Understanding Data: Concepts, Governance, Ethics, and EU Perspectives

Defining and Contextualising Data

Data is the foundation of information and a core component of modern communication, decision-making, and technological advancement. At its essence, data refers to raw, unprocessed facts or figures that can be collected, measured, and analysed. It exists in various forms—numerical, textual, visual, or symbolic—and is given meaning through context and interpretation.

Historically, data has always served a communicative function. Early humans recorded experiences through cave paintings, tally marks on bones, and inscriptions on stones to share knowledge, mark time, or map territory (Academy, 2023). In ancient Egypt, papyri were used to document information in a structured manner, resembling early forms of spreadsheets or checklists (Data, 2019). These rudimentary practices illustrate the timeless human impulse to record and convey information.

A major transformation occurred during the 20th century with the advent of digital computing. The development of machines like the ENIAC during World War II marked a turning point, as data could now be processed at electronic speeds. The evolution of computer architectures led to the creation of structured query languages (SQL), and later, big data frameworks like Hadoop, which made the storage and retrieval of large-scale data seamless and rapid.

Today, data is a dynamic asset. Its interpretation fuels data analytics, recommendation engines, and **Large Language Models (LLMs)**—once the realm of science fiction. This progression highlights the importance of a clear understanding of what data is and how it functions. Without precise definitions and governance, data can be misused, misinterpreted, or weaponized, particularly in decision-making contexts. As such, a nuanced understanding of data is essential for ethical and effective application in society.

Defining and Contextualising Governance and Data Governance

Governance, in a general sense, refers to the frameworks, structures, and processes through which organizations and societies make decisions, exercise authority, and ensure accountability. It encompasses policies, rules, and practices aimed at achieving collective goals while maintaining order and ethical standards. In the context of data, governance applies these principles to ensure data is accurate, secure, accessible, and used responsibly.

Data governance, therefore, is the systematic management of data to ensure it meets organizational standards and supports strategic objectives. It involves the coordination between business units and IT departments to maintain data quality, privacy, security, and compliance with regulations. Effective data governance creates a reliable environment in which data can be transformed into actionable insights for business and social value (Lewis, 2023).

Key components of data governance include:

• Strategy and Doctrine:

Establishing a clear strategy ensures that data governance aligns with the overall goals of the organization. Doctrine refers to a guiding philosophy or set of principles that inform how data policies are developed and implemented. Industry standards provide a benchmark to ensure consistency, interoperability, and best practices across sectors (Derof, 2023).

• Responsibility:

Clearly defining roles and responsibilities helps assign ownership for data stewardship, access control, and quality assurance. This ensures accountability throughout the data lifecycle and facilitates better communication between departments. Roles such as data owners, custodians, and stewards are often designated to oversee various aspects of governance (Derof, 2023).

• Quality Standards:

Maintaining high data quality is crucial for making accurate, reliable decisions. Frameworks like ISO 8000, DMBoK (Data Management Body of Knowledge), and CMMI-DM (Capability Maturity Model Integration for Data Management) offer structured methodologies to manage data accuracy, consistency, completeness, and relevance. These standards help organizations assess their current data maturity and work toward continuous improvement (Derof, 2023).

• Regulatory Compliance:

Compliance with legal and regulatory frameworks such as the **General Data Protection Regulation (GDPR)** in the EU and **Health Insurance Portability and Accountability Act (HIPAA)** in the US is critical to safeguard individual rights and organizational integrity. These regulations dictate how personal and sensitive data should be collected, stored, processed, and shared. Failure to comply can lead to legal penalties, reputational damage, and loss of public trust (Derof, 2023).

A critical analysis of governance also reveals its dual purpose. Organizations may approach data governance either to reduce costs or to unlock long-term strategic value. Some scholars note that the latter requires a forward-looking mindset that aligns

governance with innovation and ethical responsibility. However, some scholars caution that overly rigid governance models can stifle agility and innovation, particularly in fast-paced digital environments. This underscores the need to balance control with flexibility when implementing data governance frameworks.

Data Ethics

Data ethics is a branch of applied ethics that examines the moral obligations and implications surrounding data collection, sharing, analysis, and use. It deals with issues such as privacy, consent, fairness, and accountability, especially in contexts where data influences human lives and social structures (Project, 2019).

For instance, when individuals fill out surveys, data may be collected, analysed, and shared. But were they informed and did they consent? Were institutional ethical requirements followed? These questions are central to ethical data practice (Project, 2019).

The volume and sensitivity of data collected by companies like Google—millions of webpages and search queries stored over six months—raise concerns about transparency and user control. Ethical considerations must be embedded at every stage of data handling (Project, 2019).

Ultimately, data ethics seeks to ensure that data practices align with values of justice, dignity, and human autonomy.

