 46–51.
 44. Zhu, J.J.; Jiang, J.; Yang, M.; et al. ChatGPT and Environmental Research. *Environ. Sci. Technol.* **2023**, *57*, 17667–17670. [CrossRef]
 45. Yu, H. Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. *Front. Psychol.* **2023**, *14*, 1181712. Available online: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1181712/full>
 46. New, B. Paternalism and Public Policy. *Econ. Philos.* **1999**, *15*, 63–83.
 47. Chin, J.J. Doctor-patient relationship: From medical paternalism to enhanced autonomy. *Singap. Med. J.* **2002**, *43*, 152–155.
 48. Burrows, P. Analyzing legal paternalism. *Int. Rev. Law Econ.* **1995**, *15*, 489–508.
 49. Arneson, R.J. Joel Feinberg and The Justification of Hard Paternalism. *Leg. Theory* **2005**, *11*, 259–284.
 50. Takala, T. Right to Know and Right Not to Know. In *Encyclopedia of Applied Ethics*; Elsevier: Amsterdam, The Netherlands, 2012; pp. 829–835.
 51. Clarke, S. A definition of paternalism. *Crit. Rev. Int. Soc. Polit. Philos.* **2002**, *5*, 81–91.
 52. Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*; Oxford University Press: Oxford, UK, 2014; pp. 1–7.
 53. Sheng, A.; Wang, F. Falling in love with machine: Emotive potentials between human and robots in science fiction and reality. *Neohelicon* **2022**, *49*, 563–577.
 54. Rochi, M. Essays on Technology Paternalism: Paternalistic Effects of Smart Technologies on User Behavior. PhD Thesis, Otto-Friedrich-Universität Bamberg, Bamberg, Germany, 2024. Available online: <https://fis.uni-bamberg.de/handle/uniba/98474>
 55. Al-Abaydh, M.; Bustami, M.R. Paternalism and Liberalism: A Comparative Analysis of Political Modernity in

