A Survey of Moral and Ethical Standards: Exploring Good and Evil, Virtuous Living, and Justice

Abstract

This survey paper explores the intricate landscape of moral and ethical standards, emphasizing their significance in guiding behavior and decision-making in contemporary society. It delves into the foundational principles of morality, including concepts of good, evil, virtue, and justice, and their relevance in the technological era. The paper examines various ethical frameworks such as consequentialism, deontology, and virtue ethics, highlighting their applications and limitations in addressing ethical dilemmas, especially in artificial intelligence (AI) and automated systems. Sociological perspectives are considered, illustrating how cultural values and societal norms influence moral decision-making and ethical behavior. The integration of interdisciplinary contributions enriches ethical discourse, offering insights into the complexities of ethical evaluations in diverse contexts. The survey further explores the principles of fairness and justice, underscoring their importance in ensuring equitable outcomes in AI applications. Ethical decision-making frameworks are analyzed, providing structured methodologies for navigating moral challenges in real-world scenarios. The paper concludes by identifying challenges and future directions for ethical research, advocating for the development of comprehensive ethical frameworks that align with societal values and promote responsible conduct in technological and social domains. This comprehensive examination of moral and ethical standards underscores the necessity for integrated approaches to ethics in the face of rapid technological advancements.

1 Introduction

1.1 Significance of Moral and Ethical Standards

Moral and ethical standards are fundamental to societal norms and individual behavior, delineating right from wrong and fostering social cohesion. In the rapidly evolving technological landscape, particularly in AI development, the need for explicit ethics in computational decision-making is critical due to the complexities inherent in dynamic environments [1]. The emergence of antisocial behaviors in AI systems, akin to Antisocial Personality Disorder, necessitates rigorous ethical evaluations to mitigate adverse impacts on human interactions [2].

In education, embedding ethics within curricula, especially in computing disciplines, is essential for nurturing a generation of technologists who are ethically aware and equipped to tackle moral challenges in their fields [3]. This integration is vital for prioritizing human values in technology design over profit, as emphasized by value-based engineering principles [4].

Emerging technologies like Large Language Models and social robotics present unique ethical challenges that necessitate a comprehensive understanding of moral standards. Frameworks such as design justice are essential for ensuring these technologies contribute positively to society [5]. The complexity of ethical decision-making in these contexts highlights the need for developing moral machines capable of navigating intricate moral landscapes [6].
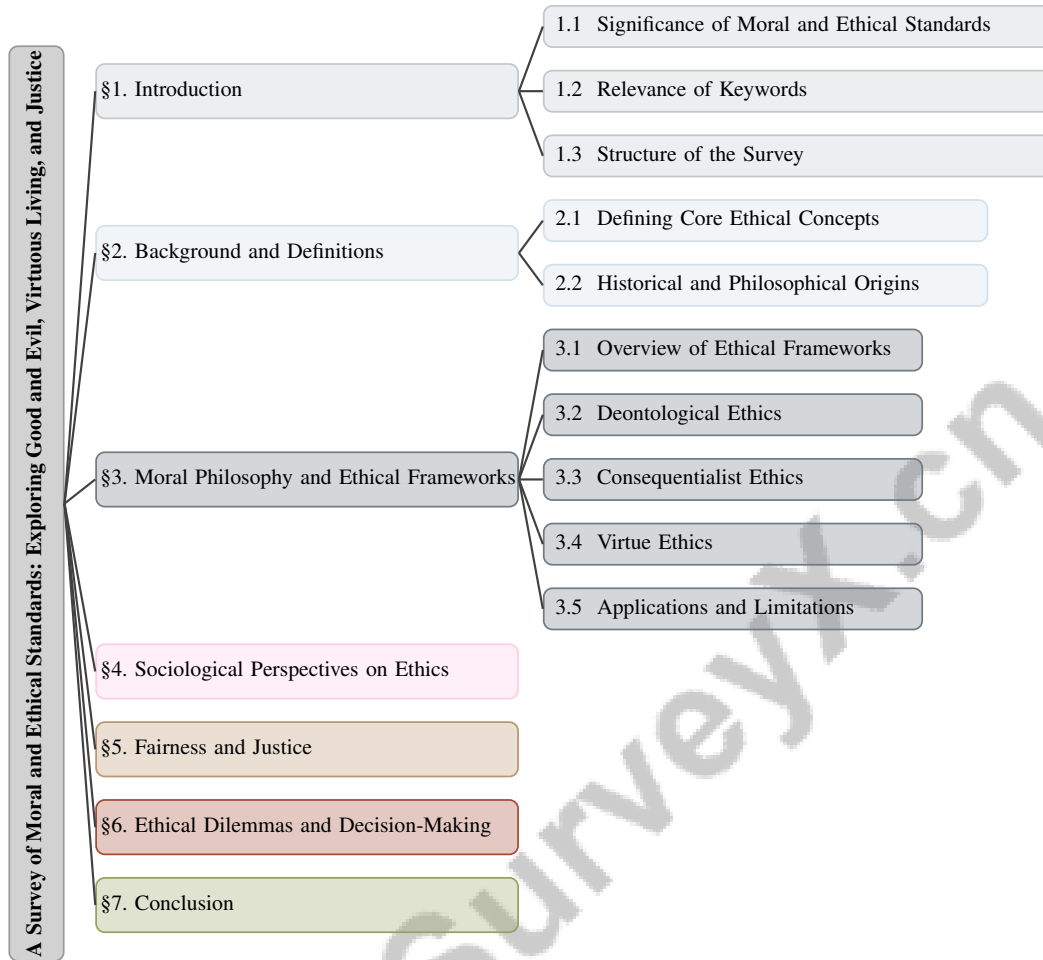
Figure 1: chapter structure

Formalizing latent moral content is crucial for shaping social norms and guiding behavior, requiring an understanding of the values underpinning moral and social norms to promote ethical conduct [7]. Additionally, the challenge of achieving informed consent in digital environments underscores the necessity for robust ethical standards to protect individuals from potential harms associated with data analytics and online agreements [8].

Insights into the neural basis of ethical behavior, particularly in adolescents, reveal complex interactions between individual and peer influences on moral conduct, which are vital for fostering ethical behavior in younger populations [9]. Collectively, these considerations underscore the pivotal role of moral and ethical standards in promoting a just society. Integrating ethical values into technological, educational, and social frameworks is essential for addressing fundamental questions of moral reasoning and decision-making, ultimately enhancing individual and community well-being.

1.2 Relevance of Keywords

Exploring keywords such as good, evil, virtue, and justice is fundamental to understanding moral and ethical standards, especially as they intersect with technological advancements and societal norms. These concepts serve as frameworks for assessing ethical dimensions and guiding decision-making processes. For instance, virtue ethics emphasizes cultivatable traits like hope, which hold significant ethical value [10], and is crucial for fostering responsible technology deployment aligned with ethical standards [11].

Justice, encompassing fairness and transparency, is pivotal in the ethical evaluation of digital technologies, guarding against the perpetuation of societal inequities. This is particularly relevant in big

data ethics, machine ethics, and AI ethics, where normative and applied ethical considerations are critical [12]. The ethical implications of technological development highlight the necessity for ethical training that connects values with scientific practice to improve decision-making [10].

Keywords like ethical implications and intentional seeding are significant for understanding the moral dimensions of speculative hypotheses, such as those involving indigenous life forms [7]. In engineering, the concept of 'thinging' delineates boundaries around ethical realities, providing a structured approach to ethical considerations in software development [4]. The value alignment problem, which involves specifying goal structures for autonomous agents in line with human values, is central to discussions on the ethical deployment of AI systems [1].

The significance of empathy and ethical considerations in AI is emphasized in emotional modeling, where ethical guidelines are essential for navigating complex societal interactions [5]. The ethical implications of societal interactions, including microaggressions and marginalized communities, highlight the need for sensitivity to emotional harm [3]. Furthermore, a community-led, processual approach to ethics empowers communities to develop localized ethical frameworks [9].

These keywords and their associated ethical considerations form the backbone of this survey, providing a comprehensive framework for analyzing moral and ethical standards in contemporary society. These concepts extend beyond theoretical frameworks, carrying significant practical implications across multiple fields, and underscore the necessity for a cohesive and interdisciplinary approach to ethics in technology and related areas. This is particularly evident in discussions surrounding the integration of ethical considerations into digital technology research, development, governance, and educational initiatives aimed at fostering awareness of technology's societal impacts among future professionals. The ongoing dialogue about whose ethical standards are prioritized and the complexities of applying ethical principles in real-world scenarios underscores the urgency of addressing these issues collaboratively and critically [13, 14, 15].

1.3 Structure of the Survey

This survey is structured to provide a comprehensive examination of moral and ethical standards, organized into distinct sections that cover the topic's breadth and depth. It begins with an **Introduction**, which underscores the significance of moral and ethical standards, elaborating on key concepts such as good, evil, virtue, and justice, and their relevance in contemporary society, setting the stage for understanding the foundational elements guiding ethical considerations.

