ORGDYN 53 (2024) 101041

Contents lists available at ScienceDirect

Organizational Dynamics

journal homepage: www.elsevier.com/locate/orgdyn

Ethical compass: The need for Corporate Digital Responsibility in the use of Artificial Intelligence in financial services

Zsófia Tóth *, Markus Blut

Durham University Business School, University of Durham, United Kingdom

ARTICLE INFO

Keywords:
Artificial Intelligence
Business ethics
Corporate Digital Responsibility
Accountability
Moral intensity
Locus of morality

ABSTRACT

Service research and business ethics literature intersect concerning the question of artificial intelligence (AI) service robot accountability. In financial services, there is a broad spectrum of potential ethical issues, from data usage to customer vulnerabilities. This article scrutinizes the impact of morality and where accountability resides in the use of AI service robots in financial services. To address this challenge, we discuss the role of Corporate Digital Responsibility (CDR) for firms and illustrate how to implement a conceptual framework on the ethical implications of AI service robot applications, drawing on normative ethical theory. The framework elaborates on how the locus of morality (from human to AI agency) and moral intensity combine within context-specific AI service robot applications, and how this might influence associated accountability. We provide examples of AI robots' use for different purposes, differentiating between four 'accountability clusters': (1) professional norms, (2) business responsibility, (3) inter-institutional normativity, and (4) supra-territorial regulations cluster. We also discuss the CDR implications in different clusters. Ethical implications of using AI service robots and associated accountability challenges are relevant for a network of actors—from customers and designers to firms and the government. Implementation of the framework incorporates a range of internal and external stakeholders that firms need to consider. We also provide a CDR roadmap to incorporate a time perspective and to inform implementation efforts.

Introduction

Artificial intelligence (AI) is increasingly being integrated into, and reshaping, the financial services landscape. The dynamic development of AI has also accentuated the pressing need for corporate digital responsibility (CDR), with special regards to the questions of the identification of accountability and human agency. AI ethics forms part of CDR because (un)ethical AI use directly influences various stakeholders, including consumers, employees, organizational partners, and society at large. Even though AI already plays a significant role in the financial services industry, its future influence is expected to grow further. For example, AI systems can analyze vast amounts of data to assess and manage risks more effectively, and are therefore used for risk assessment and management. Financial institutions use AI to identify fraud, assess creditworthiness, and improve lending decisions. AI is also used extensively in high-frequency trading and algorithmic trading strategies to analyze market data and execute trades at speeds and frequencies that are impossible for humans to attain. Moreover, AI can improve the provision of personalized financial advice based on an individual's financial situation, goals, and risk tolerance. However, although these opportunities enhance the quality of service provision, the use of AI in financial services also presents various data and privacy risks, as well as other undesirable side-effects (McKinsey, 2021).

The importance of ethical AI use has been stressed by various business organizations, such as the Financial Conduct Authority (FCA), which regulates the conduct of nearly 50,000 UK businesses. During the Artificial Intelligence Public-Private Forum launch, the FCA (2022) emphasized their desire for consumers to benefit from digital innovation and competition, including data-based and algorithmic innovation, and highlighted the fact that consumers must be confident that they are obtaining fair access, price, and quality, and that firms are acting in their best interest. The FCA encourages transparent, fair, and secure decision-making in situations where financial services firms make decisions using data-based or algorithmic methods. The FCA also advocates for ethical data use, particularly regarding vulnerable consumer groups.

This article scrutinizes the impact of morality in the use of AI in financial services, as well as the accountability of those implementing

* Corresponding author.
E-mail addresses: zsofia.toth@durham.ac.uk (Z. Tóth), markus.blut@durham.ac.uk (M. Blut).

https://doi.org/10.1016/j.orgdyn.2024.101041
Received 14 February 2024; Accepted 21 February 2024
Available online 29 February 2024

AI. Next, we discuss the concept of CDR and how it defines the context for the present article. We also explain the conceptual framework that was developed based on the ethical implications of AI applications. This framework draws on normative ethical theory and empowers firms to maximize CDR. The AI accountability framework, as applied to the financial services sector, provides managers with a structured approach to navigate the complexities of integrating AI into their strategies and operations. We illustrate how this framework can be implemented in financial services. We provide managers with a guiding CDR roadmap containing specific measures to implement and steps to take when considering ethical AI use, and we illustrate this with several use cases. The roadmap considers different accountability clusters, the time perspective, and key principles, strategies, and actions.

How is CDR reshaping management practices?

