

Course: Data and Information Literacy

Answers to Lecture 1: Introduction to Data and Information Literacy

Lecturer: Dr. Johnson Masinde

1. What are the key ethical principles that should guide the collection and use of personal data?

- **Privacy:** This involves limiting data collection to what is necessary and relevant, and using the data in ways that align with the individual's expectations and consent.
- **Informed Consent:** Informed consent requires that individuals are fully aware of how their data will be collected, used, and shared.
- **Transparency:** Transparency involves openly communicating with individuals about how their data is collected, processed, and used.
- **Data Minimization:** Data minimization is the principle of collecting only the data necessary for a specific purpose.
- **Security:** Data security is essential for protecting personal data from unauthorized access, theft, or misuse.
- **Fairness and Non-Discrimination:** The principle of fairness requires that data collection and use should not lead to unjust or discriminatory outcomes.

2. What challenges arise when applying the principle of informed consent in the era of big data and advanced analytics?

- **Complexity of Data Use:** It is difficult to provide individuals with a clear understanding of all the ways their data may be used, especially when future analyses or applications may not be foreseeable. This undermines the ability to obtain fully informed consent.
- **Data Repurposing:** In the big data ecosystem, data collected for one purpose is often repurposed for secondary uses, which may not have been originally disclosed to the individual.
- **Difficulty in Withdrawing Consent:** In many cases, once data has been collected and incorporated into large datasets, it is challenging for individuals to revoke their consent or remove their data.
- **Understanding of Consent Forms:** Consent forms in the context of big data often contain complex legal and technical jargon that the average person may struggle to understand.
- **Data Aggregation and Anonymization:** In big data analytics, individual data is often aggregated or anonymized to remove personal identifiers. However, re-identification techniques can sometimes reverse this anonymization, allowing personal information to be uncovered.

- **Volume and Frequency of Consent Requests:** With the increasing use of data collection tools across different platforms, individuals are frequently asked to provide consent for multiple applications, services, and websites. This can lead to "consent fatigue,"

3. **What are some common ethical issues related to data ownership and intellectual property, and how can organizations address these issues to ensure fair use of data?**

- **Ambiguity in Data Ownership:** One of the central ethical challenges is the ambiguity surrounding data ownership. When individuals share their data with organizations (e.g., social media platforms or mobile apps), they often lose control over how it is used or shared. It is unclear who owns the data—the individual who generated it or the company that collects and processes it. This ambiguity can lead to ethical dilemmas regarding who has the right to use, sell, or modify the data.
- **Intellectual Property Rights of Data Creators:** Data creators, such as researchers, organizations, or individuals, may generate valuable datasets. However, questions of intellectual property arise when third parties use or monetize this data without proper attribution or compensation. In academic and research settings, this can lead to disputes over data ownership, especially when multiple parties contribute to a dataset's creation.
- **Misuse of Data:** Ethical issues arise when data is used for purposes other than what was originally intended. For example, data collected for research purposes may later be used for commercial activities without the data subject's knowledge or consent. This raises concerns about respecting the intent behind data collection and ensuring it is not exploited for unethical purposes.
- **Data Monetization Without Compensation:** Organizations often collect and monetize data without providing any compensation to the individuals who generated it. This raises ethical concerns, particularly when companies profit from personal data without acknowledging or rewarding the data source. In sectors like social media and online platforms, this issue is prevalent, with users providing valuable personal information in exchange for "free" services.
- **Access to Data:** Data ownership issues also involve determining who has the right to access and use data. Ethical concerns arise when data is hoarded by organizations, preventing other researchers, businesses, or the public from accessing valuable datasets. This can stifle innovation and limit the societal benefits that could be derived from open data sharing.
- **Control and Accountability:** Data ownership carries with it the responsibility to control how data is used and to be accountable for any misuse. Ethical issues arise when organizations fail to protect the data they collect or when they use it irresponsibly, such as by sharing it with third parties without proper oversight.