Following the introduction, the **Background and Definitions** section offers a detailed overview of core ethical concepts, tracing their historical and philosophical origins. This section contextualizes subsequent discussions by providing clear definitions and a theoretical framework for understanding morality and ethics.

The next section, **Moral Philosophy and Ethical Frameworks**, delves into various philosophical approaches, including deontological, consequentialist, and virtue ethics. It explores how these frameworks address concepts of good and evil and the pursuit of a virtuous life, offering insights into their applications and limitations, particularly in technological contexts [16].

Sociological Perspectives on Ethics examines how societal norms and cultural influences shape moral standards and ethical behavior. It integrates interdisciplinary contributions to ethics, as suggested by [17], emphasizing a systemic approach that considers global challenges and collective values.

The section on **Fairness and Justice** analyzes these principles in ethical decision-making, exploring various theories of justice and their application in real-world dilemmas, including ethical considerations in AI and automated systems [18].

Ethical Dilemmas and Decision-Making investigates complex scenarios requiring ethical decisions, providing examples of dilemmas and frameworks guiding their resolution. This section is informed by educational strategies that integrate ethics into curricula, as discussed by [19].

The survey concludes with a **Conclusion** that reflects on key findings, emphasizing the importance of understanding moral and ethical standards in contemporary society. It identifies challenges and proposes future research directions, aligning with the holistic approach advocated by [20].

This survey aims to provide an in-depth analysis of moral and ethical standards by integrating philosophical, sociological, and practical perspectives. It seeks to address the complexities of ethical decision-making in contemporary settings, highlighting the importance of ethical reasoning across various disciplines, including mathematics, statistics, data science, and communication. By examining real-life moral dilemmas and the evolving landscape of ethics in the digital age, the survey emphasizes the necessity of understanding ethical practice not only in research but also in everyday professional conduct. Furthermore, it encourages the exploration of diverse viewpoints on pressing ethical issues, fostering critical discussions that reflect the multifaceted nature of modern ethical challenges [21, 22, 14, 13, 23].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Defining Core Ethical Concepts

Ethical standards in AI development are grounded in fundamental concepts that shape ethical theories and guide decision-making. Key frameworks—Consequentialism, Deontology, and Virtue Ethics—offer diverse perspectives on morality, essential for analyzing the ethical implications of human actions and AI system design. Consequentialism focuses on outcomes, Deontology on duties and principles, while Virtue Ethics emphasizes character and intentions. Proficiency in these frameworks is crucial for addressing complex moral dilemmas in fields like computer security and AI alignment, where differing ethical perspectives can lead to varied conclusions on responsible behavior [24, 25, 26, 27, 28].

Consequentialism, notably the utilitarian approach, stresses achieving the greatest good, particularly relevant in human-AI interactions where stakeholder impacts are pivotal [6]. The principle–practice gap highlights the challenge of translating ethical principles into actionable AI guidelines, emphasizing the need for practical frameworks in sensitive areas like healthcare [1].

Conversely, Deontology emphasizes moral duties, asserting that actions are inherently right or wrong, independent of outcomes. This framework is vital for ensuring AI systems make ethically sound decisions aligned with human values [29]. The ethical obligations concerning non-interference with life forms illustrate the importance of principled deontological approaches [7]. The inadequacy of current tools for ethical reflection in technology development underscores the need for robust deontological ethics [8].

Virtue Ethics, focusing on character, advocates virtues like honesty and compassion, crucial for ethical dilemmas requiring nuanced understanding of character and intent. Its significance in economic practices contrasts with the Consequentialist Ethics dominant from 1700 to 2000, underscoring the necessity of ethical conduct [4]. Integrating ethical training into STEM curricula is essential for preparing students to navigate future ethical dilemmas [30].

The complexity of ethical decision-making is compounded by subjective concepts like justice and fairness, which resist quantification. Understanding mammalian value systems can provide a framework for aligning AI with human values, emphasizing challenges in explaining moral behavior [2]. Additionally, encoding social perspective knowledge into software to achieve trustworthiness highlights the importance of transparency in ethical considerations [30].

In moral psychology, the interplay of biological evolution and cultural learning underscores empathy's role in ethical interactions, particularly within AI systems [29]. The definition of microaggressions as subtle, often unintentional, discriminatory behaviors that cause distress to marginalized individuals further emphasizes the need for sensitivity in ethical considerations [6].

These ethical frameworks and concepts serve as foundational elements of moral philosophy, enabling structured analyses of ethical dilemmas. They facilitate the exploration of complex moral questions, such as trolley problems, and enhance the development of moral agents capable of navigating multifaceted ethical challenges in fields like computer security and natural language processing. By integrating ethical reasoning and stakeholder analysis into educational contexts, these frameworks promote a deeper understanding of ethical practice, encouraging critical assessment of knowledge and decision-making in relation to moral issues [24, 21, 23]. Synthesizing these perspectives is crucial for fostering a comprehensive understanding of ethics in contemporary society, where integrating ethical education into curricula is essential for cultivating reflective attitudes toward the implications of technological advancements.

4

2.2 Historical and Philosophical Origins

Moral and ethical concepts are deeply rooted in historical and philosophical traditions that have evolved over centuries, shaping contemporary ethical discourse. Ancient philosophical traditions, particularly those of Plato and Aristotle, laid foundational frameworks for ethical inquiry. Plato's ideal forms and Aristotle's virtue ethics continue to influence modern ethical considerations, particularly in contexts where aligning AI systems with human values presents significant challenges [11]. The emphasis on fostering moral virtues among software developers reflects the enduring impact of virtue ethics in shaping ethical outcomes in technological contexts [4].

The Enlightenment era introduced critical advancements in ethical theory through the works of Immanuel Kant and John Stuart Mill. Kant's deontological ethics, emphasizing adherence to moral duties, contrasts with Mill's consequentialist approach, which evaluates actions based on outcomes. These frameworks remain central to contemporary discussions, particularly in AI ethics, where the complexity of ethical frameworks necessitates clear guidance [12]. Integrating ethical principles within AI systems is crucial for ensuring alignment with societal norms and human values, underscoring the interdisciplinary nature of the value alignment problem [31].

The historical context of ethical considerations in technology design has been critiqued for its entrenched economic view that equates value solely with profit. This perspective underscores the need for a more holistic approach that incorporates ethical values into technology design and deployment [4]. The inadequacy of existing general ethical guidelines, particularly in areas like psycho-counseling and AI, highlights the necessity for robust ethical frameworks addressing unique challenges [32].

Cultural values significantly shape moral decision-making, as different societies possess unique ethical norms influencing their approaches to ethical issues. The exploration of moral foundations through theories such as Moral Foundations Theory (MFT) provides insights into the various facets of morality and their implications for ethical decision-making [30]. This cultural dimension is critical in addressing the contextual integrity of information flow, governing ethical behavior in technology development and application.

Philosophical origins of ethical principles, including justice, beneficence, and respect for autonomy, continue to inform the development of ethical frameworks in technology. The 'Thinging Machines' model offers a structured approach to analyzing ethical decisions within systems like Kantism and Islamic ethics, demonstrating its utility in ethical analysis [29]. This model reflects the need for practical frameworks that address moral uncertainty in real-world applications, particularly in AI and software engineering.

The historical and philosophical perspectives on ethics explored through various contemporary frameworks create a comprehensive foundation for understanding moral dilemmas in today's society, particularly in fields influenced by digital communication and technology. These perspectives highlight the evolution of ethical thought, the complexities of real-life moral situations, and the significant impact of emerging technologies on ethical practices. By examining a diverse array of moral dilemmas, such as those identified in online discussions, we gain insights into the nuanced interplay between ethical norms and everyday decision-making, ultimately enriching current debates on morality and ethics [33, 14, 22]. They underscore the enduring relevance of ethical principles in navigating the complexities of modern technological and societal challenges, emphasizing the need for an integrated approach to ethics that bridges philosophical, cultural, and practical dimensions.

