CODE OF ETHICS
for Data-Based Value Creation

www.data-innovation.org

d●i
data innovation alliance

OVERVIEW

The "Code of Ethics for Data-Based Value Creation" consists of the following documents: 1) Overview; 2) Basics; 3) Recommendations; 4) Implementation; 5) Context.

The **BASICS** document provides information on the purpose and benefits of the Code, the basic ethical orientations, the relationship of the Code to data protection, and the structure of the recommendations. The Code is available in German, English, French, and Italian.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. PURPOSE OF THE CODE

The "Code of Ethics for Data-Based Value Creation" is intended for organisations that offer services or products based on data. The Code helps organisations develop products and services that are consistent with the ethical expectations of customers, employees or society. The Code can be used by any organisation that works with data. It gives large companies a common guideline. It also provides assistance to small and medium-sized enterprises (SMEs) that have little or no expertise in the ethical aspects of data use. The purpose of the Code is to systematically address the ethical issues that arise when creating or using data-based products and services. We define "data-based products and services" as any product or service that is produced by using data, data-derived models or data-generated knowledge as a necessary part of the value creation process. The Code provides recommendations on how these ethical issues can be specifically addressed.

1.2. WHY SHOULD A COMPANY USE THE CODE?

Data management is networked. The quality of companies' data-based products and services depends in many ways on other organisations. Accordingly, it is important that in this network of interdependence that companies can trust each other. If, for example, data is collected in an illegal manner, the company that acquires the data can also be caught in the maelstrom of a data scandal. If a company develops a forecasting tool that its customers use to discriminate against third parties, either unknowingly or negligently, it cannot completely escape the ethical responsibility for doing so.

Prudent companies therefore strive to reduce the risk of a negative reputation resulting from the violation of the ethical expectations of customers, employees or society. Ethically responsible companies can thus also improve their reputation as employers. This helps them attract employees who are cognisant of the ethical behaviour of their employer.

Data protection legislation already provides a basis for correct conduct. However, this does not cover all ethical issues, and ethical grey areas often remain – for example, when working with anonymised data, where data protection laws hardly apply. This Code helps companies to illuminate these grey areas and gain a better understanding of the ethical issues that can arise from the use of data.

Companies collectively following this Code will increase the confidence of consumers and policymakers in the value-added use of data. This is beneficial in the long term for all companies concerned, because this type of self-regulation is more effective and less costly than government legislation, which will increase in the absence of trust.

1.3. WHICH COMPANIES NEED THE CODE?

Not all companies require the same degree of ethical orientation in the development of data-based products and services. The following checklist assists a company to self-assess the extent to which this Code can be applied:

	Yes	No
1. The company creates and uses products or services that use data-based algorithms to make decisions that affect people's lives.		
2. The company provides third parties with forecasting or decision support systems that make decisions about people or support such decisions.		
3. The company uses forecasting or decision support systems supplied by third parties to make decisions about people.		
4. The company offers services widely used by the population, e.g., online services based on the processing of data (not necessarily personal data).		
5. The company uses data to provide services in socially and politically sensitive areas, e.g., journalism or healthcare.		
6. The company collects data from devices (e.g., Internet of Things) that enable comprehensive monitoring of individuals, regardless of whether this data can be attributed to individuals or not.		
7. The company collects, aggregates or sells data from other companies, which is then used to make decisions about individuals.		

A company that answers "yes" to at least one of these questions should consider this Code of Ethics to be relevant. The structure of the Code allows for a focus on the relevant sections depending on the company's product or service.

1.4. HOW SHOULD A COMPANY IMPLEMENT THIS CODE?

This Code consists of a group of documents. This modularity allows the Code to be used in a targeted manner and allows for individual in-depth study according to the needs of each company. The following documents form the Code:

1) **OVERVIEW:** The Overview serves as a simple orientation aid. Two posters show the most important basic orientations and the structure of the concrete recommendations along the data life cycle.

2) **BASICS:** This document sets out the principles of the Code. Section 1 explains the purpose of the Code and its conditions of application. Section 2 defines the six values on which the recommendations are based and indicates references to the applicable data protection law. Section 3 shows the relationship between these values and the activities of a data-based value creation process by means of a schematic representation of the data life cycle.