EU Data Governance and Ursula von der Leyen

Contemporary EU data governance is grounded in the principle of **digital sovereignty**—Europe's ability to act independently in the digital domain while upholding fundamental values such as privacy, fairness, and transparency. This vision was clearly articulated by European Commission President Ursula von der Leyen in her 2020 **State of the Union** address.

Digital sovereignty:

There have been growing concerns about the economic and social influence of non-EU technology companies. These firms have had extensive control over the personal data of European citizens, posing potential security threats and hindering the growth of European high-tech companies. This situation also limits the ability of European lawmakers to enforce their own regulations. Digital sovereignty refers to Europe's capacity to act autonomously in the digital realm. It encompasses both protective

mechanisms and proactive tools aimed at fostering innovation and ensuring technological independence (Madiega, 2020).

Ursula von der Leyen, President of the European Commission, identified digital policy as a key political priority for her 2019–2024 term. She pledged to achieve high levels of technical sovereignty in critical digital sectors. According to recent reports from the European Election Commission, competition from global technology players—who often disregard EU rules and core values—has become a major policy challenge. These companies prioritize data appropriation and monetization, which conflicts with the EU's human-centric digital agenda (Madiega, 2020).

In response, the European Parliament has raised concerns about security risks stemming from the growing presence of Chinese technologies in Europe. It has committed to reducing this dependency by fostering the development of secure, inclusive, ethical, and high-performance software, with strong emphasis on data security and artificial intelligence. The COVID-19 pandemic, which struck Europe in the spring of 2020, underscored the critical role that high-tech companies play in sustaining social life, business continuity, and public administration. This experience accelerated the EU's focus on achieving sovereign digital capabilities (Madiega, 2020).

In the post-pandemic recovery, the European Council has called for decisive action to secure the EU's strategic autonomy. This includes substantial investments in digital infrastructure, technological innovation, and capacity building—key pillars of Europe's broader recovery and resilience strategy (Madiega, 2020).

Secure digital identity:

The EU's digital wallets empower citizens to share specific documents, access services, and disclose selective information—such as age—without revealing their entire identity. This approach ensures that individuals have full control over their personal data and can track who has access to it. Some practical use cases of the EU digital identity include accessing public services, such as retrieving birth certificates, updating addresses, opening bank accounts, filing income tax returns, and applying to universities. Not only can these services be used in a citizen's home country, but also in other EU member states. Additionally, the digital identity allows for the storage of medical prescriptions, which can be accessed and used across Europe, as well as facilitating tasks like renting a car or checking into a hotel (Leyen, 2020).

Millions of Extra Jobs:

One of the key benefits of the EU's digital governance strategies is the creation of millions of new jobs, driven by advancements in technologies such as AI, data analytics, and digital infrastructure. As generative AI and other digital tools transform

industries—particularly those involving repetitive, rule-based tasks in sectors like manufacturing, technology, healthcare, and finance—it's crucial that lawmakers actively monitor and analyse these shifts in order to create policies that address skill gaps and workforce needs. For example, the EU could encourage educational curricula that emphasize creativity, critical thinking, and interdisciplinary skills, preparing students for high-demand professions in fields such as teaching, medicine, law, AI development, and embedded systems. This approach not only fosters job creation but also aligns with the EU's broader goals of promoting economic growth, innovation, and social inclusion, ensuring that citizens are well-equipped to thrive in an increasingly digital world.

Secure digital Identity: EU digital wallets enable citizens to share documents, access services and share parts of their information such as age without revealing their entire identity. Citizens have full control and be aware of who has their data. Some use cases of the EU digital identity include public services such as retrieving birth certificates, updating addresses, opening a bank account, filing income tax returns, applying for university not only in the home country but also in another EU member state, storing medical prescriptions that could be used anywhere in Europe, renting a car, checking into a hotel, etc (Commision, 2020).

Bibliography

Academy, C. (2023). Retrieved from https://www.youtube.com/watch?v=74B9cne1PPE&t=5s

Data, H. o. (2019, August 8). Retrieved from https://humansofdata.atlan.com/2019/08/historical-humans-of-data-great-pyramid-of-giza/

Lewis, K. (2023, August 17). Retrieved from https://www.youtube.com/watch?v=kiYPjNj9AmU

Derof, B. L. (2023, July 23). Retrieved from https://www.youtube.com/watch?v=iR8H-nzQZf8&t=163s

Project, O. L. (2019). Retrieved from https://www.youtube.com/watch?v=l-k_1RQmmVY

Leyen, U. v. (2020). Retrieved from https://www.youtube.com/watch?v=1X6F3_Tugwk&t=1028s

Commision, E. (2020). Retrieved from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en

Madiega, T. (2020, July). Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf