Individual Technology Ethics Case Study
Fall 2024

Masoomeh Akbari

1 Introduction and Framing

Jibo is a compact social robot designed for home use, intended to serve as a personal companion and assistant. It sits on a counter or shelf, integrating smoothly into domestic environments. Equipped with advanced artificial intelligence, Jibo can recognize faces and voices, allowing it to adapt to and interact uniquely with multiple users. Its key features include a video camera, microphone, and a 360-degree microphone array, enabling it to respond to commands, take photos, and even engage in conversations. These capabilities are supported by continuous cloud connectivity, through which Jibo stores and processes data, enhancing its ability to "learn" and respond more accurately over time [4].

One of Jibo's defining characteristics is its anthropomorphic design. The robot's expressive movements, friendly "face," and human-like gestures make it a personable, emotionally engaging presence, almost like a family member. While this design creates an inviting user experience, it also introduces ethical concerns. Jibo's cloud-based data processing and anthropomorphic design raise issues around user privacy and potential manipulation through its lifelike appearance, both of which are explored further in the next section. (178 words)

2 Ethics Analysis

2.1 Issue 1: Privacy and Data Security Concerns

As a cloud-connected device, Jibo continuously collects, stores, and processes data, including video and audio recordings, to provide personalized responses and interactions. This feature raises serious privacy concerns, especially given the sensitive nature of the

Figure 1: This picture is taken from here

data collected within private homes. When data is sent to external servers, there is a risk of data breaches, unauthorized access, or misuse of personal information, which could lead to privacy violations or potential misuse of users' data.

According to Diakopoulos's work on transparency in algorithmic decision-making (ADM), Jibo's data handling would greatly benefit from clear, open practices that allow users to understand what data is collected, how it's processed, and the specific purposes it serves [3]. Diakopoulos emphasizes that transparency supports accountability, an important factor in reducing privacy risks and helping users make informed choices about the technology they use daily.

Furthermore, Millar's discussion on social failure modes in technology highlights how Jibo may fail to meet user expectations of privacy, especially in a personal setting like the home [8]. While Jibo is designed to feel like a friendly companion, its constant data collection could create a feeling of being watched, which conflicts with the trusted space of a home. Bender et al. add to this concern by warning that large-scale AI systems, particularly those connected to big data sources, often lack strong protections to prevent unintentional privacy leaks and misuse of sensitive data [1]. Together, these factors show the complex, sometimes hidden ways that data-driven AI systems like Jibo can affect user privacy, needing careful thought in their design and use. (258 words)

2.2   Issue 2: Manipulation through Anthropomorphism

Jibo's design uses anthropomorphic, or human-like, features to build emotional connections with users, making it feel more like a "friend" than just a device. This approach encourages frequent interaction and even loyalty, but it also raises ethical concerns about the influence on vulnerable groups such as children and the elderly. When users feel emotionally connected to Jibo, they may be more likely to trust it, share personal details, or rely on it for companionship. This attachment can make users more willing to reveal private information or create a reliance on the device, which could negatively impact their independence and overall well-being.

Leong and Selinger's research on "dishonest anthropomorphism" discusses how human-like traits in robots can mislead users into seeing them as more intelligent or emotionally aware than they actually are [7]. Jibo's expressive movements, eye-like display, and friendly gestures foster a bond that can shape user behavior in ways they might not fully realize. According to Leong and Selinger, such designs encourage interactions that don't match the robot's real capabilities, which can lead to misunderstandings or misplaced trust.

Additionally, Lacey and Caudwell's work on "dark patterns" in robotics highlights how Jibo's "cute" design works as a gentle yet strong way to attract users' interest [6]. Cuteness appeals to human emotions, instinctively drawing users in, which may cause them to overlook trade-offs around privacy and independence. This appealing design taps into deep-rooted human responses, making users more likely to welcome Jibo into their personal lives without questioning the full implications. These elements of Jibo's design show how anthropomorphic features can blur the line between machine and companion, impacting users in complex emotional and psychological ways that they may not anticipate. (279 words)

3   Ethics Design Element

To address the issue of **privacy and data security concerns**, a design change could involve processing as much data locally on Jibo as possible, rather than relying on cloud storage. By allowing Jibo to store and process user data directly within the device, sensitive information—like audio recordings and facial recognition data—could be kept within the home, minimizing the need to send data to external servers. This would reduce the risk of data breaches, unauthorized access, and privacy violations, as user information would not need to leave the personal, trusted space of the home.