CDR is a management concept that help firms comprehensively address the ethical challenges related to digital technology use. CDR helps managers to ensure the responsible use of new technologies to the benefit of organizations, society, and the environment, while mitigating negative consequences. Rooted in computer and business ethics, CDR comprises a set of shared values and norms that guide a firm's operations with respect to the creation and operation of digital technology and data. The concept considers specific ethical dilemmas and firm responsibilities related to technology use. CDR represents a firm's commitment to responsibly use technology (Wirtz et al., 2023).

A firm's CDR performance can be assessed in various ways. For instance, the German Association for the Digital Economy BVDW (2023) created a CDR award to recognize firms' achievements regarding their CDR activities. The BVDW identifies three general CDR domains, including strategy and governance, change and transformation, and values and principles. Additionally, several specific subdomains are considered, including AI use, digital well-being, environment and resources, communication, responsibility for data, privacy and security, responsible innovation, future of work, and digital empowerment and inclusion. The BVDW provides best-practice examples of top-performing firms, as well as useful criteria and guidelines for CDR implementation. In terms of CDR strategy and governance, the science and technology company Merck KGaA has been recognized for its international digital ethics advisory panel, which helped to steer its increasingly digitized business, while the car manufacturer BMW has received recognition for its development of AI guidelines.

The clarity of organizational responsibilities in managing and leveraging digital technologies, data privacy and cybersecurity considerations, as well as the socially responsible use of digital technologies, are high priorities for the maximization of CDR. AI ethics largely falls under the umbrella of CDR, since it focuses on the ethical deployment of AI systems in various settings. An important overlapping area is accountability, which the present article aims to disentangle for the case of AI's use in financial services. Further shared themes are data governance that informs algorithmic decision-making, transparency, explainability, social impact, and a human-centric approach. The latter theme includes the prioritization of human safety, dignity, and well-being. It also includes, at the very least, Asimov's law, which states that AI may not injure any human being or, through inaction, allow any human being to suffer harm.

How can CDR guide ethical AI implementation in financial services?

The financial services sector has always adapted in response to new digital technologies. For example, banks have adopted various technologies to improve customer experience and back-office processes. This started with the introduction of ATMs in the 1960s, and continued through card-based payments in the 1970s, online banking in the 2000s, and mobile banking in the 2010s. Now, AI-powered technologies are expected to transform the industry. AI can support banking process automation and enhance human decision-making in terms of both speed and accuracy. AI can increase revenue generation for banks by delivering highly personalized services to customers while simultaneously reducing costs through improving automation, minimizing error rates, and optimizing resource allocation. Furthermore, AI can leverage the vast amount of data available to unlock untapped opportunities and generate valuable insights, thereby enabling the discovery of new avenues for growth. The annual worldwide value of AI for banking is estimated to be as high as $1 trillion, and is expected to have the strongest impact on marketing and sales, followed by risk management, human resources, and finance and IT. However, AI usage in financial services has the potential to violate certain ethical principles.

CDR is of particular importance in financial services given the nature of the service and its impact on people's lives. AI algorithms can have adverse consequences and undercut individuals' rights and dignity. As an example, a car insurance algorithm might be designed to consider a consumer's credit score as a more significant factor than their history of drunk driving, begging the question of whether appropriate factors are being used for such evaluations. Such ethical issues can be avoided when technical teams and software developers are aware of the importance of ethics and receive training on the ethical implications of algorithms. AI algorithms therefore have the potential to be less biased and more neutral than human decision-making, which also suffers from biases. Various approaches can be applied to prevent algorithmic biases. For example, bias can be removed from the data prior to model construction (i.e., data preprocessing), or improved goals can be set for models through the use of fairness metrics (Townson, 2020). Besides, regular auditing, monitoring, and stakeholder engagement provide additional assurances.

According to the BVDW, CDR implementation has several advantages, including the creation of trust, transparency, innovation, risk minimization, and value orientation, as well as sustainable resource usage. For example, when financial service firms implement CDR, they create trust among customers and employees in their utilization of the firm's products and services, as well as in the company itself. Approaching ethical issues related to these digital technologies also creates transparency regarding their social and environmental impact at all links in the value chain.

How to assess accountability?

