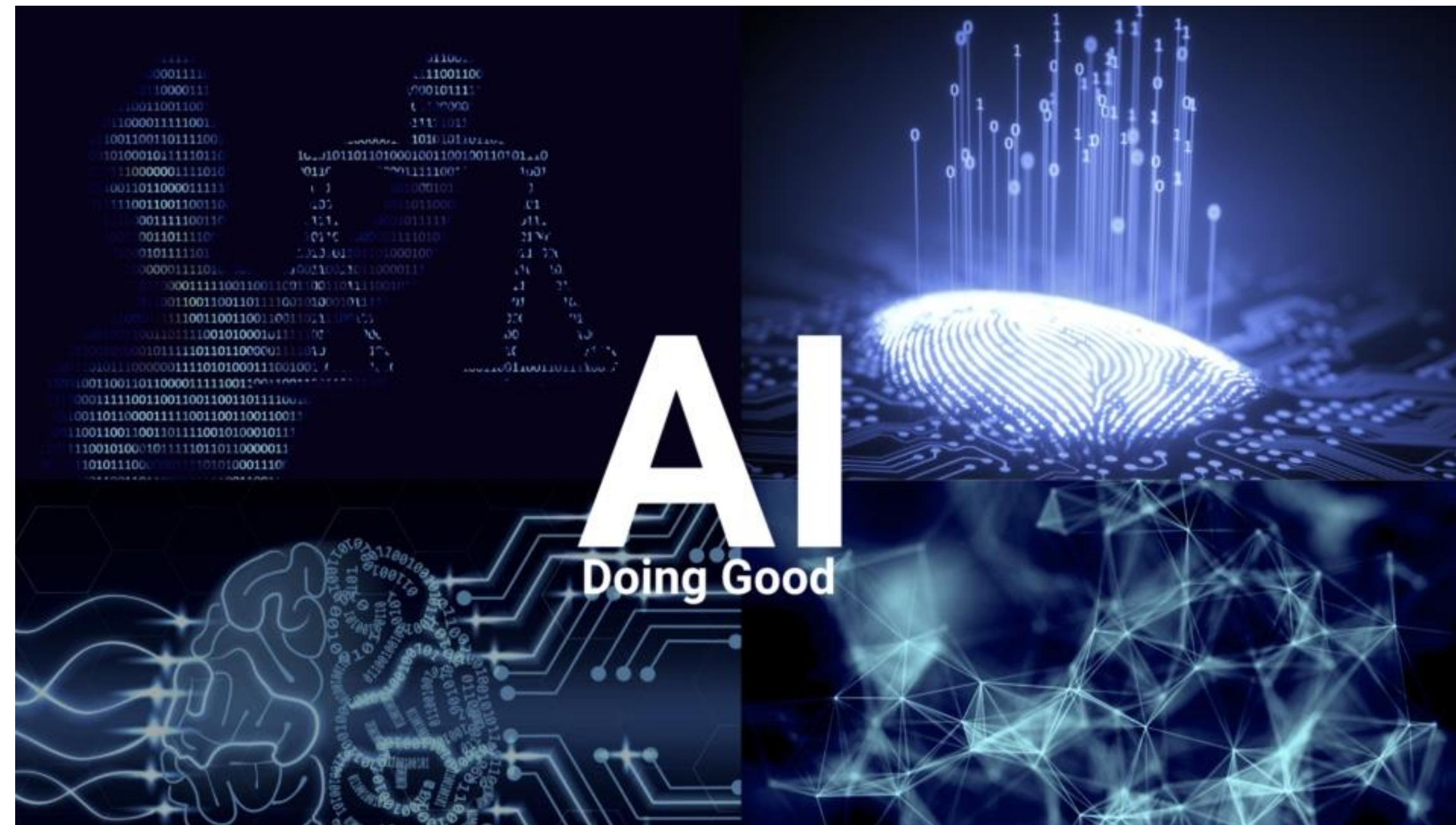


Etyka sztucznej inteligencji i geopolityka

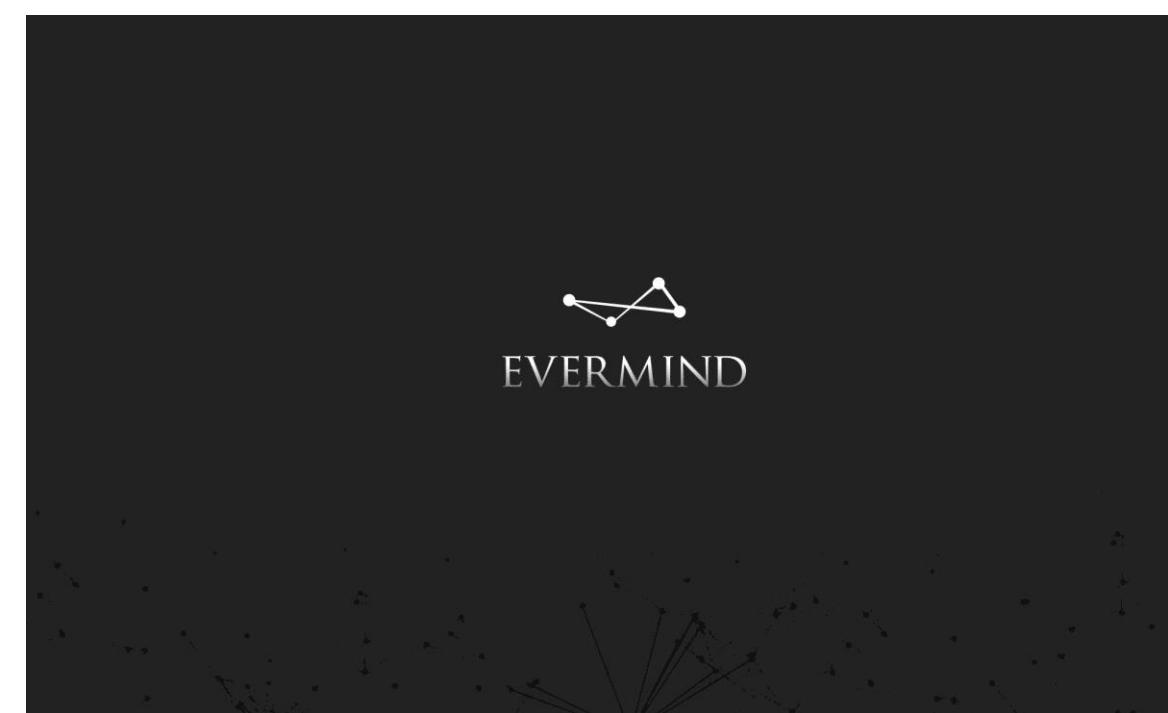


Ada Florentyna Pawlak
Uniwersytet Łódzki
Towarzystwo Etyki i Filozofii Techniki
adaflorentynapawlak@gmail.com



Towarzystwo Etyki i Filozofii Techniki

TEFT.pl



konferencja naukowa
Transhumanizm: 22/23.02.2023
Idee, Strategie, Wątpliwości