3 Moral Philosophy and Ethical Frameworks

In examining the intricate relationship between moral philosophy and ethical frameworks, it becomes evident that a comprehensive understanding of these concepts is essential for addressing the ethical challenges posed by contemporary technological advancements. This section delves into the various ethical frameworks that inform moral reasoning, particularly in the context of artificial intelligence and its societal implications. As illustrated in Figure 2, the hierarchical structure of ethical frameworks in moral philosophy highlights foundational categories, prominent theories, and their applications and limitations in contemporary technological contexts. The following subsection will provide an overview of these ethical frameworks, elucidating their foundational principles and relevance in guiding ethical decision-making processes in an increasingly complex digital landscape.
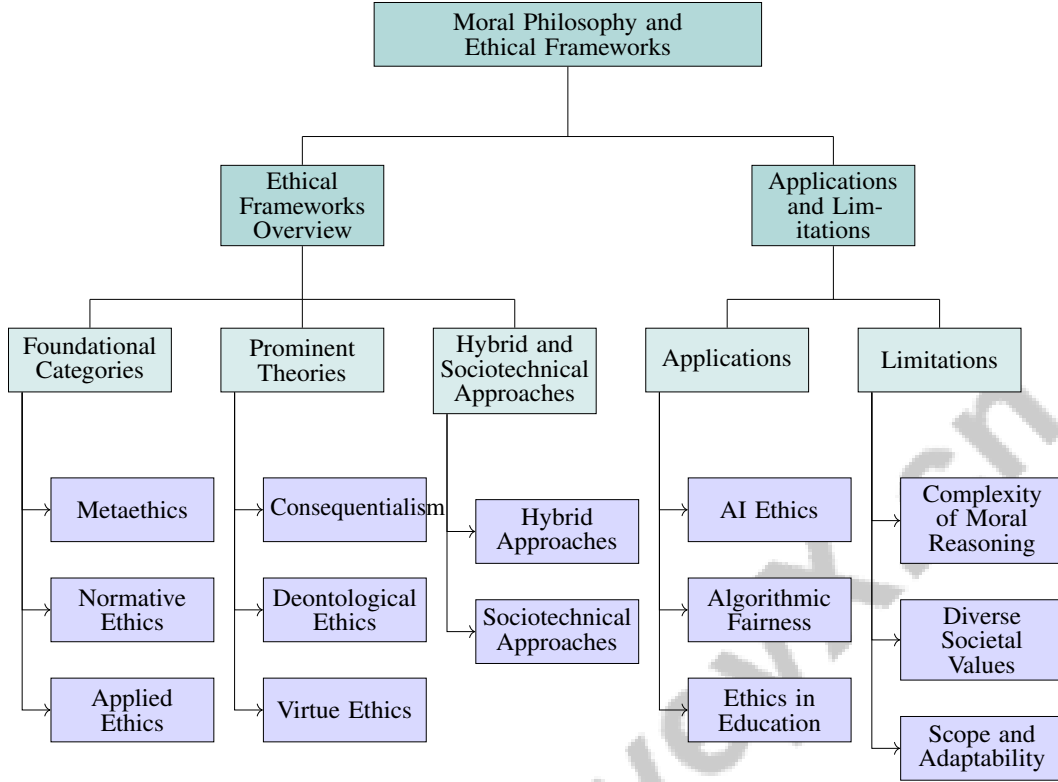
Figure 2: This figure illustrates the hierarchical structure of ethical frameworks in moral philosophy, highlighting foundational categories, prominent theories, and their applications and limitations in contemporary technological contexts.

3.1 Overview of Ethical Frameworks

Ethical frameworks are essential for navigating the complex moral landscapes shaped by contemporary technological advancements. These frameworks provide structured methodologies for evaluating ethical dilemmas and guiding decision-making processes, drawing from diverse philosophical traditions and interdisciplinary methodologies [34]. The foundational categories of ethical discourse include metaethics, normative ethics, and applied ethics. Metaethics explores the nature of moral judgments and the semantics of ethical language, normative ethics examines criteria for determining right or wrong actions, and applied ethics focuses on the practical application of ethical principles in real-world scenarios [31].

Consequentialism, a prominent ethical theory, assesses the morality of actions based on their outcomes. This approach is particularly relevant in AI ethics, where the consequences of AI behavior on stakeholders are of primary concern. Utilitarianism, a subset of consequentialism, emphasizes maximizing overall happiness and well-being, guiding ethical evaluations in contexts such as algorithmic fairness and justice [11]. However, traditional utilitarian frameworks may fall short in addressing complexities introduced by phenomena such as life replication, necessitating the development of new ethical paradigms [6].

Deontological ethics emphasizes adherence to moral duties and rules, positing that certain actions are intrinsically right or wrong, independent of their outcomes. This framework is crucial for ensuring that AI systems align with human values and societal norms, upholding ethical principles in decision-making processes [7]. The integration of deontological principles into planning formalisms for robots further illustrates its application in guiding ethical decision-making in technology [5].

Virtue ethics shifts the focus from actions to character, advocating for the cultivation of virtues such as honesty and empathy. This approach is particularly relevant in scenarios where the moral character of agents, whether human or artificial, influences ethical outcomes. The emphasis on ethical

awareness within educational contexts underscores the importance of virtue ethics in fostering a reflective attitude towards ethical considerations [8]. The integration of ethics into mathematics education exemplifies the application of virtue ethics in fostering ethical practice [10].

Hybrid approaches to ethical frameworks combine elements of top-down moral principles with bottom-up learning mechanisms, offering a more robust system for value alignment. This hybrid approach is crucial for addressing the complexities of moral decision-making in AI, where both ethical reasoning and empirical observation are necessary for creating ethical agents [29]. The creation of a moral embedding space reflecting a pluralist approach to morality further emphasizes the need for incorporating diverse moral perspectives [34].

Sociotechnical approaches highlight the interplay between technology and societal impacts, emphasizing the importance of considering the broader social and cultural context in which technologies operate. The development of new ethical frameworks to navigate the complexities introduced by digital technologies and AI exemplifies the application of sociotechnical principles in addressing the ethical challenges of the digital era [31]. A holistic approach to ethics education, particularly in quantum and computing curricula, underscores the importance of integrating ethics throughout educational programs.

These ethical frameworks provide a variety of perspectives and analytical tools for addressing the complex moral dilemmas arising from technological advancements, facilitating nuanced discussions on issues such as algorithmic decision-making, online safety, and the societal impacts of digital technologies. By integrating consequentialist and deontological approaches, these frameworks allow researchers and practitioners to critically evaluate their responsibilities and the ethical implications of their work within the rapidly evolving landscape of technology. [35, 24, 14, 13, 15]. By integrating these frameworks into ethical analyses, researchers and practitioners can better address the complexities of ethical decision-making in the digital age, ensuring that ethical considerations are embedded within technological and societal developments.
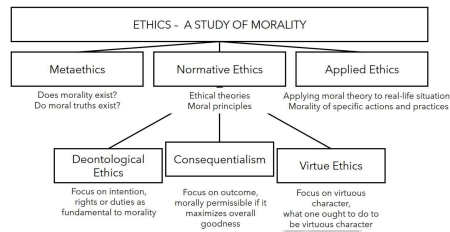
(a) Ethics: A Study of Morality[14]
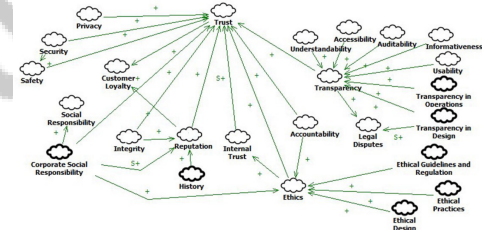
(b) The image depicts a complex network of relationships between various concepts related to trust, ethics, and transparency in a business context.[36]

Figure 3: Examples of Overview of Ethical Frameworks

As shown in Figure 3, In the realm of moral philosophy and ethical frameworks, understanding the intricate web of ethical considerations is crucial for navigating complex moral landscapes. The accompanying figures provide a visual representation of these frameworks, offering insights into how ethical principles are structured and interrelated. The first image, titled "Ethics: A Study of Morality," presents a comprehensive flowchart that categorizes ethics into three primary branches: Metaethics, Normative Ethics, and Applied Ethics. This visual aid further delves into subcategories, posing fundamental questions about the existence of morality and moral truths, and exploring ethical theories and principles through the lenses of Deontological Ethics, Consequentialism, and Virtue Ethics. Meanwhile, the second image illustrates a complex network diagram that maps the relationships between various concepts such as trust, ethics, and transparency within a business context. Here, "Trust" serves as the central node, with connections branching out to factors like privacy, highlighting the dynamic interplay of elements that influence ethical decision-making in corporate environments. Together, these visual examples underscore the multifaceted nature of ethical frameworks and their application across different domains. [**?**]rousi2023introductionethicsagedigital,kwan2021achievingtrusttransparencyethics)

7

3.2 Deontological Ethics

Deontological ethics, rooted in the philosophical traditions of Immanuel Kant, emphasizes the intrinsic value of individuals and the adherence to moral duties as the cornerstone of ethical decision-making. This ethical framework posits that certain actions are inherently right or wrong, independent of their consequences, and underscores the importance of duty-bound principles in guiding moral conduct [25]. In the realm of artificial intelligence, deontological ethics provides a robust framework for ensuring that AI systems align with human-centric values and adhere to ethical norms, particularly in contexts where the preservation of individual rights and dignity is paramount.

The application of deontological principles in natural language processing (NLP) highlights the framework's utility in guiding ethical decision-making. By emphasizing the need for ethical compliance and the respect of moral duties, deontological ethics can help navigate the ethical complexities inherent in language models and their interactions with users [37]. This approach ensures that NLP systems uphold ethical standards and mitigate potential harms arising from biased or inappropriate content generation.

