AI ethics

Questions Data Ethics Board for *high-risk AI systems* on the principle of «Fairness»

December 2025

swisscom

Ethical principles as the basis for the decisions of the Data Ethics Board

The six principles concretise Swisscom's ethics framework in the area of data and AI

Fairness

Benefits and added value

We utilize digital resources and data in such a way that added value, and benefits are created for customers and society. We apply our framework of values unrestrictedly in the use of digital technologies.

Transparency

We enable our customers and the public to fully understand our use of digital technologies and the processing of data, as well as the risks of the most important use cases.

.

Responsibility / Accountability

We assume full responsibility for the use of digital technologies and the processing of data and the results or consequences thereof by Swisscom or third parties acting on our behalf.

No discrimination

We prevent discrimination and bias. We ensure that no one is disadvantaged or defamed because of particular characteristics such as gender or skin colour.

Informal self-determination

We protect our customers and the public from unlimited processing of their data. We give data subjects the opportunity to determine the processing of their data themselves without incurring any disadvantages in their relationship with Swisscom.

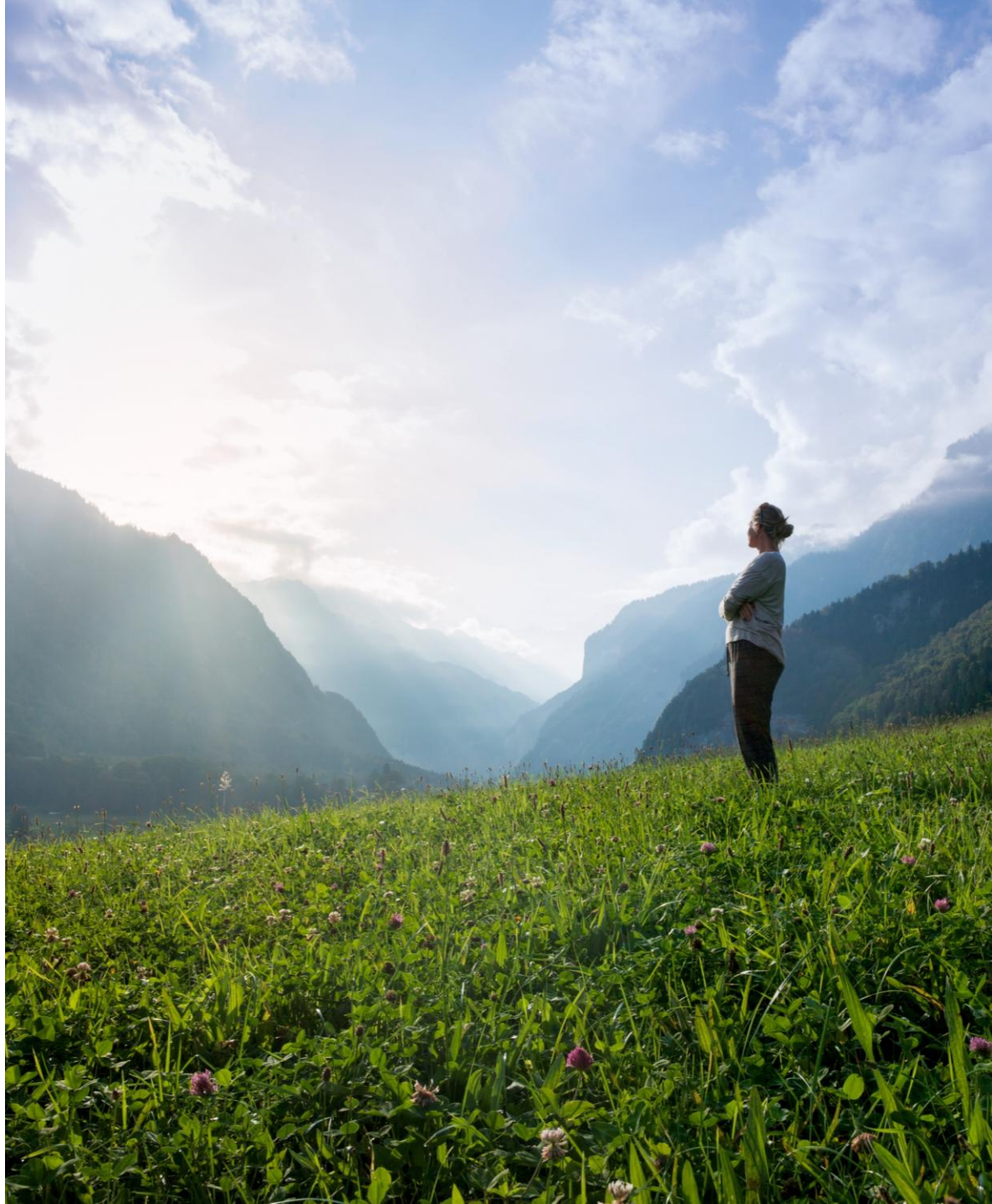
Respect for personality

We respect people's personality and privacy. We avoid interfering with or damaging physical and psychological integrity. We respect each individual’s right to their own image and voice.