3) **RECOMMENDATIONS:** This document gives concrete recommendations for each step of the data lifecycle, structured according to the six values. Case studies illustrate how the recommendations can be implemented.

4) **IMPLEMENTATION:** This document explains how the recommendations of the Code can be incorporated into a company's business processes and describes several possible variants. In particular, it shows how to establish responsibilities for data governance in a company.

5) **CONTEXT:** This document explains the background to the ethical orientations, makes reference to the multitude of existing codes of ethics in the field of data and artificial intelligence, and refers to further literature for interested parties.

2. ETHICS FOR DATA-BASED VALUE CREATION

Data-based products and services create value for customers and companies. The business model is geared to this value creation. However, such products and services can also affect other personal or social values. The ethical responsibility of a company is to take these values into account and to introduce procedures that guarantee certain ethical standards.

2.1. BASIC ETHICAL ORIENTATIONS

In this Code we are guided by three basic ethical orientations, which stand for a number of values that regularly occur in the debate on data-based value creation:

1. HARM AVOIDANCE
You should not harm individuals or communities. This basic orientation includes the values of protection (e.g., against data loss), security (e.g., of data against hackers) and sustainability (i.e., minimising negative effects on the environment, e.g., through energy-efficient data processing).

2. JUSTICE
You should consider a fair distribution of benefits and burdens. This basic orientation includes, among other things, the values of equality (e.g., protection against discrimination), fairness (e.g., by giving something in return for collecting customer data) and solidarity (e.g., by making data available to the public for collective use).

3. AUTONOMY
You should enable individuals and communities to act in a self-determined manner. This basic orientation includes, among other things, the values of freedom (e.g., through freedom of choice in the configuration of digital services), privacy (e.g., by not collecting certain data) and dignity (e.g., through information practices taking the customer seriously).

The overarching goal of these basic orientations for data-based companies is to build and maintain trust in their products and services. Accordingly, these basic orientations should shape the business actions of every company. Being ethical means pursuing these values whenever possible. Ethical orientations can imply conflicting objectives; i.e., the realisation of one value can limit the pursuit of other values that are considered equally important: this should lead to a strategy for balancing or prioritising values for which we do not offer any guidance here.

[1] The expression "a value" is equivocal, and these basic orientations can also be understood as values in one of the senses of this expression. With the term "basic orientation" we express that each of them can be understood as a "cluster" of values, which aim in the same direction.

2.2. PROCEDURAL VALUES

In everyday business life, these basic orientations appear as abstract goals and it is often unclear how they are compatible with the concrete business processes. For this reason, three procedural values form important links between the basic ethical orientations and the recommendations that are intended to ensure alignment with them:
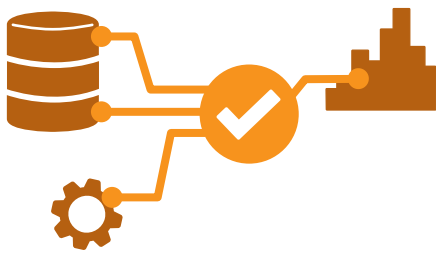
1. CONTROL

You should ensure that the internal processes surrounding the handling of data are well defined and controllable. This includes the need to create knowledge about the processes so that a control can be effective.

2. TRANSPARENCY

You should document and communicate what happens to data and how it is done. The focus of transparency is both the customer and, for example, an auditor; the concrete requirements for transparency differ according to these target groups.

3. ACCOUNTABILITY

You should define clear responsibilities for the handling of data and take responsibility in case of violations of rules. This is particularly intended to counteract the tendency for responsibilities to become blurred and unclear in the course of the digitisation of processes.

These procedural values do not deal with the question of selecting ethically proper goals for one's business, but rather how to achieve these goals. They are an expression of common corporate principles for ensuring the quality of products and services and the control of corporate risks. Accordingly, ethical requirements from the basic ethical orientations always imply corresponding requirements for control, transparency and accountability in order to ensure implementation in concrete practice. They should be used in addition to the three basic orientations to derive concrete recommendations.