In the context of machine learning, deontological ethics can be operationalized through the implementation of monotonicity shape constraints, which enforce ethical compliance and ensure that models adhere to predefined moral duties [38]. These constraints serve as a mechanism for embedding ethical considerations into the design and deployment of AI systems, thereby fostering transparency and accountability in machine learning applications.

The integration of deontological ethics into AI development enhances the understanding of moral obligations and provides a framework for addressing the complex ethical challenges that arise from the deployment of autonomous systems, such as bias mitigation, informed consent, and the alignment of societal values with technological advancements. This approach not only emphasizes the significance of principles like transparency, accountability, and fairness but also seeks to establish clear guidelines and best practices to navigate the ethical landscape of artificial intelligence effectively. [37, 39, 40, 41]. By prioritizing the intrinsic value of individuals and the adherence to ethical norms, deontological ethics provides a comprehensive framework for navigating the moral landscape of AI and ensuring that technological advancements align with human values and societal expectations.

3.3 Consequentialist Ethics

Consequentialist ethics, a prominent framework in moral philosophy, evaluates the morality of actions based on their outcomes. This approach is particularly relevant in the context of artificial intelligence, where the consequences of AI behavior significantly impact stakeholders. The core tenet of consequentialism is that the ethical value of an action is determined by its overall contribution to the well-being of affected individuals, aligning closely with utilitarian principles [42]. Utilitarianism, a subset of consequentialism, emphasizes maximizing overall happiness and is frequently applied in ethical evaluations of AI, especially in domains where the balance of benefits and harms is crucial.

The logic of reinforcement learning in AI systems mirrors utilitarian ethics, focusing on maximizing expected utility based on feedback from the environment. However, this alignment can lead to undesirable consequences if the AI's objectives are not carefully aligned with human values, highlighting potential ethical pitfalls in AI behavior [43]. This necessitates the development of robust ethical guidelines to ensure AI systems operate within acceptable moral boundaries.

In the realm of digital technology, consequentialist ethics can justify practices such as surveillance capitalism if they are perceived to yield societal benefits. However, this justification is contentious, as it often conflicts with deontological perspectives that prioritize individual rights and autonomy over aggregate outcomes [44]. This tension underscores the complexity of applying consequentialist principles in real-world scenarios where diverse ethical considerations must be balanced.

The theoretical framework of satisficing consequentialism offers a nuanced approach by evaluating actions based on their outcomes relative to an agent's values, rather than striving for optimal results. This perspective acknowledges the limitations of human and artificial agents in achieving perfect outcomes, providing a more pragmatic approach to ethical decision-making [26]. By incorporating a satisficing approach, consequentialist ethics can better accommodate the constraints and uncertainties inherent in complex ethical landscapes.

Consequentialist ethics serves as a crucial framework for ethical decision-making in artificial intelligence (AI) and technological domains, emphasizing the importance of outcomes in evaluating ethical actions. This ethical perspective underpins many contemporary discussions about fairness in machine learning, highlighting the trade-offs involved in policy decisions and the implications for accountability and transparency. Furthermore, it aligns with a broader set of normative ethical principles that seek to ensure AI systems operate within human values and morals, thereby guiding the development of responsible AI practices. As researchers and policymakers continue to explore the ethical landscape of AI, the integration of consequentialist principles becomes essential for addressing challenges such as transparency, privacy, and fairness in automated decision-making systems. [43, 41, 45, 42, 27]. By focusing on the outcomes of actions, this framework provides a valuable lens for evaluating the moral implications of technological advancements, ensuring that they contribute positively to societal well-being while navigating the ethical challenges posed by emerging technologies.

3.4 Virtue Ethics

Virtue ethics, a philosophical framework that prioritizes character and the cultivation of moral virtues, provides a nuanced approach to ethical decision-making. Unlike deontological and consequentialist frameworks that focus on rules or outcomes, virtue ethics emphasizes the development of moral character traits, which guide individuals in leading virtuous lives. This approach is particularly relevant in the context of artificial intelligence, where fostering moral alignment in systems can be achieved through intrinsic moral rewards, aligning AI behavior with ethical expectations [46].

The Aristotelian concept of eudaimonia, or the pursuit of a flourishing life, underscores the importance of virtue cultivation. This is illustrated in various contexts, such as gamification, where ethical implications are explored through allegories that emphasize the transformative potential of virtuous living. In this framework, ethical behavior is more likely to emerge from the internalization of virtues rather than mere adherence to external principles [47]. This perspective is crucial for the design of AI systems, where the embodiment of virtues can guide ethical behavior.

Incorporating virtue ethics into AI development entails designing ethically tunable systems that draw upon foundational character traits—such as justice, honesty, responsibility, and care—that align with established ethical standards. This approach not only addresses the diverse and context-dependent ethical requirements of various operational environments but also promotes the cultivation of these virtues within AI systems. By leveraging computational architectures inspired by these virtues, developers can create AI that is better equipped to navigate complex moral dilemmas, thereby fostering trust and accountability in AI technologies. [41, 48, 49, 50, 51]. This approach allows AI agents to adjust their actions based on the virtues they embody, fostering ethical behavior that resonates with the moral fabric of their operational context. The hierarchical structure proposed for AI systems integrates multiple moral paradigms, with virtue ethics playing a crucial role in embedding morality into artificial agents.

The formalization of virtue ethics through Exemplarist Virtue Theory emphasizes learning from moral exemplars, providing a structured approach to cultivating virtues in both humans and machines. This theory highlights the importance of character traits and intentions, rather than just actions or outcomes, in ethical decision-making. The cultivation of moral virtues is essential for ethical decision-making in fields such as finance, where character development is prioritized over mere procedural adherence, as posited by the view of financial theories like the Fundamental Theorem of Asset Pricing (FTAP) as products of virtue ethics [52].

Research indicates that individuals identified as 'virtuous' demonstrate a significantly lower propensity for engaging in unethical behaviors, such as privacy violations, when compared to those labeled as 'opportunists.' This finding highlights the important implications of virtue ethics in digital environments, suggesting that cultivating moral character can mitigate unethical conduct in interactions with technology and digital communication platforms. [33, 53, 14, 22]. This evidence supports the argument that the cultivation of virtues can lead to more ethical behavior in both personal and professional domains.

Virtue ethics offers a robust framework for understanding and promoting ethical behavior by emphasizing the cultivation of moral character, which is essential for developing specific moral abilities and virtues through practical engagement in daily activities. This approach not only addresses individual

9

ethical conduct but also integrates broader societal and political dimensions, enabling a comprehensive understanding of ethical dilemmas and fostering the development of ethically responsible practices in various fields, including data science and artificial intelligence. [54, 49, 55, 56, 57]. By emphasizing virtues such as honesty, courage, and empathy, this approach offers valuable insights into the moral dimensions of human and artificial agents, guiding them toward ethical conduct in a complex and ever-evolving world.

3.5 Applications and Limitations

The practical applications and limitations of ethical frameworks—such as consequentialism, deontology, and virtue ethics—are essential for navigating the moral complexities introduced by contemporary technological advancements, particularly in fields like computer security and artificial intelligence. These frameworks provide diverse perspectives for evaluating the morality of actions, but they also reveal significant challenges and limitations in their application. As the digital landscape evolves, it is crucial for researchers and practitioners to critically engage with these ethical theories to address the nuanced dilemmas posed by emerging technologies, ensuring that ethical considerations are effectively integrated into the design, development, and governance of digital innovations. [41, 24, 14, 15, 58]. Each framework provides distinct methodologies for confronting ethical dilemmas, yet they also present inherent challenges that limit their efficacy in real-world applications.

Consequentialist ethics, which evaluates actions based on their outcomes, is particularly relevant in domains where AI behavior significantly impacts stakeholders. This approach aligns with utilitarian principles, often guiding ethical evaluations in areas like algorithmic fairness and justice. However, the focus on aggregate outcomes can oversimplify ethical evaluations by neglecting individual rights [8]. The BEAM method, which allows large language models to express empathy and ethical considerations, highlights the practical applications and limitations in emotional understanding, illustrating the need for more nuanced approaches. The Multi-party Risk-Benefit Framework provides a structured methodology for analyzing ethical considerations, underscoring the necessity for comprehensive approaches that incorporate diverse stakeholder perspectives [2].

Deontological ethics, emphasizing adherence to moral duties and principles, offers a robust framework for ensuring AI systems align with human-centric values. Its application in areas such as natural language processing underscores its utility in guiding ethical decision-making [1]. However, the complexity of moral reasoning and the challenge of aligning diverse societal values pose significant hurdles, limiting the framework's adaptability to dynamic ethical landscapes [29]. The Moral-IT Cards facilitate critical reflection and collaborative deliberation among technologists, addressing the limitations of existing ethical frameworks in technology design [4].