Fairness

Benefits and added value
We utilise digital resources and data in such a way that added value and benefits are created for customers and society. We apply our framework of values unrestrictedly in the use of digital technologies.

1. What is the intended use of the AI system or what problem does it solve?
2. What benefits and added value does the AI system create for employees, customers and society?
3. What considerations have been made regarding the environmental impact (energy and water consumption) of the AI system?

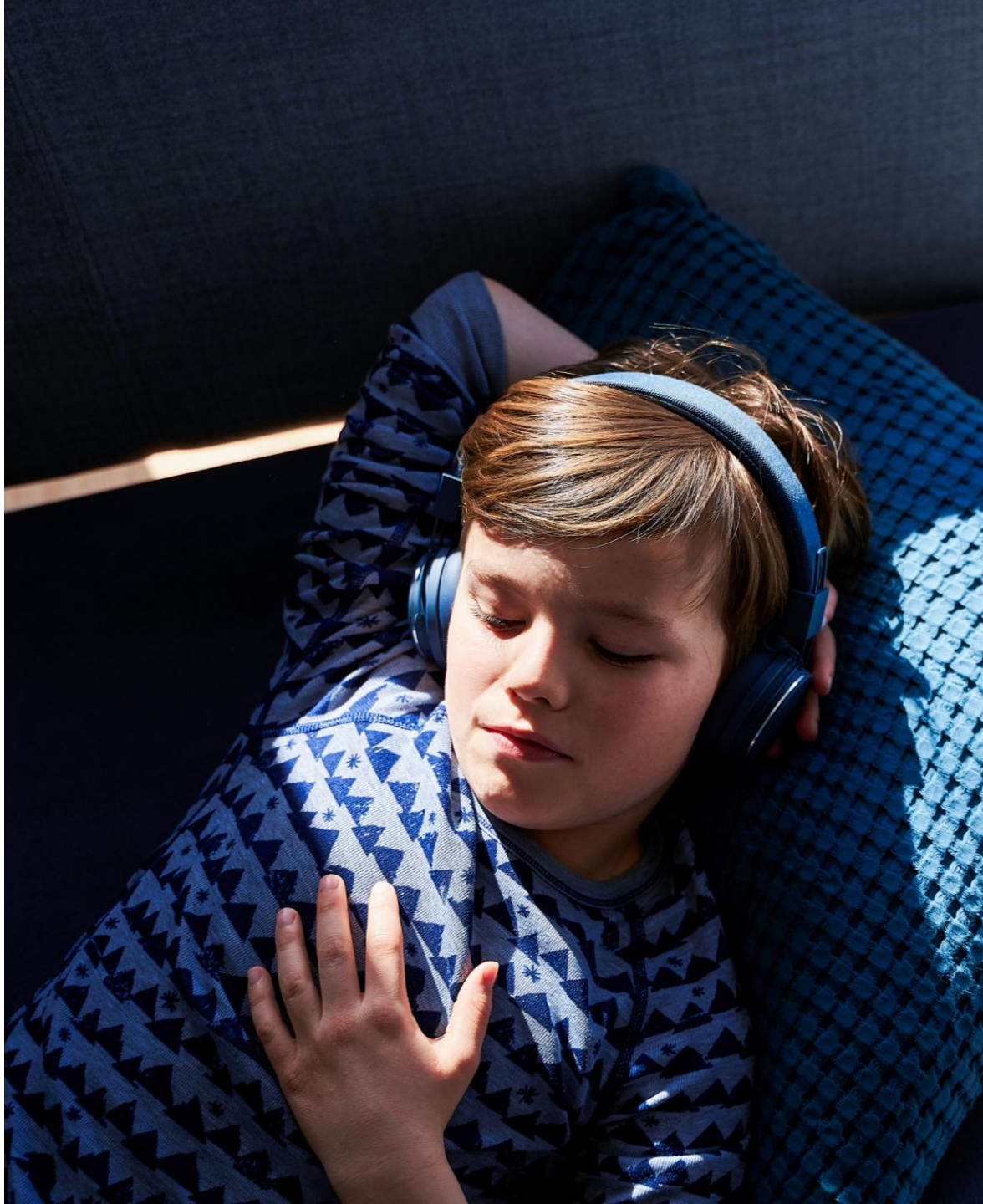


Fairness

No discrimination
We prevent discrimination and bias. We ensure that no one is disadvantaged or defamed because of particular characteristics such as gender or skin colour.