The basic orientations of harm avoidance, justice and autonomy complement the "natural" corporate goals; i.e., to meet the needs of the customers and to generate profit for the company. An ethical evaluation of data-based value creation is based on the degree to which both these basic orientations and the procedural values are achieved.

Background information on the basic orientations and the procedural values can be found in the accompanying "Context" document. This document also provides an overview of the various data ethics guidelines published to date by international organisations.

2.3. THE ROLE OF DATA PROTECTION LAW

The principles of data protection law, which must of course be respected, also contribute to promoting these basic orientations. Data protection law primarily refers to the processing of personal data, i.e., all information relating to an identified or identifiable person. A person is "identifiable" if his or her identity can be established from the data itself, from the context or in conjunction with other data without disproportionate effort. "Processing" means any handling of personal data, irrespective of the means or procedures used, in particular the obtaining, storage, use, processing, disclosure, archiving, or destruction of data.

These principles and the requirements for data-based value creation derived from them are as follows:
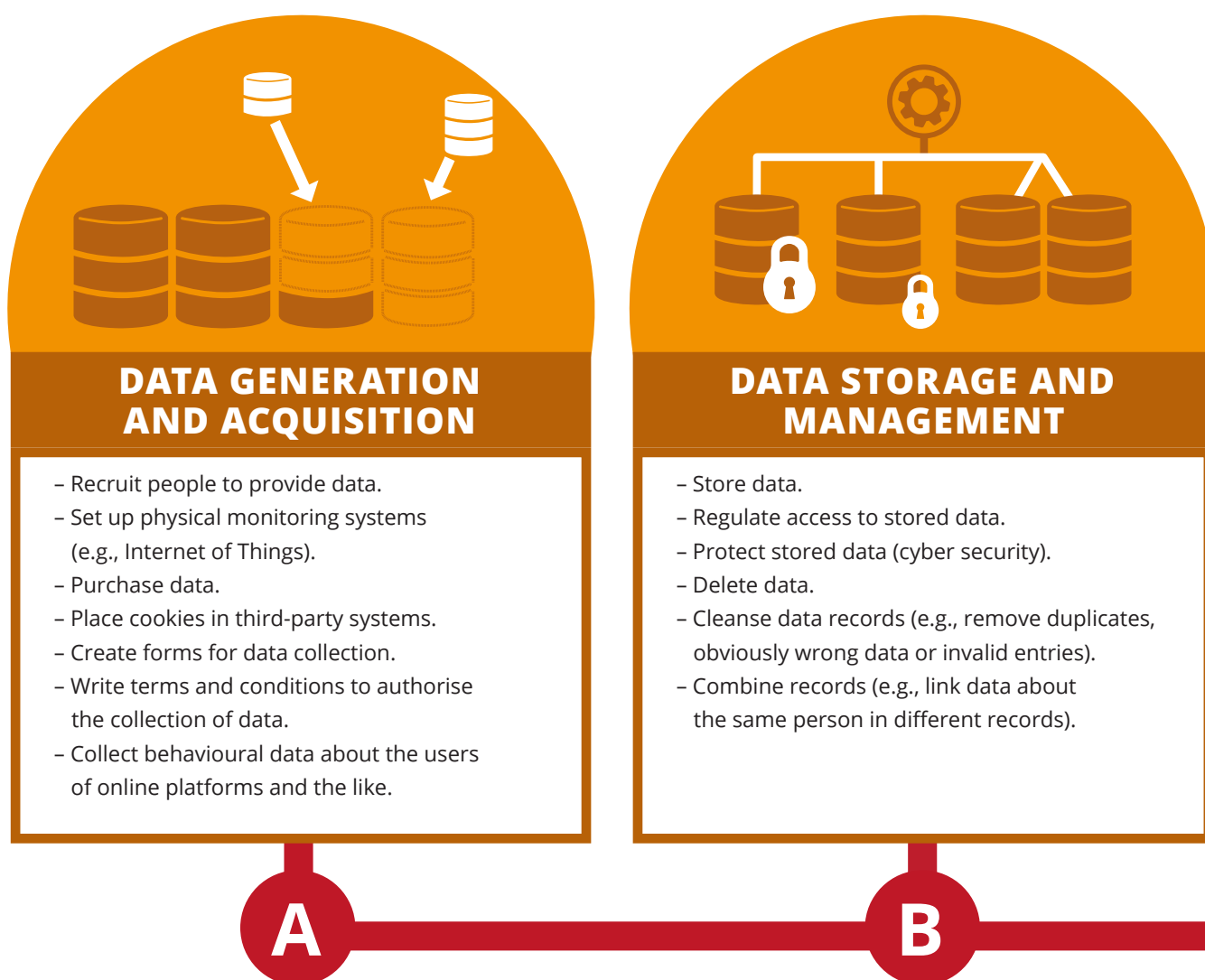
1) **LAWFULNESS:** Personal data must be processed lawfully. This can be achieved by obtaining informed consent, i.e., the person must give his or her consent to the collection of data after being adequately informed about its purpose and this consent may be revoked. In certain cases, the data processing institution may claim an overriding interest, but this must be regulated by law (e.g., the collection of data in the event of epidemics).

