Northeastern University
CS5100 Foundations of Artificial Intelligence
Summer 2025
Professor Sarita Singh
Erdun E
July 18, 2025

Assignment 8 Answer

Question 1

Artificial Intelligence (AI) systems pose unique risks due to their autonomy and ability to scale rapidly across industries. Unlike traditional software, AI can make decisions without direct human oversight, which increases the chance of unintended outcomes. For example, biased algorithms have caused discriminatory practices in hiring and lending decisions [3]. Deepfake technologies also threaten political stability and social trust by enabling realistic false media [2]. Real-world incidents such as Microsoft's Tay chatbot, which became racist within 24 hours of deployment [6], highlight how AI systems can behave unpredictably in dynamic environments.

Compared to other computer science technologies, such as conventional software systems, AI introduces a higher level of unpredictability. While software bugs can be patched and security vulnerabilities mitigated, autonomous AI systems can act in unexpected ways as they adapt to new data or environments. For instance, the 2010 "Flash Crash" in financial markets was partly attributed to AI-powered trading systems making rapid, unsupervised decisions that destabilized markets within minutes [5]. Moreover, AI could potentially automate cyberattacks on a large scale, surpassing the speed and efficiency of human hackers.

When considering bio-, nano-, and nuclear technologies, AI's risks are different in scope and accessibility. Biological and nuclear technologies often require specialized facilities and materials, making them harder to misuse. In contrast, AI systems can be distributed globally over the internet, making regulation and control much more challenging. However, biological risks, such as the creation of synthetic pathogens, could have more immediate catastrophic effects [1]. Similarly, nuclear accidents like the Fukushima disaster demonstrate the extreme physical destruction possible when such technologies fail [4].

In summary, AI's potential for widespread misuse, combined with its accessibility and scalability, creates a distinct category of risk. Although its destructive power may not match nuclear weapons or engineered viruses, the difficulty of containment and rapid adoption across sectors make it a significant societal challenge.

Question 2

Facebook's facial recognition software works by analyzing the facial geometry from photos you and others upload. This model gets trained on images where friends are "tagged," learning unique facial features like eye spacing and jawline [11]. Once trained, the system can suggest name tags on new photos automatically.

From a privacy perspective, several serious issues arise. First, users often do not provide true informed consent. Facebook faced multiple lawsuits, such as the Illinois biometric privacy class action, which led to a $650 million settlement in 2021 [8]. The platform collected and used biometric face data without explicit permission. Second, this technology enables mass surveillance: it can track people across photos and time without their awareness, raising concerns about constant monitoring and the potential misuse of data.

On a social level, facial recognition can normalize being constantly identified in public or private pictures. In Europe, even with GDPR protection, many users "agree" to facial recognition simply because it's the default option or hard to find the setting to opt out [9]. This creates social pressure to accept tagging, even when people want privacy. Worse, bad actors could use the same tools to monitor activists or minority groups, increasing risks of abusive targeting.

In summary, Facebook's photo recognition creates a tension between convenience and privacy. Understanding its technical functioning helps frame the problem, but it is the legal, social, and ethical consequences often invisible to users that truly demand careful attention.

Question 3

The weak AI hypothesis claims that machines can simulate human thinking and be useful tools for cognitive processes, but do not genuinely understand or have consciousness. In contrast, the strong AI hypothesis argues that sufficiently advanced machines could truly understand and be minded, possessing real cognitive states akin to humans [10].

The Turing test, introduced by Alan Turing in 1950, is designed to assess whether a machine's responses are indistinguishable from a human's. It specifically examines whether a computer can simulate human-like intelligence well enough to fool an interrogator [12]. Therefore, it addresses the weak AI hypothesis, focusing on behavior rather than true understanding. The test does not guarantee that the machine has conscious understanding; it only checks for externally observable performance.

In conclusion, Turing's approach highlights the functional aspect of intelligence, reinforcing the weak AI perspective. By testing what a machine can do rather than what it observes intrinsically, the Turing test offers a practical method to evaluate simulated intelligence without requiring real understanding.

Question 4

Google tailors search results by collecting and analyzing users' data, such as web history, location, cookies, and logged in account activity to build personalized profiles. These profiles include interests, demographics, and behavior preferences, enabling Google to rank and display results believed to be most relevant [14].

This personalization creates several social challenges. One key issue is the formation of "filter bubbles," meaning users may only see information aligned with their existing views. This can isolate individuals from diverse perspectives, potentially undermining public discourse and increasing political or ideological polarization [13]. Additionally, tailored content can be used to manipulate opinions advertisers or political actors could exploit these profiles to influence decision making.

Finally, users may lose autonomy as algorithms subtly guide what they see, base their judgments, and reinforce preferences without conscious awareness.