Accountability refers to individuals and organizations taking responsibility for their actions and decisions. This includes being accepting of, and open to questions about, relevant ethical implications, aiming to adhere to a set of ethical standards. Various accountability clusters can be created; these may be based on insights from the entities involved in relevant actions and decision-making (e.g., AI developers or various organizations and governments), as well as the locus of morality (the extent to which human morality is visible) and moral intensity concepts (the magnitude of impact that certain decisions/actions have on others). While the locus of morality describes the locality of moral decision-making, moral intensity describes the extent to which a single action can impact multiple victims or beneficiaries. The novelty of the AI accountability framework we apply to the financial services industry lies in the fact that it ascribes moral responsibility between humans and organizations in relation to AI usage, and combines it with moral intensity. Financial services professionals can use the three dimensions to compile different ethical issues and assess them in terms of responsibility (human versus AI) and moral intensity (low versus high impact), as well as accountability dispersal (designers/AI unit; organization/users; industry norms/regulatory bodies; intergovernmental regimes).

Locus of morality

AI in and of itself does not possess awareness and intrinsic morality. Thus, humans and organizations are always responsible for imbuing AI

with ethical guidelines, or failing to do so. However, the locus of morality exists as a spectrum, from clearly attributable human morality at one end to a blurred locus of morality between AI and humans at the other. Likewise, the ethical implications of certain AI-supported financial decisions are clearly attributable to humans in some cases, and less so in others. Human oversight is crucial for keeping the locus of morality as straightforward as possible. This ensures a clear connection to human decision-makers, where the locus of morality ultimately resides. The locus of morality therefore describes the individuals whose ethical standards inform decision-making.

Morality intensity

Moral intensity refers to the number of humans potentially affected by AI-supported decisions and other activities, human vulnerability, and the severity of effects on communities and ecosystems. For example, individuals' financial stability may be harmed, and public trust in financial institutions can be eroded. This might especially be the case in situations of high moral intensity, for instance when many people are unable to access their financial resources due to a cybersecurity issue.

Accountability dispersal

Accountability dispersal pertains to moral intensity and blurred human agency (in cases where the locus of morality is not clearly attributable to humans), wherein more stakeholders are required as accountability dispersal increases. Stakeholders may include financial institutions, AI developers, regulators, industry associations, and representatives of local communities. Some challenges can be managed within organizations, while others require inter-governmental collaboration. Firms must be cognizant of higher accountability dispersal across firms as the number of stakeholders increases.

Accountability clusters

In addition to the three major aspects of the AI accountability framework, we differentiate between four 'accountability clusters': (1) professional norms, (2) business responsibility, (3) inter-institutional normativity, and (4) supra-territorial regulations. Most ethical issues can be handled by developers and supporting managers within the professional norms cluster—for instance, by ensuring that certain features are integrated into AI use. The business responsibility cluster requires strategic-level decision-making within an organization, as well as liaisons between various departments. For example, to avoid small-scale market manipulations, brokers, developers, cybersecurity experts, and other professionals should collaborate; however, strategic-level support from leadership is also required to encourage regular monitoring. Inter-institutional normativity typically requires collaboration between different institutions for ethical AI use, and may at times require competitors to share knowledge and best practices. For instance, the deployment of AI for investment decisions may require a joint effort from different market actors to benefit customers and the sector in general. The supra-territorial regulations cluster of accountability may require international and inter-sectoral collaborations, in addition to intra- and inter-organizational effort. The creation of best practices to support diversity and inclusion in AI-enhanced decision-making in the financial services sector and beyond specifically resides in this accountability cluster.

We combine these elements in the AI accountability framework shown in Fig. 1., an application from Tóth et al. (2022). We use the financial services context to illustrate the applications to the financial services industry and provide managers with guidance. This figure allows the derivation of managerial implications from different perspectives with regard to normative ethics—specifically illegality, immorality, permissibility, and supererogation—across the four accountability clusters. We discuss implications and use cases using these two perspectives (i.e., normative ethics and accountability clusters) below. Subsequently, we discuss the locus of morality and the moral intensity of AI's use in financial services, and examine relevant use cases.

Pitfalls and opportunities of ethical AI use for financial services organizations

Ethical theories demonstrate how our perceptions of moral intensity affect individual decision-making. For instance, a bank manager following certain ethical standards might refuse to facilitate money laundering activities that would prevent a children's hospital from receiving its annual funding from a major funding body. In contrast, AI is