In addition, local data processing would give users greater control over what data Jibo collects and retains, offering more transparency and building trust. If data storage on the device reached capacity or if users wanted to delete certain records, they could easily do so without depending on remote servers. This approach would prioritize user privacy, making Jibo a safer choice for home environments by significantly lowering the chance of personal data being exposed externally. (169 words)

4   Design Element Defense

Processing data locally on Jibo rather than relying on cloud storage is a strong approach to address privacy concerns and aligns well with ethical standards in technology design. By keeping user data within the device and within the home, this change respects users' expectations of privacy in their personal space, as highlighted by Millar's analysis of social failure modes in technology [8]. Millar points out that technology designed for private spaces should respect the norms and privacy expectations inherent to those settings, reducing the sense of surveillance and fostering trust.

Additionally, Chowdhury's work on the need for understandable AI supports this approach by emphasizing that AI systems should be designed to be clear and transparent to everyday users, not only technical experts [2]. Without data being sent to external servers, users gain greater control over their personal information and can make informed choices about what data Jibo retains. This simplicity and control make the technology easier for users to understand, promoting transparency and accountability.

Moreover, Bender et al. emphasize the risks associated with large-scale data collection in cloud-connected AI systems [1]. These risks include unintentional privacy violations and potential data misuse. By storing data locally, Jibo avoids these risks, ensuring a safer interaction for users. Overall, this design change aligns with ethical principles of privacy and understandability, making it a responsible improvement for a home robot like Jibo. (228 words)

5   Anticipating Objections

A primary objection to using only local data storage on Jibo is that it limits the device's ability to adapt and improve over time. Cloud-based systems give Jibo access to extensive data and advanced processing, allowing it to "learn" from patterns observed across many users and situations. This continuous connection to the cloud enables Jibo to stay updated with the latest AI advancements, making its responses more flexible and better aligned with users' evolving needs. Restricting Jibo to local storage, on the other hand, could reduce its responsiveness and make it less likely to remain relevant as user

expectations shift. As Chowdhury emphasizes, AI should work alongside users, providing responses that are clear and helpful for everyday interactions [2]. Millar and Kerr further argue that limiting AI systems to local data can hinder their ability to achieve optimal outcomes [9].

Without cloud access, Jibo's learning is limited, which could lead to repetitive or restricted interactions. Users who view Jibo as a "family member" may feel disappointed if the device behaves more like a static machine than an adaptable companion. Jones and Millar suggest that metaphors like "family member" shape how users expect a robot to behave, and failing to meet these expectations could reduce users' interest and engagement. If Jibo does not act as the "companion" it is marketed to be, users may disengage, ultimately seeing less value in the device [5]. (232 words)

References

[1] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, pages 610–623. Association for Computing Machinery, 2021.

[2] Rumman Chowdhury. Is explainability enough? why we need understandable ai. *Forbes*, 2018.

[3] Nicholas Diakopoulos. *Transparency*. The Oxford Handbook of Ethics of AI, 2020.

[4] Jibo. Website. https://robotsguide.com/robots/jibo, 2024.

[5] Meg Leta Jones and Jason Millar. Hacking metaphors in the anticipatory governance of emerging technology: The case of regulating robots. In Roger Brownsword, Eloise Scotford, and Karen Yeung, editors, *The Oxford Handbook of Law, Regulation and Technology*. Oxford University Press, 2017. Print Publication Date: Jul 2017.

[6] Cherie Lacey and Catherine Caudwell. Cuteness as a 'dark pattern' in home robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 374–381, 2019.

[7] Brenda Leong and Evan Selinger. Robot eyes wide shut: Understanding dishonest anthropomorphism. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, page 299–308, New York, NY, USA, 2019. Association for Computing Machinery.

[8] Jason Millar. *Social Failure Modes in Technology and the Ethics of AI: An Engineering Perspective*. The Oxford Handbook of Ethics of AI, 2020.

[9] Jason Millar and Ian R. Kerr. Delegation, relinquishment and responsibility: The prospect of expert robots. 2013.