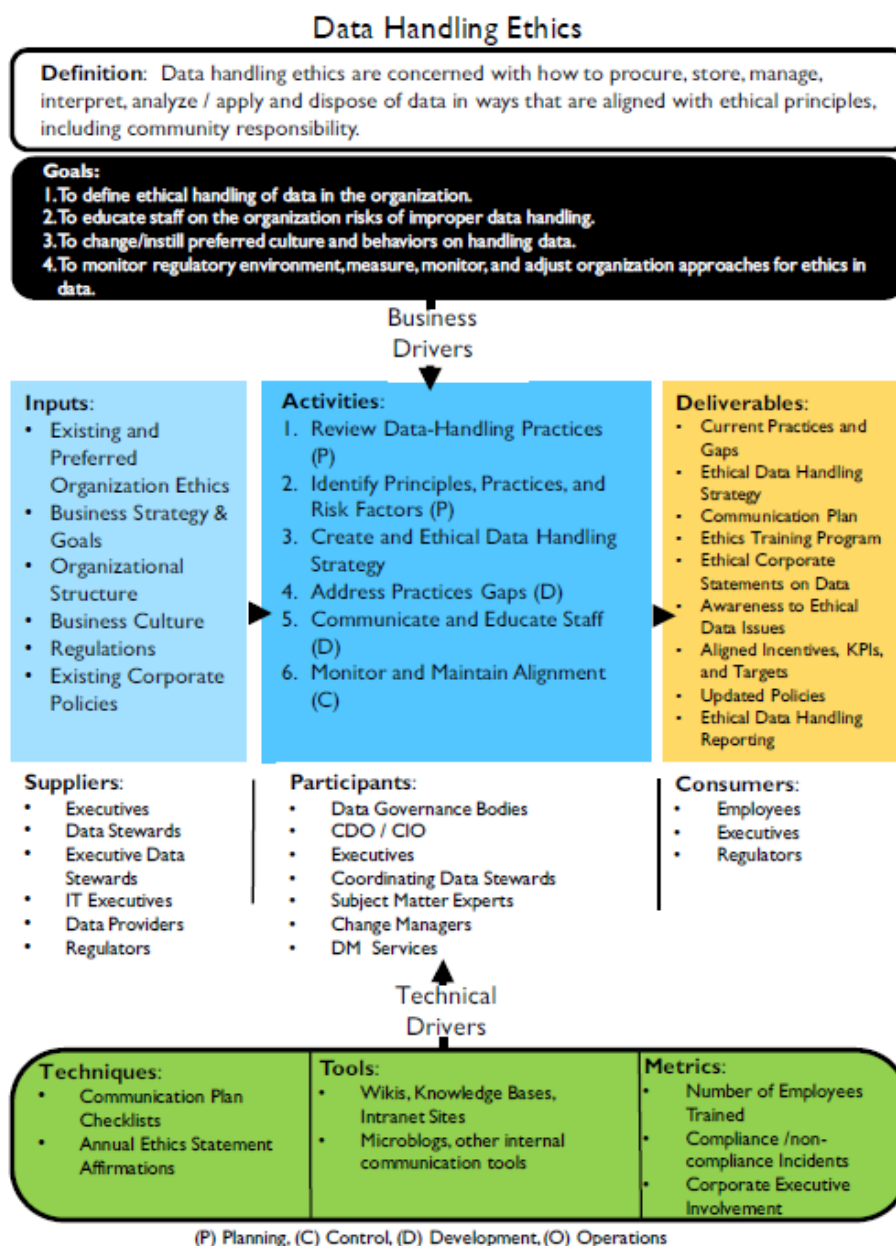


Data Handling Ethics

1 Introduction

Core Concepts:

- **Impact on people:**
 - Imperative to maintain the quality and reliability of individuals' data
 - Used to make decisions that impact people's lives
- **Potential for misuse:**
 - Negative effects on people and organisations
- **Economic value of data:**
 - Determine how and by whom that value of data can be accessed



2 Business Drivers

Ethics means “doing it right when no one is looking”.

Ethical handling of data is a competitive business advantage for an organisation:

- Increased trustworthiness and improved relationships – a competitive advantage
- Better relationships with stakeholders
- Reduced risk of misuse or a criminal breach
- Responsibilities when sharing data
- Organisation wide commitment to handling data ethically

Chief Data Officer, Chief Risk Officer, Chief Privacy Officer, Chief Analytics Officer focus on controlling risk by establishing acceptable practices for data handling.