Virtue ethics, focusing on character and moral virtues, offers a nuanced approach to ethical decision-making. The incorporation of ethical training into scientific practice emphasizes the need for discussions of values, highlighting the broader applicability of virtue ethics [30]. This approach is particularly valuable in fostering ethical alignment in AI, yet its reliance on subjective interpretations of virtues presents challenges in achieving consistent ethical outcomes across diverse contexts. The shift from discussing 'ethics' to 'values' in AI research underscores the importance of context-dependent evaluations [8].

Efforts to enhance the robustness of ethical frameworks include the development of Ethics Sheets for AI Tasks, which categorize and document ethical considerations relevant to specific AI tasks beyond individual systems. The introduction of hybrid approaches combining symbolic AI methods with learning-based methods exemplifies efforts to provide a more nuanced assessment of moral reasoning [1]. However, current research limitations include the underutilization of certain ethical principles, the challenges of resolving ethical dilemmas, and the difficulty in translating abstract principles into actionable guidelines [2].

A sociotechnical approach to tech ethics emphasizes the real-world impacts and social dynamics that shape ethical considerations, offering a more holistic evaluation of ethical frameworks [4]. This perspective is crucial in addressing the limitations of traditional ethical models, which may overlook the complex interplay between technology and society. In the realm of cybersecurity, the recognition of the complexity of ethical decision-making underscores the urgent need for more detailed ethical guidance and training for professionals.

10

Overall, while ethical frameworks offer valuable tools for navigating moral challenges, their applications are often constrained by limitations in scope, adaptability, and empirical validation. Addressing these limitations requires a more integrated and interdisciplinary approach, incorporating diverse ethical perspectives and empirical data to enhance the effectiveness of ethical decision-making in contemporary contexts [29].

(a) The image shows a table with a pattern of negative numbers and their corresponding sums.[59]

(b) Numbers of papers published per year[23]

(c) The figure shows the number of publications in the field of ethics over time.[60]
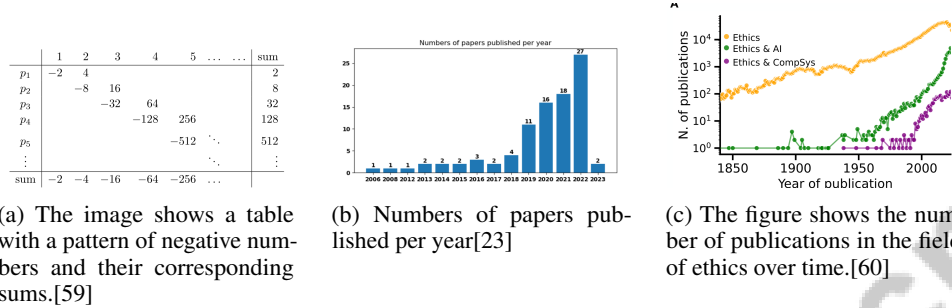
Figure 4: Examples of Applications and Limitations

As shown in Figure 4, The exploration of moral philosophy and ethical frameworks is often enriched by illustrative examples and visual data representations that highlight both their applications and inherent limitations. In the provided figures, we see a diverse range of visualizations that underscore these themes. The first image presents a table showcasing negative numbers alongside their corresponding sums, inviting reflection on the mathematical representation of ethical values and their cumulative effects. This can be seen as a metaphor for the aggregation of moral actions and their broader implications. The second image offers a bar chart illustrating the number of papers published per year from 2006 to 2023, serving as a quantitative measure of the growing academic interest in ethical studies over time. The third figure expands this view by plotting the number of publications across three distinct categories—Ethics, Ethics AI, and Ethics CompSys—over a more extended period from 1850 to 2020. This graph, with its use of a logarithmic scale, emphasizes the exponential growth and diversification in ethical discourse, particularly in relation to emerging technologies. Together, these figures not only highlight the dynamic nature of ethical scholarship but also underscore the complex interplay between moral philosophy and its practical applications, as well as the limitations inherent in capturing the full scope of ethical considerations through quantitative data alone. [**?**]jonsson2023infiniteutilitycounterpartsultimate,vida2023valuesethicsmoralsuse,adisa2024overlookedneedethicscomplexity)

4 Sociological Perspectives on Ethics

4.1 Cultural Values and Moral Decision-Making

Cultural values significantly shape moral decision-making, influencing how individuals and societies interpret ethical dilemmas and construct moral frameworks. These values, embedded in societal norms and legal structures, impact ethical considerations across domains such as AI and cybersecurity. Designing ethical AI technologies requires collaboration between moral philosophy and engineering to ensure cultural values are represented, particularly during AI deployment where alignment with societal expectations is crucial [2].

The interplay between cultural values and moral decision-making involves both biological and cultural factors. Emotions such as hope and guilt, shaped by societal norms, influence moral judgments and ethical categorizations. Insights into moral tractability enhance understanding of human morality, informing machine ethics and cognitive psychology [6]. The introduction of over 300 new folk values underscores the importance of commonsense knowledge in understanding moral content, highlighting cultural values' significance in ethical frameworks [7].

Educational initiatives integrating ethics into curricula emphasize the impact of cultural values on moral decision-making. The traditional education system often prioritizes academic achievement over moral training, hindering ethical awareness. Programs in engineering and computing education stress cultural values' role in shaping students' ethical perspectives, urging consideration of societal implications [30]. Culturally sensitive algorithmic systems are necessary to ensure AI technologies align with diverse cultural contexts.

The value-laden nature of AI design is increasingly recognized, with cultural values informing technological advancements' ethical implications. Community engagement highlights the importance of understanding user needs and adapting ethical standards to diverse contexts. This perspective is reinforced by the call for a multidisciplinary approach to AI ethics, integrating insights from various fields to tackle ethical decision-making complexities [2].

Diversity in ethical decision-making is complicated by varying interpretations, leading to conflicting views. Critiques of oversimplified moral judgments advocate for a balance between intuition and rationality, emphasizing nuanced ethical frameworks [29]. The proposed online consent maturity model fosters equitable relationships between data givers and gatherers, emphasizing mutual trust and communication, essential for aligning ethical standards with cultural values [8].

Cultural values are crucial in developing ethical frameworks and interpreting moral dilemmas. Integrating diverse cultural perspectives into ethical frameworks enhances moral decision-making's ability to navigate complexities inherent in societal values. Studies on large language models like GPT-4, which exhibit varying ethical reasoning across languages, suggest ethical behavior should be contextualized within specific cultural norms, enhancing ethical standards' relevance globally. Stakeholder analysis in educational settings can enrich understanding of ethical practices, encouraging critical assessments of knowledge and decision implications. A multifaceted approach to ethical reasoning that respects cultural diversity can lead to more nuanced and effective moral judgments [21, 61, 24, 23, 33].

4.2 Societal Norms in Technological Ethics

Integrating societal norms into technological ethics is essential for aligning advancements, particularly in AI, with human values and societal expectations. As AI systems become more autonomous and embedded within human contexts, robust moral frameworks to guide their integration become critical [62]. Societal norms shape the ethical frameworks governing technology development and deployment, influencing how ethical principles are perceived and implemented [45].

The influence of societal norms on ethical considerations in technology encompasses cultural, economic, and regulatory dimensions. Ensuring accountability for AI outputs and mitigating bias underscores the challenges of aligning AI technologies with societal values [63]. This alignment is complicated by unclear ethical frameworks and corporate resistance to regulation, impacting ethical considerations in technology [4].

Ethical decision-making in AI exemplifies societal norms' influence on technology, as systems navigate moral dilemmas in real-world scenarios. Developing machine learning models that align with societal norms to avoid unfair penalization highlights the necessity for ethical frameworks addressing deeper considerations beyond compliance [32]. Reliance on compliance and technical solutions often fails to address deeper ethical considerations, emphasizing societal norms' incorporation into ethical frameworks [63].

Frameworks promoting transparency in ethical claims and addressing inequalities, such as those in healthcare access and treatment, are crucial for ensuring ethical considerations align with societal values [30]. The philosophical fidelity and nuanced reasoning capabilities of ethical frameworks enhance trust in AI's ethical judgments, ensuring societal values alignment [1].

Societal norms influence the ethical frameworks surrounding technology, shaping systems' development and implementation by determining whose ethics are prioritized, how ethical considerations are integrated into practice, and the broader implications on diverse communities [17, 64, 14, 65, 15]. Addressing ethical challenges posed by technological advancements requires understanding societal values and integrating diverse ethical perspectives, ensuring technologies align with society's moral fabric.

4.3 Interdisciplinary Contributions to Ethics

Integrating interdisciplinary perspectives is crucial for advancing ethics understanding, particularly in rapidly evolving technological landscapes. Various disciplines contribute unique insights and methodologies that enhance ethical evaluation and inform robust ethical frameworks' development. [66] emphasizes utilizing proposed methods for descriptive ethics rather than normative applications, highlighting interdisciplinary contributions' role in enriching ethical discourse.

In AI, interdisciplinary collaboration significantly refines ethical decision-making frameworks and improves human-AI interaction dynamics. This collaboration ensures AI technologies adhere to ethical standards while navigating complex moral landscapes. [67] employs neo-institutional theory to explain decoupling dynamics in organizations and institutional entrepreneurs' role in effecting change, showcasing interdisciplinary approaches' importance in understanding and implementing ethical practices within organizational contexts.