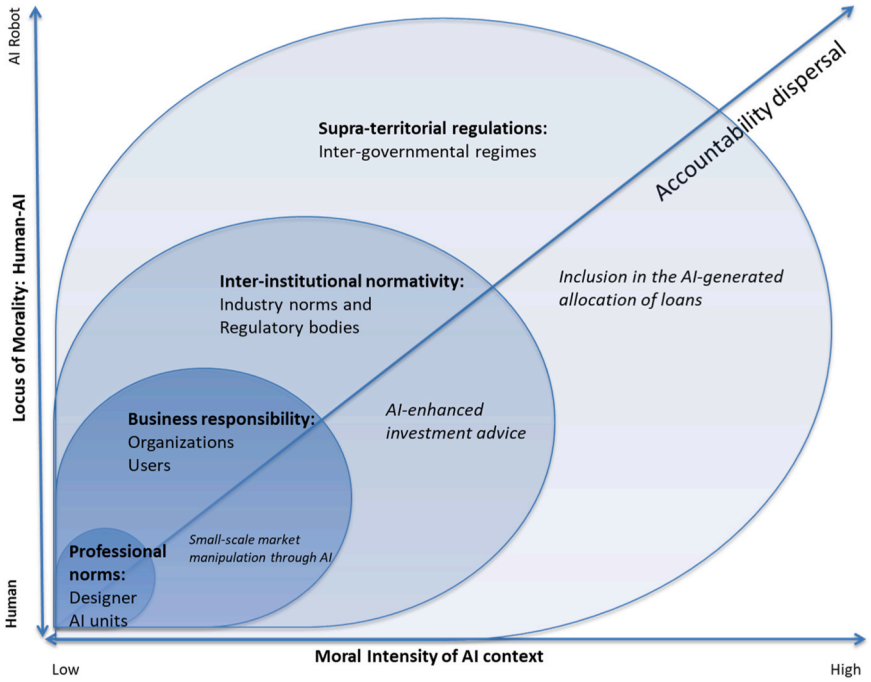


Fig. 1. Application of a framework for AI accountability (following Tóth et al., 2022).

not influenced by such perceptions unless it is programmed to account for moral intensity. On the other hand, AI would most likely avoid bribery and blackmail if programmed accordingly due to its lack of emotions, including temptation and fear. Thus, both the absence of ethical thought and the absence of greed are intriguing starting points for the 'ethical training' of AI.

Following the AI accountability framework's introduction, it is useful to further examine different use cases specific to the financial services sector, and thereafter in more detail from a normative perspective with action points. Normative business ethical categorization distinguishes between (1) illegal and immoral, (2) permissible, and (3) supererogatory cases. All of these categories are relevant for AI usage in finance. These three categories are relevant at different levels of moral intensity and AI/human agency. Relevant use cases, as well as managerial attention and where it is most needed, will be discussed next, first from a bird's eye view, and then focusing specifically on illegal/immoral and permissible dealings, as well as supererogation. Table 1 outlines four exemplary and general areas that are relevant for AI's ethical use in the financial services sector: (1) data privacy and security, (2) non-discrimination and fairness, (3) transparency and explainability, and (4) regulatory compliance. Issues associated with low moral intensity should not be easily disregarded. In fact, they may evolve into cases of high moral intensity if not managed properly and on time.

Scenarios illustrating the illegal and immoral use of AI in financial services. Potential illegal and immoral activities must be considered when creating ethical guidelines for AI usage in financial services, including how such activities can be prevented, when they take place, how they can be stopped, and how their impact can be minimized. While illegal activities are forbidden by law, immoral activities may not always be prohibited but typically go against widely accepted moral standards in a specific context. Table 2 provides an overview of exemplary AI use scenarios. For example, systematic discrimination is immoral and illegal in most countries because it violates principles of fairness and equality. It can also potentially undermine trust in financial systems, thereby giving rise to risk and far-reaching socio-economic implications that necessitate the creation of regulations to prevent discriminatory practices in AI-driven financial decision-making. An example of this is the evaluation of mortgage applications. Companies should aim to minimize such practices to maintain a high CDR.

Scenarios illustrating the permissible use of AI in financial services. An intrinsically morally permissible action is something that is allowable and is not immoral or illegal. Permissible actions meet ethical requirements, though they may not exceed these expectations. Below, we demonstrate ethically permissible cases in financial services. AI-powered customer service and AI's ability to remain unaffected by abuse are mentioned above as ethically permissible actions, and special care should be taken to note what is considered ethically permissible AI in this context. If customers abuse AI, then it will not incur emotional and mental health-related damages on the provider's side, since AI does not have emotions. However, such cases must still be handled with care. Customers abusing AI virtual assistants without any consequences may encourage further abusive behaviors, including in human-to-human interactions, that should not be tolerated. Moreover, algorithmic trading can potentially have a high impact on market conditions, but is neither illegal nor immoral in most contexts. Algorithmic trading therefore has high accountability dispersal: its regulation requires the collaboration of several organizations, potentially at an international level. An example of the permissible use of AI in financial services is the application of AI-enhanced investment advice tools. While these require less human employment and thus a smaller workforce, they allow greater efficiency, scalability, and cost savings, and improved accessibility to a broader range of investors. Table 3 contrasts various use scenarios.

