

Data Ethics Canvas

Identify possible ethics related issues & actions needed to mitigate them

1. INTENT

Who does the system you build affect?

- Intended audience?
- Externalities?
- How much uncertainty can be tolerated in terms of different stakeholders?
- Who will be affected and how by false positives/negatives?

Positive effects?

- For whom?
- How to increase them
- How to monitor/measure?

Possible negative effects?

- For whom?
- How to minimize/prevent?
- How to monitor/measure?

What are the expected results of the system you are designing/building?
What is the state you will be measuring against?

2. DATA

What are your data sources?
Why are you using this data?

Where is the data from? Who are represented? What are its limitations? (for example over/underrepresentation, bias, gaps...)

Will the system create a feedback loop that affects the distribution of data that will be collected in the future?

Does your data have sensitive characteristics that should be taken into account?

Are there any fields that should be eliminated from the data?

Can you use anonymization or pseudonymization techniques to avoid needless evaluation or processing of individual data?

Would this use of data surprise the data subject?
Why?

Does the context of collection match the context of use?

Did this data come from a system prone to human error?

Was participation of the subjects voluntary? Was it influenced by incentive or quota system?

Data Ethics Canvas

Identify possible ethics related issues & actions needed to mitigate them

3. TRANSPARENCY & ACCOUNTABILITY

<p>Is explainability needed? Who needs it? How much are you able to explain your system?</p>	<p>Does this system need to be audited? By whom?</p>	<p>Who will be responsible and contacted if this system/product/service is seen causing harm?</p> <p>Who in this project resolves ethical issues?</p>	<p>Are there legal issues, regulations or policies that should be taken into account when using this data?</p> <p>How do we make sure users can exercise their data rights in meaningful way?</p>

4. SAFETY & MONITORING

<p>Can the system be "played"?</p>	<p>How are you ensuring data safety?</p>	<p>Is it possible for malicious actors to intentionally input bad data in the system?</p>	<p>How will the system be monitored after initial completion?</p>

5. ACTIONS

<p>Based on going through this canvas, what are the possible unintended consequences of this system and how severe are they?</p>	<p>Is there a clear red flag that needs to be escalated immediately?</p>	<p>What actions are you taking based on the issues found?</p> <p>Make sure someone is named responsible of all of them.</p>	<p>Additional findings</p>