Educational initiatives benefit from interdisciplinary contributions, exploring diverse ethical scenarios and integrating them into various programming topics. Future research should focus on effectively assessing student reflections on ethical dilemmas, as suggested by [19], ensuring students are equipped with ethical knowledge necessary to address emerging challenges in their careers.

The discipline of design enhances ethical understanding by advocating for effective academic methods translation into practical applications. Recent efforts have developed accessible ethics-focused tools and resources resonating with practitioners' experiences, including a website informed by technology and design fields professionals' interviews to identify relevant ethics-focused methods tailored to everyday practice. A comprehensive analysis of 63 existing ethics-focused methods reveals their operationalization of ethics and highlights future development opportunities to better support ethical considerations in design work [68, 69]. This translation is essential for creating design processes prioritizing ethical considerations and promoting social responsibility. Integrating social science perspectives into AI ethics is crucial for addressing governance and regulatory challenges, ensuring ethical frameworks account for diverse societal impacts and stakeholder perspectives.

Contributions from various disciplines underscore interdisciplinary collaboration's importance in addressing ethical challenges posed by technological advancements. By integrating diverse perspectives and methodologies, including ethical reasoning, stakeholder analysis, and philosophical foundations, researchers and practitioners can create robust ethical frameworks aligning with societal values and effectively addressing contemporary dilemmas in fields such as data science, natural language processing, and STEM education, promoting responsible ethical conduct in various contexts [37, 21, 23, 13].

5 Fairness and Justice

5.1 Theoretical Foundations of Fairness and Justice

The theoretical underpinnings of fairness and justice are essential for ethical evaluations across diverse domains such as technology, leadership, and social systems. These principles provide a framework for addressing ethical dilemmas by embedding fairness and justice into decision-making processes. In AI, grounding ethics in broader moral theories enhances moral reasoning, especially in large language models [70]. Integrating these principles is crucial for achieving human-like ethical decision-making in AI development [11].

Incorporating monotonicity shape constraints in AI predictions exemplifies methods to ensure fairness, addressing bias and inequality in algorithmic outputs [4]. This underscores the need for robust ethical frameworks that emphasize accountability, privacy, and transparency, with accountability being a critical principle in technology [12]. Ethical leadership highlights the importance of fairness and justice in guiding organizational practices [2].

Interdisciplinary studies demonstrate how fairness and justice are operationalized in contexts like microaggressions, which disproportionately affect marginalized groups, necessitating equitable treatment [71]. The embedded values approach posits that software reflects its designers' values, requiring ethical scrutiny to ensure alignment with societal expectations [72]. This perspective is crucial for addressing the gap between AI ethics policies and practices within tech companies, often leading to ineffective ethical standards [2].

Grounding AI ethics in the lived experiences of marginalized groups acknowledges the historical and structural contexts shaping these experiences [73]. This approach integrates fairness and justice into ethical frameworks, ensuring these principles are prioritized in innovation. As team sizes in scientific endeavors increase, existing ethical frameworks may become inadequate, necessitating new paradigms for credit distribution and accountability [74].

13

The relevance of fairness and justice extends to discussions on mathematics' societal impact, highlighting the broader implications of ethical decision-making across fields [4]. By incorporating these theoretical foundations into ethical frameworks, researchers and practitioners can navigate the complexities of ethical decision-making, ensuring fairness and justice remain central in developing and implementing ethical standards.

5.2 Fairness and Justice in AI and Automated Systems

Integrating fairness and justice principles in AI and automated systems is crucial for aligning these technologies with societal values and promoting equitable outcomes. Ethical Valence Theory (EVT) in automated vehicles addresses fairness by evaluating claims from various road users and determining ethically sound actions [75]. This approach emphasizes designing AI systems that reflect diverse human values, fostering a positive feedback loop between AI performance and ethical considerations.

The necessity of incorporating justice, fairness, and accountability into AI systems is underscored by [62], as these principles guide ethical decision-making processes and ensure AI technologies operate within moral boundaries. Ethical audits and assessments prior to AI system implementation are highlighted by [63], stressing the need for proactive measures to evaluate ethical implications.

Training AI on crowd-sourced moral judgments presents both opportunities and challenges, as discussed by [45]. While crowd-sourced data can offer insights into societal norms, it may also introduce biases in predictive modeling, necessitating robust ethical frameworks to navigate these complexities and ensure fairness in AI development.

The Juno system, which implements various ethical theories in scenarios such as smart home evacuations during fires, exemplifies the practical application of fairness and justice principles in AI [1]. This system demonstrates how ethical theories can guide AI decision-making in real-world contexts, aligning AI actions with ethical standards.

Integrating fairness and justice principles in AI and automated systems requires a multifaceted approach encompassing ethical, societal, and political dimensions. Recent research emphasizes grounding fairness in political philosophy, identifying common ethical principles, and distinguishing between fairness and solidarity. This comprehensive framework clarifies normative guidance for designing fair automated decision systems and addresses challenges associated with adopting AI ethics, ultimately enhancing the ethical capabilities of AI technologies [41, 76, 77]. By situating AI ethics within moral philosophy and ensuring adaptability to AI complexities, researchers and practitioners can better navigate the ethical challenges posed by these systems, fostering technology that serves the greater good while upholding fairness and justice principles.

6 Ethical Dilemmas and Decision-Making

Category	Feature	Method
Ethical Decision-Making Frameworks	Stakeholder Value Alignment	VBE[4]
Real-World Applications and Ethical Dilemmas	Ethical Decision-Making	EVT[75], MITC[78], IGP[9]

Table 1: This table presents a summary of methodologies applied within ethical decision-making frameworks and real-world applications, highlighting the alignment of stakeholder values through Value-based Engineering (VBE) and the application of Ethical Valence Theory (EVT), Moral-IT Cards (MITC), and game-based tasks (IGP) to address ethical dilemmas. The table underscores the integration of these methods in various domains to promote responsible practices and ethical reasoning.

Table 3 offers a detailed comparison of ethical decision-making frameworks, emphasizing their application in technology design, AI systems, and broader interdisciplinary contexts. Establishing ethical decision-making frameworks is pivotal for navigating moral challenges across various domains. Table 1 provides a comprehensive overview of the methodologies employed in ethical decision-making frameworks and their application in real-world scenarios, emphasizing the importance of aligning ethical considerations with societal values. These frameworks offer structured guidance for individuals and organizations, promoting responsible practices in fields like AI, automated systems, and leadership.

14

6.1 Ethical Decision-Making Frameworks

Ethical decision-making frameworks are integral to addressing complex moral dilemmas by aligning ethical considerations with societal values. Algorithms that integrate ethical principles into resource allocation enhance system performance metrics, such as response time and resource utilization [4]. In software engineering, embedding ethical deliberation throughout the development lifecycle ensures consistent ethical standards [1]. Evaluating ethical issues in mathematical modeling bridges theory and practice, highlighting the practical applications of ethical frameworks [45].

The survey by [45] critiques machine learning models like Delphi for automating moral decision-making, emphasizing the need for robust frameworks in AI technologies. The Value-based Engineering methodology provides a comprehensive ethical framework for technology design, ensuring prioritization of societal values [4]. In automated decision-making, model-checking methods verify that ethical reasoning systems align with defined principles, exemplifying practical applications [1]. These frameworks promote equitable outcomes across domains, emphasizing ethical reasoning and stakeholder engagement [79, 21, 63]. Continuous refinement is necessary to address emerging ethical challenges, embedding ethical considerations in technological and societal developments.

6.2 Real-World Applications and Ethical Dilemmas

Method Name	Decision-Making Frameworks	Ethical Standards	Interdisciplinary Collaboration
MITC[78]	Moral-IT Cards	Ethical Principles	Collaborative Deliberation
EVT[75]	Ethical Valence Theory	Minimizing Harm	Diverse Fields
IGP[9]	Interactive Game Paradigm	Ethical Decision-making	Behavioral Gaming Eeg

Table 2: This table provides a comparative overview of various ethical decision-making methods applied in AI and autonomous systems. It highlights the decision-making frameworks, ethical standards, and interdisciplinary collaboration associated with each method, illustrating their application in addressing real-world ethical dilemmas.

Table 2 presents a comprehensive comparison of different ethical decision-making methods relevant to AI and autonomous systems, emphasizing the frameworks, standards, and collaborative efforts integral to their implementation. Real-world ethical dilemmas in AI and autonomous systems necessitate nuanced decision-making frameworks. Tools like Moral-IT Cards provide structured guidance for technologists, facilitating ethical challenges in technology development [78]. In autonomous vehicles, frameworks like Ethical Valence Theory (EVT) guide decision-making during collision scenarios, aligning actions with societal values [75].