Scenarios illustrating the supererogatory use of AI in financial services. The term 'supererogatory' refers to an act that is good but not morally required. It is often colloquially known as 'going the extra mile' in terms of ethical conduct; in other words, going beyond what is expected or required. There are certain AI use cases in the financial services sector through which supererogatory outcomes can be achieved. For example, AI can support sustainability-focused investment decisions. This has high (positive) long-term moral intensity. It goes beyond what is expected of firms by enabling more comprehensive data-driven approaches with increased transparency to inform investment decision-making. Utilizing AI for such purposes maximizes a firm's CDR. Table 4 shows this and other exemplary use scenarios.

Accountability clusters and a CDR roadmap

There are several implications of AI usage in financial services organizations. Specific forms of accountability and responsibility should ideally be established in AI systems. These can be achieved through measures in different accountability clusters, as discussed below.

Professional norms measures

Professional norms measures should focus on transparent and responsible practices, even for mundane tasks. Managers should ensure that ethical considerations are ingrained in the design and developmental processes enhanced by AI technologies. Complying with equality and privacy should be present 'by design'. AI ethics guidelines should be developed early on, along with regular and scheduled ethical reviews. Both internal teams and certain key stakeholders (e.g., suppliers) should be updated on guidelines that require planning for training and for raising awareness. Customer experience, business efficiency, and innovativeness can be increased through ethical AI usage. Customers generally appreciate decreased response times and detailed information, though some may have reservations about communicating with a non-human entity. The absence of human support may be considered permissible if other customer benefits outweigh this shortcoming. These

Table 1
Exemplary areas of ethical AI use in the financial services sector.

Issue	Managerial attention	Locus of morality	Moral intensity	Accountability dispersal
1. Data privacy and security	Protection and privacy of customer data should be prioritized. Each financial institution should have guidelines for data management and storage, ensuring that data is only used for legitimate purposes.	Human agency has limited transparency	High	High
2. Non-discrimination and fairness	AI-enhanced financial decisions, such as loan allocation, should be free from bias. Therefore, no discrimination should occur based on protected attributes such as race and gender.	Human agency has some / limited transparency	High	High
3. Transparency and explainability	Customers should have the option to understand how financial decisions relevant to them are made in a way that is easy to understand for someone who does not possess professional financial and accountancy skills.	Human agency has some transparency	Medium	Medium/High
4. Regulatory compliance	There are financial regulations that AI systems must adhere to, such as anti-money laundering regulations. While there may be some country-level and regional specificities, regulatory compliance should be a priority in the development and use of AI systems.	Human agency has higher transparency	Low/ Medium	Low

Table 2
Exemplary illegal and immoral use scenarios of AI in financial services.

Issue	Managerial attention	Locus of morality	Moral intensity	Accountability dispersal
1. Systematic discrimination	Unless trained otherwise, the characteristic learning approach of AI is regressive, drawing on predictions based on data generated from the past. Thus, if the data used for AI-supported decisions shows discriminatory practices, AI will not 'think' that decisions should be made differently.	Human agency has very limited transparency	High	High
2. Major cyber-attacks and large-scale identity theft	Social engineering can become increasingly sophisticated with support from AI. Therefore, financial institutions should invest in monitoring the new tactics to which their customers and stakeholders are exposed.	Human agency has limited transparency	High/ Medium	High/Medium
3. Money laundering	AI algorithms could be used to obscure illicit financial transactions, making it more challenging to detect money laundering activities. Illicit funds may be more efficiently disguised with the use of AI, which can make certain transactions appear legitimate by circumventing the pattern detection of transactions.	Human agency has some transparency	Medium	Medium
4. Insider trading and market manipulation	AI may utilize non-public information as well as public information to help with the more accurate prediction of stock market trends. The use of non-disclosed private information for trading purposes is illegal.	Human agency has some transparency	Medium	Medium
5. Small-scale market manipulation	AI algorithms could be employed to engage in small-scale market manipulation tactics, such as the creation of artificial buying/selling pressure on specific stocks. While the impact may be limited to a small number of traders/investors, it is still illegal and immoral.	Human agency has higher transparency—due to the small-scale of the issue, it is easier to identify the humans behind the manipulation	Low	Low

Table 3
Exemplary permissible use scenarios of AI in financial services.