1. How was the AI system analysed (by the manufacturer or by Swisscom) for possible bias* in the predictions?
2. How is it ensured that the AI system does not create echo chambers and does not reinforce existing stereotypes?
3. What measures have been taken (by the manufacturer or Swisscom) to prevent modelling bias*?
4. What training data has been or will be used (by the manufacturer or Swisscom) for the model?
5. What measures have been taken (by the manufacturer or Swisscom) to ensure that this data is fair and equitable (bias-free)? If the AI system was purchased: Are there any additional measures taken by Swisscom?
6. What data does the AI system produce?

* possible reasons for bias see next page



Possible reasons for bias based on the data for training, validating and testing the AI system

[Source: Florent Thouvenin, Stephanie Volz; WHITE PAPER Discrimination, June 2024]

Historical bias: Data based on outdated, discriminatory realities was used. Social changes may have been neglected, thereby reproducing problems of the past.

Representation bias: An unbalanced data set was used that does not take equal account of all groups of people. This may be, for example, because less data is available for certain groups of people than for others.

Label: The data used was labelled unbalanced, e.g. personal data with protected characteristics was labelled negatively.

Feedback: In self-learning systems, a problematic feedback loop can occur if the results of the AI system are used as part of a new database for further training of the AI system. This can maintain or even reinforce discrimination.

Transfer context bias: The AI system is used in the wrong context. For example, if the algorithm was developed to assess a specific group of people, it can lead to inaccurate or incorrect results when applied to another group of people.

Aggregation bias: Conclusions are drawn from a standardised model to groups of people who should be considered differently.

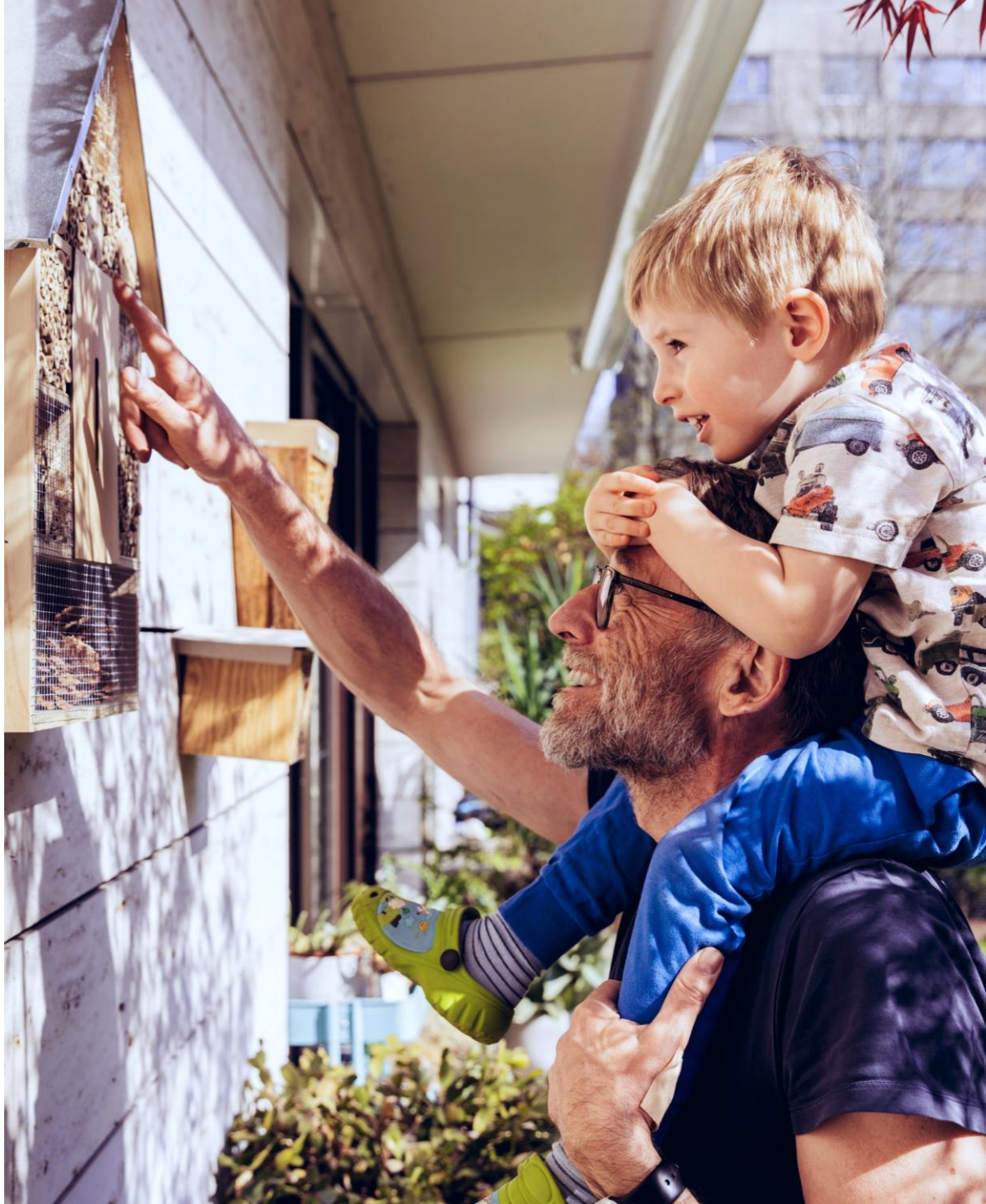
Automation bias: Too much trust is placed in the results of the AI system because the quality of the decisions is overestimated. For example, decisions are adopted by the AI system even though it only makes suggestions ("decision support system").