3 Essential Concepts

3.1 Ethical Principles for Data

- **Respect for persons:** Treat people in a way that respects their dignity and autonomy as human individuals. Protect the dignity of those with “diminished Autonomy”.
- **Beneficence:** Firstly, do no harm. Secondly, maximise possible benefits and minimise possible harms.
- **No maleficence:** Minimise harm
- **Justice:** The fair and equitable treatment of people.

European Data Protection Supervisor, 2015 – the “engineering, philosophical, legal and moral implications” of developments in **data processing and Big Data**:

- Future-oriented regulation of data processing and respect for the rights to privacy and data protection.
- Accountable controllers who determine personal information processing
- Privacy conscious engineering and design of data processing products and devices
- Empowered individuals

3.2 Principles behind Data Privacy Law

Privacy regulations have been driven by human rights violations experienced during WW2 in Europe. The European Union’s Organisation for Economic Co-operation and Development (OECD) established 8 Guidelines and Principles for Fair Information Processing.

- Limitations on data collection
- An expectation that data will be of a high quality
- The requirement that when data is collected it will be for a specific purpose
- Limitations on data usage
- Security safeguards
- An expectation of openness and transparency
- The right of an individual to challenge the accuracy of data related to himself or herself
- Accountability of organisations to follow the guidelines

These underpin the GDPR:

3.2.1 General Data Protection Regulation of the EU

GDPR Principle	Description of Principle
Fairness, Lawfulness, Transparency	Personal data shall be processed lawfully, fairly, and in a transparent manner in relation to the data subject.
Purpose Limitation	Personal data must be collected for specified, explicit, and legitimate purposes, and not processed in a manner that is incompatible with those purposes.
Data Minimization	Personal data must be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed.
Accuracy	Personal data must be accurate, and where necessary, kept up-to-date. Every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purpose for which they are processed, are erased or rectified without delay.
Storage Limitation	Data must be kept in a form that permits identification of data subjects [individuals] for no longer than is necessary for the purposes for which the personal data are processed.
Integrity and Confidentiality	Data must be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures.
Accountability	Data Controllers shall be responsible for, and be able to demonstrate compliance with [these principles].

Other focusses for fair information practices include:

- Simplified consumer choice to reduce burden placed on consumers
- Recommendation to maintain comprehensive data management practices throughout the lifecycle
- A Do Not Track option
- Requirements for affirmative express consent
- Concerns regarding the data collection capabilities of large platform providers; transparency and clear privacy notices and policies
- Individuals' access to data
- Educating consumers about data privacy practices
- Privacy by Design

3.3 Online Data in an ethical context

- **Ownership of data:** Right to control own personal data
- **The right to be forgotten:** To have information about an individual erased from the web, particularly to adjust online reputation
- **Identity:** Identity must be correct, or one can opt for a private identity
- **Freedom of speech online:** Expression rather than online bullying or insulting

3.4 Risks of Unethical Data Handling Practices

Data should be measured against Data Quality dimensions such as accuracy and timeliness to ensure it is trustworthy. Unethical data practices include:

3.4.1 Timing

Omission or inclusion of data points based on timing, e.g., market manipulation of end-of-day stock price.

Chapter 2

3.4.2 Misleading Visualisations

e.g. changing scale, comparing facts without clarifying their relationship

3.4.3 Unclear definitions or Invalid Comparison

- Provide context which informs meaning
- Statistical “smoothing” or over-training of a statistical model

3.4.4 Bias: an inclination of outlook

- Data collection for pre-defined results
- Biased use of data collected
- Hunch and search
- Biased sampling methodology
- Context and Culture

3.4.5 Transforming and Integrating Data

- Limited knowledge of data’s origin and lineage
- Data of poor quality
- Unreliable metadata
- No documentation of data remediation history