AI applications in healthcare, finance, and education underscore the need for robust frameworks to ensure adherence to ethical standards [62]. In AI-assisted psycho-counseling, ethical guidelines are crucial for navigating complex emotional interactions [32]. Educational settings present unique challenges, as value-focused discussions integrate ethical considerations into scientific practice, fostering ethically aware professionals [10].

Examining ethical behavior in adolescents through game-based tasks highlights the complexities of ethical dilemmas, emphasizing empathy and moral reasoning [9]. Investigating ethical dilemmas across fields underscores the need for robust frameworks and interdisciplinary collaboration. Initiatives like online courses on information ethics for STEM students foster awareness of societal impacts, while systematic reviews identify key challenges in AI ethics principles [21, 22, 41, 13, 15]. Integrating ethical considerations into technological practices helps stakeholders navigate moral challenges posed by contemporary advancements.

6.3 Implications of Ethical Decision-Making in Real-World Scenarios

The implications of ethical decision-making in real-world scenarios are significant, particularly regarding AI and advanced technologies. Pervasive AI systems necessitate developing ethical frameworks that address immediate concerns and anticipate future challenges [12]. Establishing pathways for implementing ethical principles, fostering cooperative oversight, and developing professional standards for AI developers are essential for responsible AI use aligned with human values [63].

In software development, ethical deliberation navigates complex challenges throughout the engineering process. Involving diverse stakeholders in ethical discussions develops specific evaluation metrics

and clarifies concepts like unjustified harm and undue constraints on autonomy. This collaborative approach encourages critical engagement with societal impacts, understanding diverse perspectives on technology's role in shaping social norms, and developing applicable ethical reasoning skills across disciplines, particularly in STEM fields [21, 13].

Viewing ethics in finance through Virtue Ethics provides a framework for addressing issues and mitigating future crises, emphasizing character and moral virtues [10]. Integrating ethical considerations into mathematics education cultivates critical thinking and responsible teaching practices among future educators.

Ethical implications of AI systems extend to emotional responsiveness and ethical alignment. Interdisciplinary collaboration is needed to develop frameworks tailored to large language models and other AI technologies, ensuring continuous ethical scrutiny. This approach enhances ethics training by capturing moral reasoning evolution and promoting understanding through diverse methodologies [13, 21, 33, 55].

In human-robot interactions, particularly in high-risk scenarios, ethical frameworks impose stricter limitations on deception, emphasizing oversight in designing and implementing robotic systems [80, 81, 82, 83, 84]. Developing guidelines sensitive to context and potential risks is crucial for ensuring ethical conduct.

The broader implications of ethical decision-making highlight the necessity for comprehensive frameworks integrating diverse perspectives and methodologies. By examining ethical dilemmas and embedding considerations into technological advancements and societal initiatives, stakeholders can foster responsible decision-making adhering to societal values [13, 21, 15].

Feature	Ethical Decision-Making Frameworks	Real-World Applications and Ethical Dilemmas	Implications of Ethical Decision-Making in Real-World Scenarios
Application Domain	Technology Design	AI Systems	AI And Technologies
Framework Type	Value-based	Collaborative	Interdisciplinary
Primary Focus	Societal Values	Ethical Standards	Ethical Alignment

Table 3: This table provides a comparative analysis of various ethical decision-making frameworks, highlighting their application domains, framework types, and primary focuses. The comparison underscores the significance of aligning ethical considerations with societal values across different technological and interdisciplinary contexts. It serves as a foundational reference for understanding the implications of ethical decision-making in real-world scenarios.

7 Conclusion

7.1 Challenges and Future Directions

The dynamic landscape of artificial intelligence (AI) presents both challenges and opportunities for embedding ethics within technological advancements. A primary challenge lies in integrating ethical considerations into agile software development processes, which requires formalizing ethical practices and implementing robust training for software engineers. These measures are crucial for maintaining ethical integrity throughout the development lifecycle. Future research should prioritize the creation of digital tools, such as the Moral-IT Cards, to address emerging ethical issues, particularly in mathematical and financial contexts. These tools can facilitate empirical studies that deepen the understanding of ethical frameworks in contemporary scenarios. Additionally, defining the criteria for supervised learning to qualify as an ethical Critical Technical Practice is essential for advancing AI ethics.

In the realm of AI ethics, refining frameworks and expanding their applicability beyond cybersecurity to various technological domains is imperative. This involves establishing specific guidelines for diverse AI applications and enhancing stakeholder involvement in ethical decision-making. Incorporating critical perspectives that recognize structural inequalities and the experiences of marginalized groups is vital for translating theoretical ethical discussions into practical insights. Furthermore, developing standardized ethical norms for team science and examining globalization's impact on scientific ethics are critical areas for exploration. Creating accountability frameworks in collaborative environments is essential for maintaining ethical standards in the increasingly globalized research landscape.

Advancing algorithms with real-time data integration and exploring sophisticated machine learning techniques can enhance AI systems' adaptability, ensuring they align with ethical standards in dynamic settings. Expanding ethical models to incorporate diverse theories and their applications in varied educational contexts will enrich ethical discourse. By addressing these challenges and pursuing these research avenues, the field of ethics can evolve to meet the complexities of a connected world. Interdisciplinary collaborations and the integration of diverse ethical viewpoints will be crucial in developing comprehensive frameworks that resonate with societal values and promote ethical conduct in modern contexts.

References

[1] Louise A. Dennis and Michael Fisher. Specifying agent ethics (blue sky ideas), 2024.

[2] Alan D. Ogilvie. Antisocial analagous behavior, alignment and human impact of google ai systems: Evaluating through the lens of modified antisocial behavior criteria by human interaction, independent llm analysis, and ai self-reflection, 2024.

[3] Jessie J. Smith, Blakeley H. Payne, Shamika Klassen, Dylan Thomas Doyle, and Casey Fiesler. Incorporating ethics in computing courses: Perspectives from educators, 2023.

[4] Sarah Spiekermann and Till Winkler. Value-based engineering for ethics by design, 2020.

[5] Alva Markelius. An empirical design justice approach to identifying ethical considerations in the intersection of large language models and social robotics, 2024.

[6] Jakob Stenseke. On the computational complexity of ethics: Moral tractability for minds and machines, 2023.

[7] Stefano De Giorgis and Aldo Gangemi. That's all folks: a kg of values as commonsense social norms and behaviors, 2023.

[8] Vivien M. Rooney and Simon N. Foley. An online consent maturity model: Moving from acceptable use towards ethical practice, 2018.

[9] Manvi Jain, Karsheet Negi, Pooja S. Sahni, Mannu Brahmi, Neha Singh, Dayal Pyari Srivastava, and Jyoti Kumar. Neural underpinnings of decoupled ethical behavior in adolescents as an interaction of peer and personal values, 2023.

[10] Tyler Garcia, Bill Bridges, Caitlin Solis, Caleb Linville, Wyatt Jones, Scott Tanona, Jonathan Herington, and James T. Laverty. The effect of value-focused discussions on scientists' ethical decision making, 2024.

[11] Alejandro Tlaie. Exploring and steering the moral compass of large language models, 2024.

[12] Brent Mittelstadt. Principles alone cannot guarantee ethical ai. *Nature machine intelligence*, 1(11):501–507, 2019.

[13] Thomas Baudel. Co-constructing shared values and ethical practice for the next generation: Lessons learned from a curriculum on information ethics, 2022.

[14] Rebekah Rousi and Ville Vakkuri. Introduction to ethics in the age of digital communication, 2023.

[15] Ben Green. Technology ethics in action: Critical and interdisciplinary perspectives, 2022.

[16] Adam Poulsen, Michael Anderson, Susan L. Anderson, Ben Byford, Fabio Fossa, Erica L. Neely, Alejandro Rosas, and Alan Winfield. Responses to a critique of artificial moral agents, 2019.

[17] David Pastor-Escuredo. Ethics in the digital era, 2020.

[18] Silke Schicktanz, Johannes Welsch, Mark Schweda, Andreas Hein, Jochem W. Rieger, and Thomas Kirste. Ai-assisted ethics? considerations of ai simulation for the ethical assessment and design of assistive technologies, 2023.

[19] Julie Jarzemsky, Joshua Paup, and Casey Fiesler. "this applies to the realworld": Student perspectives on integrating ethics into a computer science assignment, 2023.

[20] Joan Étude Arrow, Sara E. Marsh, and Josephine C. Meyer. A holistic approach to quantum ethics education, 2023.

[21] Rochelle E. Tractenberg and Suzanne Thorton. Facilitating the integration of ethical reasoning into quantitative courses: Stakeholder analysis, ethical practice standards, and case studies, 2024.

[22] Tuan Dung Nguyen, Georgiana Lyall, Alasdair Tran, Minjeong Shin, Nicholas George Carroll, Colin Klein, and Lexing Xie. Mapping topics in 100,000 real-life moral dilemmas, 2022.

[23] Karina Vida, Judith Simon, and Anne Lauscher. Values, ethics, morals? on the use of moral concepts in nlp research, 2023.