Fairness

Respect for personality
We respect people's personality and privacy. We avoid interfering with or damaging physical and psychological integrity. We respect each individual’s right to their own image and voice.

1. How is human control over the results of the AI system and its use intended?
2. What measures are in place to make it easier for users to interpret the results (relevant determinants)?
3. What effects does the AI system have on people in the context of its intended use and what effects could it have in the worst case?
4. What impact can the AI system have on people if it is used manipulatively** or incorrectly?
5. Have precautions been taken to prevent users from blindly trusting the AI system (e.g. critical review of a performance assessment)?

** Typical examples of manipulation are so-called "dark patterns", which are intended to induce users to take actions that are contrary to their own interests. To achieve this, users are deceived or otherwise significantly impaired in their ability to make free decisions. However, dark patterns are not a problem specific to AI systems.



Fairness

Informal self-determination
We protect our customers and the public from unlimited processing of their data. We give data subjects the opportunity to determine the processing of their data themselves without incurring any disadvantages in their relationship with Swisscom.

1. Is there an opt-out option for users who do not wish to interact with the AI system? What (preferably equivalent) alternative is offered to users?
2. Can users themselves determine whether and, if so, how their data is used (e.g. for AI training)? How user-friendly are these options? Can the selected preferences or consent be adjusted or revoked subsequently?
3. To what extent can the AI system be manipulative** (e.g. to collect more data than necessary)?

** Typical examples of manipulation are so-called "dark patterns", which are intended to induce users to take actions that are contrary to their own interests. To achieve this, users are deceived or otherwise significantly impaired in their ability to make free decisions. However, dark patterns are not a problem specific to AI systems.

Fairness

Transparency
We enable our customers and the public to fully understand our use of digital technologies and the processing of data, as well as the risks of the most important use cases.

1. How easy is it for (end) users to see and understand (e.g. appropriate and comprehensible information) that they are interacting with an AI system?
2. Question for providers: To what extent is the information provided to the operator/deployer (operating instructions) accessible and understandable?
3. Question for deployers: To what extent is the information provided to (end) users appropriate to the target group?
4. How is it ensured that the transparency requirements are regularly reviewed?

Fairness

Responsibility / Accountability

We assume full responsibility for the use of digital technologies and the processing of data and the results or consequences thereof by Swisscom or third parties acting on our behalf.

1. What is the potential for misuse of the AI system if it is used outside of its intended purpose?
2. How was the risk assessed that the AI system can be misled by carefully constructed inputs and thus achieve unwanted results?
3. Where can questions and comments from (end) users about the AI system be sent?
4. What instruction or training is provided for whom (e.g. end users) when using the AI system (AI literacy)?
5. What suitable processes and mechanisms have been set up (by the manufacturer or Swisscom) that take effect in the event of damage or adverse effects (incident process, replacement solution, communication, compensation)?
6. To what extent is a regular review of the AI system (by the manufacturer or Swisscom) ensured in terms of benefits, function, damage prevention, etc.?