2) **RECOGNISABILITY:** It must be visible or evident for the person that his or her data is being collected. This principle is closely linked to the procedural value of transparency. Accordingly, the transparency-related recommendations can also be justified by data protection law. For example, individuals must be informed when their data is used to create a profile.

3) **PURPOSE LIMITATION:** The purpose of the data collection must be evident to the data subject and the data must be processed only for the specified or foreseeable purposes. The recommendations of this Code on the procedural value of "control" often also serve to implement this principle in practice.

4) **PROPORTIONALITY:** The data processing must be suitable and necessary for achieving the intended purpose and must be reasonable for the data subjects. In the EU, the General Data Protection Regulation is largely in line with the principle of data minimisation, according to which only data necessary for the defined purpose may be collected and only stored for the time necessary for this purpose.

This Code does not contain any specific data protection guidelines. Some recommendations, however, can be understood as a concretisation of data protection principles.

---

[2] These general principles are largely identical in Swiss and EU data protection law. For differences in the concrete implementation, we refer to the relevant literature. It should also be noted that Swiss data protection law is currently being revised with the aim of aligning it with the new basic data protection regulation of the European Union.
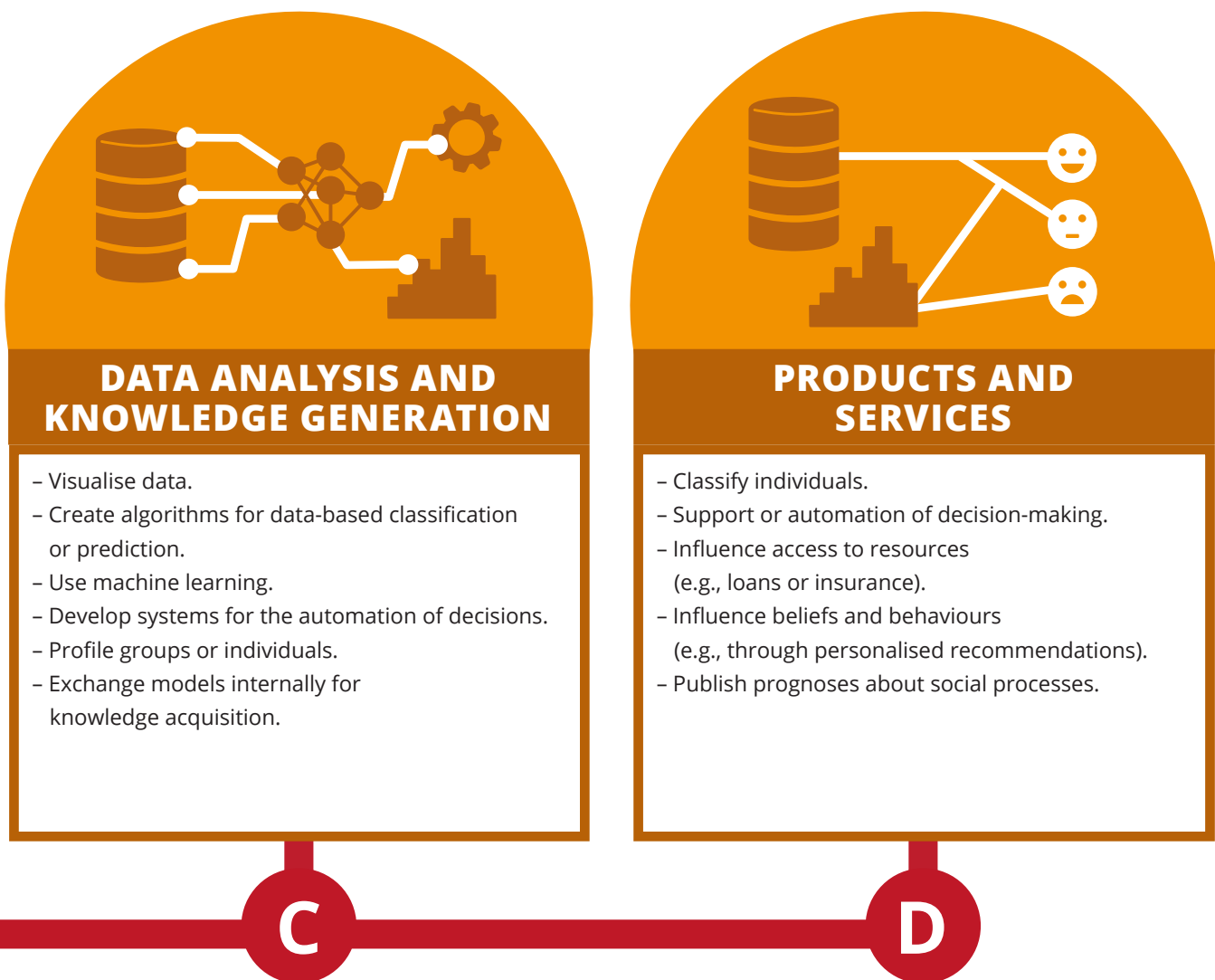
3. STRUCTURING OF DATA-BASED VALUE GENERATION

