

## Ethics

### Assignment:

- Pick 2 Ethical Models
- Compare them for similarities and differences.
- Which do you prefer and why?
- Which one is more suited for data ethics and why?

I have chosen for the **Product Impact Tool**, together with the **Ethical Readiness Check** as my first “model” of choice. With my focus on the Ethical Readiness Check.

For my second “model” I have found a **Data Ethics Framework**, which is a specific framework for incorporating Ethics in Data Science work.

## Ethical Readiness Check

“The Ethical readiness Check consists of the following two basic questions in terms of means and ends of a technical innovation:

- Is the means for a good end?
- Is it the good means for the end?

These two questions can be unfolded and developed. For example: What is the goal? Are there conflicting goals and values? May there be alternate goals, a double agenda? Is the technical means effective and fitting for the goal? What are the actual effects, also side effects, regardless the intentions?

This last question is about the effects of technology as is the topic of the Product Impact Tool. The Ethical Readiness Check with its series of questions about means and ends can therefore be seen as an simple entry to **The Product Impact Tool**, an entry moreover with a focus on the connection between impact and ethics.” (Eggink, Dorrestijn, van der Heijden, & Ouwens, 2022)

These days the ethical readiness check has more than two questions, it has been updated since 2022, see [Ethical Readiness Check](#).

The following questions are currently used (in Dutch):

- Dient de innovatie een goed doel, of is het bedenkelijk of omstreden?
- Is het doel eenduidig, of zijn er conflicterende waarden?
- Zijn er (heimelijke) bijbedoelingen?
- Werkt het middel voor het beoogde doel?
- Kan het middel worden misbruikt, voor oneigenlijk doelen?
- Welke onbedoelde (neven)effecten kan het middel hebben?

It is good to know that these questions have to be revisited during a Data Science artefact life cycle, as for instance in CRISP-DM. At minimum, at the start during Business Understanding, somewhere between Modelling and Evaluation and in the end just before Deployment.

## Data Ethics Framework

For my second choice I have whent with: **“A framework for managing ethics in data science projects”** (Sudhaman, Panigrahi, & Subramanian, 2023).

“Data ethics is the study and promotion of ethical practices with regards to data (including its creation, use, sharing, dissemination, processing, curation, and collection), algorithms (such as robotics, deep learning, machine learning, intelligent agents, and artificial intelligence), and corresponding practices (such as professional codes, hacking, programming, and responsible innovation, for instance, right conduct or right values<sup>6</sup>);). Thus, data science ethics can be grouped under the ethics of data, algorithms, and practices.<sup>6</sup> In this paper, we focus primarily on data ethics (data collection, building a data model, its evaluation, and deployment).”

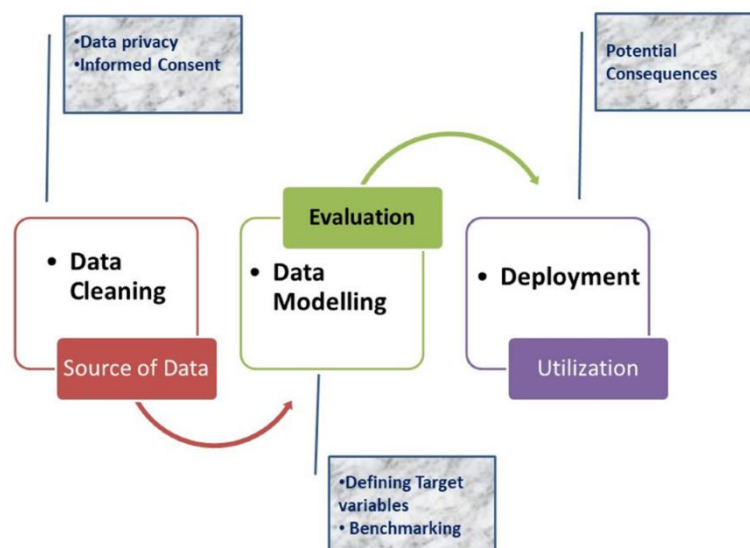


Figure 1 Data Ethics Framework

“The information presented in Figure 1 allows one to observe that there are three primary stages involved in most data science projects, namely data cleansing, data modeling, and model deployment. The data collection process is the first and foremost step for data scientists. All else being equal, the quality of your data will determine the accuracy and reliability of your analyses. Data cleaning is an essential first step in fostering an organization-wide culture that values fact-based decision-making. All else being equal, the quality of your data will determine the accuracy and reliability of your analyses.”

The Data Ethics Framework, is a framework with prescribed ethical considerations in a data science project along the lines of CRISP-DM. With a highlight on Data Preparation, Modelling and Deployment.

## Similarities & Differences

### Similarities

- Both ethics models have a connection with CRISP-DM phases, in a sense that both frameworks prescribe that during a data science project cycle, ethics should be considered in several phases.
- Both ethics models are meant for the IT industry.

### Differences

- The Ethical Readiness Check is a list of questions to activate ethical awareness and incorporate ethical points of view, to take into consideration when making decisions within the data science project.
- The Ethical Readiness Check is more abstract than the more detailed focus points prescribed by the Data Ethics Framework.
- The Data Ethics Framework is much more Data Science oriented than the Ethical Readiness Check.

### Preference

My personal preference goes to the Ethical Readiness Check, it's abstract form lets me incorporate ethical decision making in any project. I also like the maturity concept behind it, that ethical awareness can grow alongside the growth of your data science projects advancement. Meaning new insights from the project could alter an earlier ethical understanding. Perhaps new opportunities or threats are discovered and/or better understood at a later stage of the project. Although I do like the detailed CRISP-DM additions of the Data Ethics Framework.

### Best fit

Honestly having read upon both ethics models, I don't see a problem in combining both. I propose the following setup for Ethics in CRISP-DM:

CRISP-DM	Ethical Readiness Check	Data Ethics Framework
Business Understanding	Check 1	
Data Understanding		
Data Preparation		Data Privacy Informed Consent
Modelling	Check 2	Defining Target Variables Benchmarking
Evaluation		
Deployment	Check 3	Potential Consequences