Issue	Managerial attention	Locus of morality	Moral intensity	Accountability dispersal
1. Algorithmic trading	Algorithmic trading relies on computer systems to buy shares automatically when preset market conditions are met. Enhanced by AI, these purchasing mechanisms work with higher efficiency and can have a significant impact on market conditions due to the potential for market volatility.	Human agency has very limited transparency	Medium/ High	High
2. Investment recommendations and portfolio management	Investment recommendations and portfolio management also benefit from AI support. Customers can receive broader and more fine-tuned offerings. However, with less human involvement, the number of required employees may be lower.	Human agency has some transparency	Medium	Medium
3. AI-powered customer service	AI-powered customer services may respond to inquiries more quickly and provide more accurate information. AI customer service agents do not get tired or annoyed. As they do not have emotions, they are unaffected by abuse. If human customer service is still available for non-standard inquiries alongside AI agents, customer experience can increase.	Human agency has higher transparency, especially in case of hybrid AI-human assistant set ups	Low	Low

issues are relevant to both professional norms and business responsibility measures. To address these issues systematically, firms must define the underlying CDR principles (e.g., code of ethics, ethical use of data, accessibility standards, legal compliance) before defining strategic initiatives (e.g., training/education, transparent communication, privacy by design, cross-functional collaboration) and engaging in ongoing activities (e.g., continuous improvements, digital responsibility reporting).

Business responsibility measures

Business responsibility measures require organization-level strategic decision-making for AI ethics. Good inter-departmental communication, as well as a supporting organizational culture, are vital for achieving this. Financial services firms should encourage collaboration between different professionals, such as brokers, developers, cybersecurity experts, and other professionals, as well as colleagues in senior leadership positions, to ensure AI ethics principles are applied across various activities, from data management to marketing communications at a strategic level. The success of financial organizations is highly dependent on their reputation and the trust that customers and stakeholders place in them. Ethical AI use can ensure that related reputation damage and trust issues are prevented. In contrast, inappropriate AI applications may have long-term consequences for a financial organization's reputation and for customer trust. Depending on a country's regulations, not all problematic AI use cases may be illegal, but immoral dealings can similarly damage reputation and trust. Ethical AI usage can also increase talent attraction and retention, since employees might prefer to be associated with organizations with high ethical conduct—ethical AI usage is increasingly becoming part of this practice. Business responsibility measures should account for this talent attraction/retention challenge. Ideally, financial institutions should appoint a CDR representative to oversee relevant processes. A CDR representative would also ensure human oversight, contribute to the work of ethical review boards, and organize intervention as needed. Related measures should be taken regarding key CDR principles (e.g., stakeholder engagement, sustainable digital practices, and ethical marketing practices), strategic initiatives (e.g., AI literacy programs, digital well-being initiatives, cybersecurity responsibility, and board oversight), and ongoing activities (e.g., social impact measurement, sustainable digital infrastructure, and compliance audits).

Table 4
Exemplary supererogatory use scenarios of AI in financial services.

Issue	Managerial attention	Locus of morality	Moral intensity	Accountability dispersal
1. AI employed for social impact and sustainable investments	AI algorithms can be used to identify investment opportunities that create a positive social and environmental impact alongside financial returns. This promotes the concept of impact investing, channeling capital toward ventures that address social and environmental challenges.	Human agency has limited transparency	High	High
2. Financial inclusion efforts and AI-supported large-scale financial education	In a similar vein, AI has the capacity to accelerate financial inclusion efforts by pinpointing areas for future improvement that humans may overlook. AI may also increase the intensity and sophistication of financial education, making it more accessible to groups of people who may not have otherwise received such support.	Human agency has some transparency that may be limited depending on the circumstances	High	High
3. AI-enhanced customer protection	With the help of AI, individuals' finance-related issues, such as falling victim to predatory lending, suffering from gaming addiction, or dealing with excessive debt, may be spotted more effectively, with customer consent permitting.	Human agency has some transparency, especially when human contacts are also available for customers	Medium	Medium/Low
4. Internal employee training on AI ethics for those who are not involved in using/developing AI	Training sessions about the ethical use of AI at a financial institution, even for those who do not use or develop AI, may still have a positive impact on organizational culture. AI tools can potentially be employed for the organization and content delivery in such training sessions.	Human agency has high transparency	Low	Low