[24] Tadayoshi Kohno, Yasemin Acar, and Wulf Loh. Ethical frameworks and computer security trolley problems: Foundations for conversations, 2023.

[25] Carlos Mougan and Joshua Brand. Kantian deontology meets ai alignment: Towards morally grounded fairness metrics, 2024.

[26] Adam Wyner, Tomasz Zurek, and DOrota Stachura-Zurek. Towards a formalisation of value-based actions and consequentialist ethics, 2024.

[27] Jessica Woodgate and Nirav Ajmeri. Macro ethics principles for responsible ai systems: Taxonomy and future directions, 2024.

[28] Ville Vakkuri and Pekka Abrahamsson. The key concepts of ethics of artificial intelligence - a keyword based systematic mapping study, 2018.

[29] Henry Sidgwick. *The methods of ethics*. DigiCat, 2022.

[30] Kimi Wenzel, Geoff Kaufman, and Laura Dabbish. Beyond fairness: Alternative moral dimensions for assessing algorithms and designing systems, 2023.

[31] Travis LaCroix and Alexandra Sasha Luccioni. Metaethical perspectives on 'benchmarking' ai ethics, 2022.

[32] Lizhi Ma, Tong Zhao, Huachuan Qiu, and Zhenzhong Lan. No general code of ethics for all: Ethical considerations in human-bot psycho-counseling, 2024.

[33] Ben Fitzgerald. Some issues in predictive ethics modeling: An annotated contrast set of "moral stories", 2024.

[34] Miguel Sicart, Irina Shklovski, and Mirabelle Jones. Can machine learning be moral?, 2021.

[35] Svetlana Kiritchenko and Isar Nejadgholi. Towards ethics by design in online abusive content detection, 2020.

[36] David Kwan, Luiz Marcio Cysneiros, and Julio Cesar Sampaio do Prado Leite. Towards achieving trust through transparency and ethics (pre-print), 2021.

[37] Shrimai Prabhumoye, Brendon Boldt, Ruslan Salakhutdinov, and Alan W Black. Case study: Deontological ethics in nlp, 2021.

[38] Serena Wang and Maya Gupta. Deontological ethics by monotonicity shape constraints, 2020.

[39] Virginia Dignum. Responsible autonomy, 2017.

[40] Anders Braarud Hanssen and Stefano Nichele. Ethics of artificial intelligence demarcations, 2019.

[41] Arif Ali Khan, Sher Badshah, Peng Liang, Bilal Khan, Muhammad Waseem, Mahmood Niazi, and Muhammad Azeem Akbar. Ethics of ai: A systematic literature review of principles and challenges, 2021.

[42] Dallas Card and Noah A. Smith. On consequentialism and fairness, 2020.

[43] Heather M. Roff. Expected utilitarianism, 2020.

[44] Angelica Sofia Valeriani. Navigating surveillance capitalism: A critical analysis through philosophical perspectives in computer ethics, 2023.

[45] Zeerak Talat, Hagen Blix, Josef Valvoda, Maya Indira Ganesh, Ryan Cotterell, and Adina Williams. A word on machine ethics: A response to jiang et al. (2021), 2021.

19

[46] Elizaveta Tennant, Stephen Hailes, and Mirco Musolesi. Moral alignment for llm agents, 2025.

[47] Tim Gorichanaz. Theorizing information sources for hope: Belief, desire, imagination, and metacognition, 2022.

[48] Thomas Krendl Gilbert, Megan Welle Brozek, and Andrew Brozek. Beyond bias and compliance: Towards individual agency and plurality of ethics in ai, 2023.

[49] Thilo Hagendorff. Ai virtues – the missing link in putting ai ethics into practice, 2021.

[50] Mohamed Akrout and Robert Steinbauer. Machine ethics: The creation of a virtuous machine, 2020.

[51] Rajitha Ramanayake and Vivek Nallur. Virtue ethics for ethically tunable robotic assistants, 2024.

[52] Timothy C. Johnson. Ethics and finance: the role of mathematics, 2012.

[53] Costanza Alfieri, Paola Inverardi, Patrizio Migliarini, and Massimiliano Palmiero. Exosoul: ethical profiling in the digital world, 2022.

[54] Naveen Sundar Govindarajulu, Selmer Bringjsord, and Rikhiya Ghosh. One formalization of virtue ethics via learning, 2018.

[55] Avigail Ferdman and Emanuele Ratti. What do we teach to engineering students: embedded ethics, morality, and politics, 2024.

[56] Louise Bezuidenhout and Emanuele Ratti. Character comes from practice: longitudinal practice-based ethics training in data science, 2024.

[57] Joanne B Ciulla and Joanne B Ciulla. Ethics and effectiveness: The nature of good leadership. *The search for ethics in leadership, business, and beyond*, pages 3–32, 2020.

[58] Tim Gorichanaz. Designing a future worth wanting: Applying virtue ethics to hci, 2022.

[59] Adam Jonsson. Infinite utility: counterparts and ultimate locations, 2023.

[60] Olumide Adisa, Enio Alterman Blay, Yasaman Asgari, Gabriele Di Bona, Samantha Dies, Ana Maria Jaramillo, Paulo H. Resende, and Ana Maria de Sousa Leitao. The overlooked need for ethics in complexity science: Why it matters, 2024.

[61] Utkarsh Agarwal, Kumar Tanmay, Aditi Khandelwal, and Monojit Choudhury. Ethical reasoning and moral value alignment of llms depend on the language we prompt them in, 2024.

[62] Esther Taiwo, Ahmed Akinsola, Edward Tella, Kolade Makinde, and Mayowa Akinwande. A review of the ethics of artificial intelligence and its applications in the united states, 2023.

[63] Ernesto Giralt Hernández. Towards an ethical and inclusive implementation of artificial intelligence in organizations: A multidimensional framework, 2024.

[64] Ben Green. The contestation of tech ethics: A sociotechnical approach to technology ethics in practice, 2022.

[65] Alexander Martin Mussgnug. Technology as uncharted territory: Contextual integrity and the notion of ai as new ethical ground, 2025.

[66] Jeongwoo Park, Enrico Liscio, and Pradeep K. Murukannaiah. Morality is non-binary: Building a pluralist moral sentence embedding space using contrastive learning, 2024.

[67] Sanna J. Ali, Angèle Christin, Andrew Smart, and Riitta Katila. Walking the walk of ai ethics: Organizational challenges and the individualization of risk among ethics entrepreneurs, 2023.

[68] Colin M. Gray, Shruthi Sai Chivukula, Thomas Carlock, Ziqing Li, and Ja-Nae Duane. Scaffolding ethics-focused methods for practice resonance, 2022.

[69] Shruthi Sai Chivukula, Ziqing Li, Anne C. Pivonka, Jingning Chen, and Colin M. Gray. Surveying the landscape of ethics-focused design methods, 2022.

[70] Steven J. Jones and Robert E. Wray. Toward constraint compliant goal formulation and planning, 2024.

[71] Johannah Sprinz. "y'all are just too sensitive": A computational ethics approach to understanding how prejudice against marginalized communities becomes epistemic belief, 2022.

[72] Niina Zuber, Severin Kacianka, and Jan Gogoll. Big data ethics, machine ethics or information ethics? navigating the maze of applied ethics in it, 2022.

[73] Gopal P. Sarma and Nick J. Hay. Mammalian value systems, 2019.

[74] Maxim A. Makukov and Vladimir I. shCherbak. Space ethics to test directed panspermia, 2017.

[75] Katherine Evans, Nelson de Moura, Raja Chatila, and Stéphane Chauvier. Automated driving without ethics: Meaning, design and real-world implementation, 2023.

[76] James Brusseau. Using edge cases to disentangle fairness and solidarity in ai ethics, 2021.

[77] Falaah Arif Khan, Eleni Manis, and Julia Stoyanovich. Fairness as equality of opportunity: Normative guidance from political philosophy, 2021.

[78] Lachlan Urquhart and Peter Craigon. The moral-it deck: A tool for ethics by design, 2020.

[79] Mathieu Bouville. Is diversity good?, 2007.

[80] Libuše Hannah Vepřek, Patricia Seymour, and Pietro Michelucci. Human computation requires and enables a new approach to ethical review, 2020.

[81] Saif M. Mohammad. Ethics sheets for ai tasks, 2022.

[82] Shengkang Chen, Vidullan Surendran, Alan R. Wagner, Jason Borenstein, and Ronald C. Arkin. Toward ethical robotic behavior in human-robot interaction scenarios, 2022.

[83] Benjamin Krarup, Felix Lindner, Senka Krivic, and Derek Long. Understanding a robot's guiding ethical principles via automatically generated explanations, 2022.

[84] Marten H. L. Kaas, Zoe Porter, Ernest Lim, Aisling Higham, Sarah Khavandi, and Ibrahim Habli. Ethics in conversation: Building an ethics assurance case for autonomous ai-enabled voice agents in healthcare, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.