There are also important ethical concerns. First, privacy is threatened because extensive data is collected, often without clear informed consent. Second, personalization may compromise fairness and lead to discrimination: different users might receive different information or opportunities, for example in job postings or news coverage. Third, consent becomes ambiguous when defaults are set to "opt-in" or hidden within complex settings, and users often are unaware they're part of personalized pipelines.

In sum, while Google's tailored services offer convenience and relevance, they also pose significant social and ethical risks filter bubbles, manipulation, autonomy loss, privacy breaches, consent ambiguity, and algorithmic fairness concerns must be critically addressed.

Question 5

Smartphones collect GPS data via their built-in location sensors and expose this information to apps through permission settings. Many navigation or social apps continuously record precise latitude and longitude coordinates and may share this with third parties such as advertisers or analytics services [7]. Even when GPS data is coarsened or anonymized, researchers have shown that a small number of location points can uniquely identify most individuals, e.g., just four points identify 95% of users [15].

This raises several social issues. First, constant location tracking can allow advertisers, app developers, or even acquaintances to monitor users over time. Second, malicious actors or stalkers may exploit these apps for harassment, threatening personal safety. Third, the ubiquity of tracking can erode social trust, as people may feel watched or pressured to modify their behavior if others can always know where they are.

There are also serious ethical concerns. While users often grant location permission without fully understanding the extent or purpose, true informed consent is rarely obtained. GPS data can become a tool for mass surveillance, governments and corporations alike can misuse it for oversight or controlling populations. Additionally, such data can serve to discriminate, service providers might offer different pricing or access based on one's location or movement patterns, impacting fairness and equal treatment.

In sum, GPS-based tracking via smartphones brings convenience but also deep social risks and ethical trade-offs. Understanding technical mechanics helps highlight the urgency of transparent consent, robust data protection, and policy measures to prevent misuse.

References

[1] Center for Strategic and International Studies. *Understanding the National Security Commission on Emerging Biotechnology Report*. 2025. URL: https://www.csis.org/analysis/understanding-national-security-commission-emerging-biotechnology-report.

[2] Robert Chesney and Danielle Citron. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security". In: *California Law Review* 107.1 (2019), pp. 1753–1819. URL: https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security.

[3] Kate Crawford. "Artificial Intelligence's White Guy Problem". In: *The New York Times* (2016). Opinion. URL: https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html.

[4] International Atomic Energy Agency. *The Fukushima Daiichi Accident*. Tech. rep. IAEA, 2015. URL: https://www-pub.iaea.org/mtcd/publications/pdf/pub1710-reportbythedg-web.pdf.

[5] Andrei Kirilenko et al. "The Flash Crash: The Impact of High-Frequency Trading on an Electronic Market". In: *The Journal of Finance* 72.3 (2017), pp. 967–998. URL: https://doi.org/10.1111/jofi.12545.

[6] Melissa Locker. *Microsoft's Tay AI Chatbot Goes Rogue, Tweets Racist Remarks*. Time. 2016. URL: https://time.com/4274052/microsoft-tay-chatbot-racist/.

[7] Stylianos Monogios et al. "A Case Study of Intra-library Privacy Issues on Android GPS Navigation Apps". In: *arXiv* (2021). URL: https://arxiv.org/abs/2109.03664.

[8] Rosen, Grace and Gorov and Salkin, LLP. *Cases in re Facebook Biometric Information Privacy Litigation*. Rosen, Gorov Salkin, LLP. 2021. URL: https://www.rgrdlaw.com/cases-in-re-facebook-biometric-info-privacy-litig.html.

[9] Kerry Rusbridger and Simon Parkin. "Live in the EU? Facebook Is After Your Face Data (Again)". In: *Wired* (2018). URL: https://www.wired.com/story/facebook-facial-recognition-opt-out-settings-lawsuit-turn-off-gdpr-eu/.

[10] John R. Searle. "Minds, Brains, and Programs". In: *Behavioral and Brain Sciences* 3.3 (1980), pp. 417–457. URL: https://web-archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf.

[11] Time Staff. "How Facebook Knows What You Look Like". In: *Time* (2015). URL: https://time.com/3951006/facebook-visual-recognition/.

[12] Alan M. Turing. "Computing Machinery and Intelligence". In: *Mind* 59.236 (1950), pp. 433–460. URL: https://www.csee.umbc.edu/courses/471/papers/turing.pdf.

[13] Wikipedia contributors. *Filter bubble*. Wikipedia, The Free Encyclopedia. Retrieved Month Day, 2025. 2025. URL: https://en.wikipedia.org/wiki/Filter_bubble.

[14] Wikipedia contributors. *Google Personalized Search*. Wikipedia, The Free Encyclopedia. Retrieved Month Day, 2025. 2025. URL: https://en.wikipedia.org/wiki/Google_Personalized_Search.

[15] Wikipedia contributors. *Privacy and location-based services*. Wikipedia. 2025. URL: `https://en.wikipedia.org/wiki/Privacy#Internet`.