Inter-institutional normativity measures

The progressive implementation of ethical AI principles, policies, and practices at the inter-institutional level is highly desirable. To achieve this, cross-institutional collaborations should be established between financial firms and regulatory bodies to develop shared AI principles that can later be fine-tuned at firm level, and informal/formal platforms for the sharing of best practices should be set up. For instance, shared ethical standards can be developed to provide AI-enhanced investment advice. Regulatory compliance at supply chain/network level, and the implementation of transparent and explainable AI practices to both internal and external stakeholders, should be prioritized. Financial and digital literacy used to be separate concepts; however, with the increasing use of AI, digital-related, and especially AI-related, literacy is becoming a necessary component of financial literacy. Financial institutions, in collaboration with non-profit organizations and policy-makers, should work to make AI-enhanced financial education more accessible. Relevant training should not stop at the boundaries of company premises but should reach schools and higher education institutions, including non-specialist student groups. Some of the benefits of ethical AI usage in financial organizations include more sophisticated risk mitigation/management and regulatory compliance, which are highly relevant for inter-institutional normativity. AI contributes to increased forecasting accuracy, the fine-tuning of risk modeling, and superior data segmentation for portfolio composition. Appropriate AI usage also complies with national and international laws and policies in a highly regulated sector—this enables companies to avoid illegal dealings. As mentioned above, firms should take measures regarding related CDR principles (e.g., supply chain responsibility and industry standards adoption), strategic initiatives (e.g., collaborative initiatives, cybersecurity collaboration, and common incident response protocols), and ongoing activities (e.g., certification programs and technology transfer initiatives).

Supra-territorial regulation measures

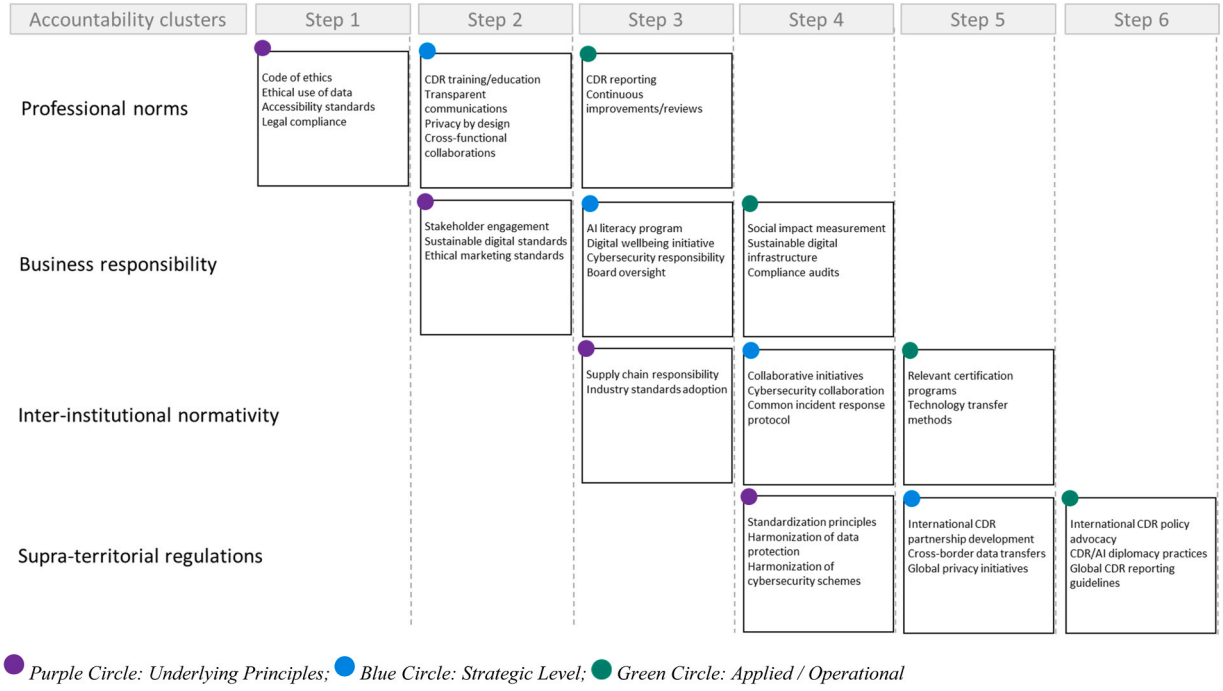
Inter-governmental regimes must be developed to oversee regular bias audits and bias-mitigation procedures in AI algorithms at regulatory level. Ensuring the collection of diverse and representative data and compliance with anti-discrimination laws in AI's applications to financial services should also be addressed at this level (e.g., at EU level; through the United Nations). Part of this effort is to ensure a fair representation of various countries and cultures to guide international AI-related governance decisions. The role of policy-makers is pivotal at supra-territorial level, and even more so than managerial roles due to the intensity of the regulatory effort. For example, avoiding discrimination requires high-level supra-territorial regulations where various factors should be considered. Besides paying attention to protected characteristics, such as race, sexual orientation, and gender, other characteristics, including age, may also be relevant from an anti-discriminatory viewpoint. As AI becomes ever more popular and more widely used among younger generations, their digital footprints accumulate more data for AI to learn from. Thus, younger generations may have an advantage in skills development associated with AI usage. However, the data they generate are more widely applied and therefore primarily reflect on the preferences of these younger generations. Through machine learning, financial institutions should ensure that their older customers – and other less represented groups – are not forgotten. By using fair machine learning approaches, market research and other techniques, institutions should ensure access, transparency, and explainability for customers from different walks of life and backgrounds. AI-enhanced financial education may become an element of CDR and an investment to inform future generations that are themselves increasingly more informed. The creation of social impact is vital in the supererogatory category of AI use in financial organizations. With the help of AI, efforts for financial inclusion and environmental sustainability can potentially be greatly increased, which is important at all levels, including with respect to supra-territorial regulations. Related CDR principles (e.g., standardization initiatives, harmonization of data protection, and harmonized cybersecurity frameworks), strategic initiatives (e.g., international partnerships, cross-border data transfers, and global privacy standards), and ongoing activities (e.g., international policy advocacy, digital/AI diplomacy, global digital responsibility, and reporting guidelines) should be defined.