This Code focuses on ethical issues related to the use of data. The ethical recommendations are structured by the four main steps of the data life cycle, each step being characterised by specific ethical challenges. In the cycle, the activities in each step build on the results of previous steps, but feedback loops are both common and necessary, but are not further discussed.

DATA GENERATION AND ACQUISITION

– Recruit people to provide data.
– Set up physical monitoring systems (e.g., Internet of Things).
– Purchase data.
– Place cookies in third-party systems.
– Create forms for data collection.
– Write terms and conditions to authorise the collection of data.
– Collect behavioural data about the users of online platforms and the like.

DATA STORAGE AND MANAGEMENT

– Store data.
– Regulate access to stored data.
– Protect stored data (cyber security).
– Delete data.
– Cleanse data records (e.g., remove duplicates, obviously wrong data or invalid entries).
– Combine records (e.g., link data about the same person in different records).

A

B

A The result of step 1 is digitised data, which can be enriched with metadata. This can be factual data or personal data.

B The result of step 2 is a database that includes access rules and appropriate security mechanisms.

The ethical focus will usually shift as the individual steps are taken. For example, if personal data are processed in the first and second steps, many ethical questions arise. For the other steps, however, all data are often relevant, including anonymous personal data or factual data. An important ethical focus here is the possibility that such data may be erroneous, leading to incorrect insights or models, the use of which can then lead to ethically problematic consequences (e.g., discrimination against individuals). In step 4, therefore, the potential for abuse of the data-based products and services is also an ethical focus.

DATA ANALYSIS AND KNOWLEDGE GENERATION

– Visualise data.
– Create algorithms for data-based classification
  or prediction.
– Use machine learning.
– Develop systems for the automation of decisions.
– Profile groups or individuals.
– Exchange models internally for
  knowledge acquisition.

PRODUCTS AND SERVICES

– Classify individuals.
– Support or automation of decision-making.
– Influence access to resources
  (e.g., loans or insurance).
– Influence beliefs and behaviours
  (e.g., through personalised recommendations).
– Publish prognoses about social processes.

C

D

C The result of step 3 is a data product. We use this term to describe any mechanism for generating knowledge from the database or this knowledge itself with the aim of enabling data-based value creation.

D The result of step 4 is an effect of a data product. This term is used to describe the effects of data-based products or services on the real world outside the company.