- Western and Non-Western Societies. *Comp. Sociol.* **2025**, *24*, 351–383.
56. Hill-Yardin, E.L.; Hutchinson, M.R.; Laycock, R.; et al. A Chat(GPT) about the future of scientific publishing. *Brain Behav. Immun.* **2023**, *110*, 152–154.
 57. Fossa, F. Artificial moral agents: Moral mentors or sensible tools? *Ethics Inf. Technol.* **2018**, *20*, 115–126.
 58. Strickland, E. IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care. *IEEE Spectr.* **2019**, *56*, 24–31.
 59. Hung, L.; Liu, C.; Woldum, E.; et al. The benefits of and barriers to using a social robot PARO in care settings: A scoping review. *BMC Geriatr.* **2019**, *19*, 232.
 60. Sunstein, C.R. Choice engines and paternalistic AI. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 888. [CrossRef]
 61. Fu, R.; Huang, Y.; Singh, P.V. AI and Algorithmic Bias: Source, Detection, Mitigation and Implications. *INFORMS Tut. Oper. Res.* **2020**, 39–63.
 62. Goldberg, L.R. Thinking, Fast and Slow, by D. Kahneman. *Quant. Financ.* **2013**, *13*, 177–179.
 63. Banaji, M.R.; Greenwald, A.G. *Blindspot: Hidden Biases of Good People*; Bantam: New York, NY, USA, 2016.
 64. Mittelstadt, B.D.; Allo, P.; Taddeo, M.; et al. The ethics of algorithms: Mapping the debate. *Big Data Soc.* **2016**, *3*, 2053951716679679.
 65. Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*; St. Martin's Press: New York, NY, USA, 2018.
 66. Al-Abyadh, M.; Bustami, M.R. The Revival of Paternalism in Shia of Iraq Post-2003: Al-Sistani and the Formation of Democracy. *Russ. Law J.* **2023**, *11*, 2654–2668.
 67. Mackenzie, C.; Stoljar, N. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*; Oxford University Press: Oxford, UK, 2000.
 68. Friedman, M. *Autonomy, Gender, Politics*; Oxford University Press: Oxford, UK, 2003.
 69. Sen, A. *Development as Freedom*; Oxford University Press: Oxford, UK, 1999.
 70. Nussbaum, M.C. Creating capabilities: The human development approach and its implementation. *Hypatia* **2009**, *24*, 211–215.
 71. Chen, X.; Ji, Y.; Liu, R.; et al. Mitochondrial dysfunction: roles in skeletal muscle atrophy. *J. Transl. Med.* **2023**, *21*, 503. [CrossRef]
 72. Fanzani, A.; Conraads, V.M.; Penna, F.; et al. Molecular and cellular mechanisms of skeletal muscle atrophy: An update. *J. Cachexia Sarcopenia Muscle* **2012**, *3*, 163–179.
 73. McGuire, L.P.; Johnson, E.M.; Frick, W.F.; et al. Temperature alone is insufficient to understand hibernation energetics. *J. Exp. Biol.* **2021**, *224*, jeb239772. [CrossRef]
 74. Robbins, C.T.; Lopez-Alfaro, C.; Rode, K.D.; et al. Hibernation and seasonal fasting in bears: The energetic costs and consequences for polar bears. *J. Mammal.* **2012**, *93*, 1493–1503.
 75. Al-Abyadh, M.; Kalista Noor, L.; Fawzi, A. Between Cognitive Augmentation and Atrophy: AI's Role in Reshaping Human Intelligence. *Interciencia* **2025**, *408*, 179–202.
 76. Parmaksız, P.M.Y. Paternalism, Modernization, and the Gender Regime in Turkey. *Aspasia* **2016**, *10*, 40–62.
 77. Mathew, S.; Taylor, G. Power distance in India: Paternalism, religion and caste: Some issues surrounding the implementation of lean production techniques. *Cross Cult. Strateg. Manag.* **2019**, *26*, 2–23.
 78. Crider, C. Two paths for health AI governance: Paternalism or democracy. *Futur. Healthc. J.* **2024**, *11*, 100180.
 79. Taeihagh, A. Governance of artificial intelligence. *Policy Soc.* **2021**, *40*, 137–157.
 80. Calo, R. Artificial Intelligence Policy: A Primer and Roadmap. *SSRN Electron. J.* **2017**, 1–28.
 81. Jin, D.; Halvari, H.; Maehle, N.; et al. Self-tracking behaviour in physical activity: A systematic review of drivers and outcomes of fitness tracking. *Behav. Inf. Technol.* **2022**, *41*, 242–261.
 82. Asimakopoulos, S.; Asimakopoulos, G.; Spillers, F. Motivation and User Engagement in Fitness Tracking: Heuristics for Mobile Healthcare Wearables. *Informatics* **2017**, *4*, 5.
 83. Owens, J.; Cribb, A. 'My Fitbit Thinks I Can Do Better!' Do Health Promoting Wearable Technologies Support Personal Autonomy? *Philos. Technol.* **2019**, *32*, 23–38.
 84. Acosta-Enriquez, B.G.; Arbulú Ballesteros, M.A.; Huamaní Jordan, O.; et al. Analysis of college students' attitudes toward the use of ChatGPT in their academic activities: effect of intent to use, verification of information and responsible use. *BMC Psychol.* **2024**, *12*, 255. [CrossRef]
 85. Butlin, P.; Long, R.; Elmoznino, E.; et al. Consciousness in Artificial Intelligence: Insights from the Science of Consciousness. *arXiv Preprint* **2023**, *arXiv:2308.08708*.
 86. Smith, D.H.; Schillaci, G. Why Build a Robot With Artificial Consciousness? How to Begin? A Cross-Disciplinary Dialogue on the Design and Implementation of a Synthetic Model of Consciousness. *Front. Psychol.* **2021**, *12*, 530560.

87. Anwar, N.A.; Badea, C. Can a Machine Be Conscious? Towards Universal Criteria for Machine Consciousness. *arXiv Preprint* **2024**, *arXiv:2404.15369*.
88. Bojic, L.; Stojković, I.; Jolić Marjanović, Z. Signs of Consciousness in AI: Can GPT-3 Tell How Smart It Really Is? *SSRN Electron. J.* **2023**, 1–37. Available online: <https://www.ssrn.com/abstract=4399438>



Copyright © 2025 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.