These activities and their metrics are summarized in the CDR roadmap in Table 5, and classified according to the underlying accountability cluster and potential ethical impact. The roadmap summarizes how managers could progressively implement ethical AI principles, policies, and practices. It provides guidance for decision-makers to address AI-specific ethical challenges in the financial services sector, thereby contributing to CDR best practice development.

Summary and conclusion

In conclusion, CDR and AI accountability are becoming increasingly important in the evolving technological environment of the financial services sector. Companies in this sector and beyond can demonstrate their commitment to responsible AI use via CDR, within which

Table 5
CDR Roadmap.



inclusivity and positive societal impact are integral. With the intensified use of AI, it becomes increasingly vital to ensure that these systems operate ethically (Lobschat et al., 2021). Accountability and the division of responsibilities should be made as clear as possible. The AI accountability framework applied to the financial services sector provides a structured approach to identifying and addressing ethical challenges associated with AI. The human oversight of AI-supported mechanisms is crucial for ensuring that the locus of morality remains identifiable. Additionally, the implementation of ethical conduct for financial institutions should involve a joint effort from stakeholders (Mueller, 2022). Policy-makers should encourage the sharing of best practices for ethical AI use among financial institutions at national and international levels. To support this, future research should explore different stakeholder perspectives on ethical AI usage and relevant cultural differences. Accountability and transparency of algorithmic trading, the ethical considerations of automated financial advice, and the ethical use of data in financial AI deserve future research attention, too. More extensive cross-cultural analyses and long-term impact assessments should be undertaken.

CRediT authorship contribution statement

Zsófia Tóth: Conceptualization, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Markus Blut:** Conceptualization, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

none.

Data availability

No data was used for the research described in the article.

References

(For a research-based discussion of corporate digital responsibility, see) Lobschat L, Mueller B, Eggers F, Brandimarte L, Diefenbach S, Kroschke M, Wirtz J. Corporate digital responsibility. J. Bus. Res. 2021;*122*:875–88.
(For more information about the impact of AI on banking, see) McKinsey. Global Banking Practice: Building the AI bank of the future. McKinsey & Company; 2021.
(For a research-based discussion of corporate digital responsibility, see) Mueller B. Corporate digital responsibility. Bus. Inf. Syst. Eng. 2022;*64*(5):689–700.
(For a research-based discussion and framework on AI robot accountability, see) Tóth Z, Caruana R, Gruber T, Loebbecke C. The dawn of the AI Robots: towards a new framework of AI robot accountability. J. Bus. Ethics 2022;178(4):895–916.
(For a research-based discussion of corporate digital responsibility, see) Wirtz J, Kunz WH, Hartley N, Tarbit J. Corporate digital responsibility in service firms and their ecosystems. J. Serv. Res. 2023;*26*(2):173–90.
(For additional information about the CDR award, see) BVDW, 2023. Corporate Digital Responsibility.⟨https://www.cdr-award.digital/en/⟩.
(For further information about the Financial Conduct Authority's recommendations on AI use, see) FCA, 2022. Data analytics and artificial intelligence (AI). ⟨https://www.fca.org.uk/firms/data-analytics-artificial-intelligence-ai⟩.
(For additional information on AI in banking, see) Townson S., 2020. AI can make bank loans more fair. *Harvard Business Review*. hbr.org.