3.4.6 Obfuscation / Redaction of Data

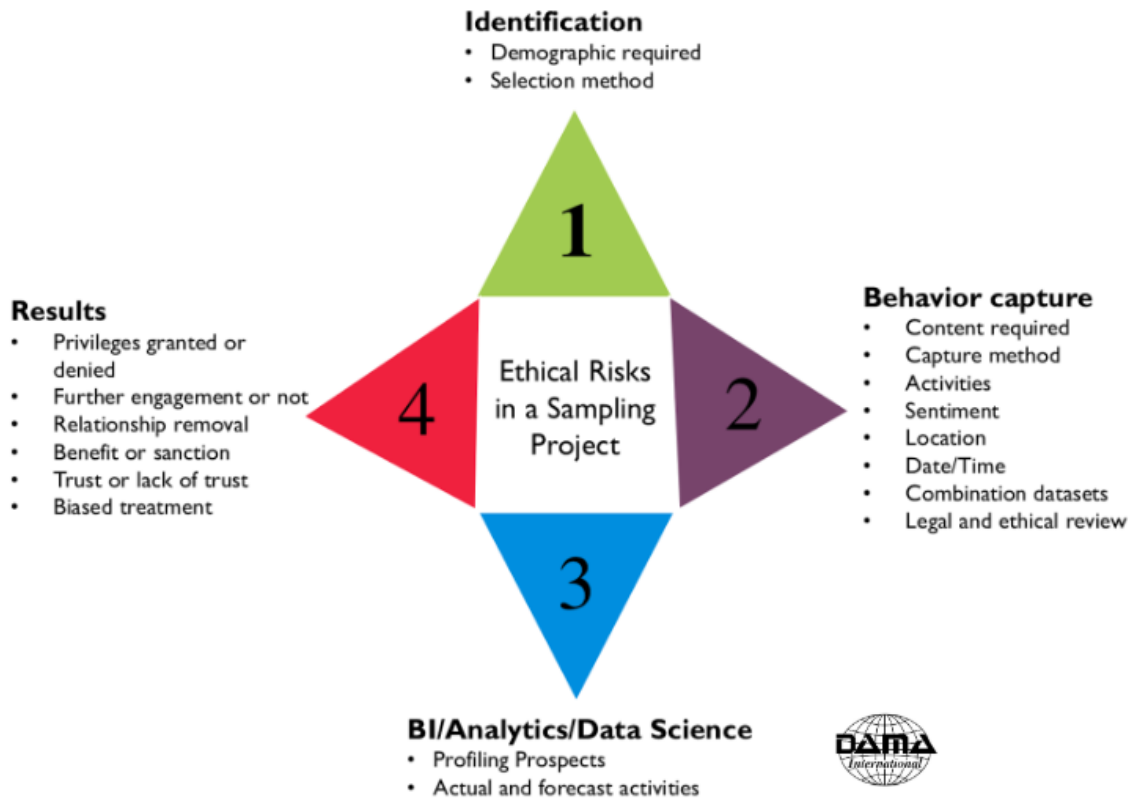
- Data aggregation
- Data Marking
- Data masking

3.5 Establishing an Ethical Data Culture

- **Review current state data handling practices:**
 - Understand degree of current practices to ethical and compliance drivers
 - Employees understanding of ethical implications of existing practices
- **Identify principles, practices and risk factors:**
- **Create an ethical data handling strategy and roadmap:** Ethical principles and expected behaviour related to data
 - **Values statements:** What the organisation believes in
 - **Ethical data handling principles:** How the organisation approaches challenges related to data
 - **Compliance framework:** Factors that drive organisational compliance obligations
 - **Risk assessment:** Likelihood of a specific problem occurring
 - **Training and communications:** Train on code of ethics. Communications must reach all employees
 - **Roadmap:** Timeline of activities that can be approved by management
 - **Approach to auditing and monitoring**
- **Adopt a socially responsible ethical risk model:** BI, analytics and data science responsible for data that describes:
 - Who people are (race, country or origin, religion)
 - What people do (Political, social, criminal activities)
 - How people live (Money, purchases, who they communicate with)
 - How people are treated (outcomes of analysis which privilege or prejudice them for future business)
- The risk model is used to determine:
 - Whether to execute the project
 - How to execute the project

- Actively identify potential risks
- Protect whistle-blowers
- Identify possible bias

3.5.1 Ethical Risk Model for Sampling Projects



Address all potential ethical risks in the areas of consideration, with a particular focus on negative effects on customers or citizens

3.6 Data Ethics and Governance

Oversight falls under Data Governance and Legal counsel:

- Keep up-to-date on legal changes
- Ensure employees are aware of their obligations
- Data Governance must set standards and policies to provide oversight of data handling practices
- Particularly to review plans and decisions proposed by BI, analytics and Data Science
- CDMP Certification requires subscription to code of ethics

Data Handling Ethics Checklist

Deliverables:

- Current Practices and Gaps
- Ethical Data Handling Strategy
- Communication Plan
- Ethics Training Program
- Ethical Corporate Statements on Data
- Awareness to Ethical Data Issues
- Aligned Incentives, KPIs, and Targets
- Updated Policies
- Ethical Data Handling Reporting

- Strategy
 - Current Practices & Gaps
 - Ethical Data Handling Strategy
- Organization
- Culture & Change
 - Communications Plan
- Working Methods
 - Ethics Training Plan
 - Ethical Principles, Policies, and Procedures
- Results
 - Ethical Corporate Statements on Data
 - Awareness to Ethical Data Issues
- Measurement
 - Aligned Incentives, KPIs, and Targets
 - Ethical Data Handling Reporting