Infographic: Why addressing ethical questions in AI will benefit organizations, Capgemini, July 2019

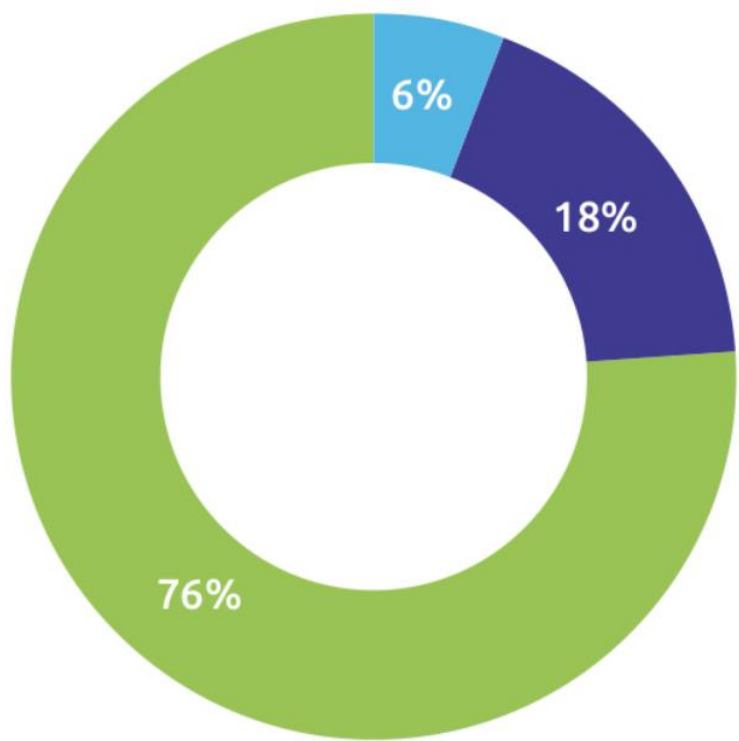


Why addressing **ethical**
questions in **AI** will benefit
organizations



Consumers want regulations on the use of AI

Do you think there should be a new law or regulation to regulate the use of AI by organizations?

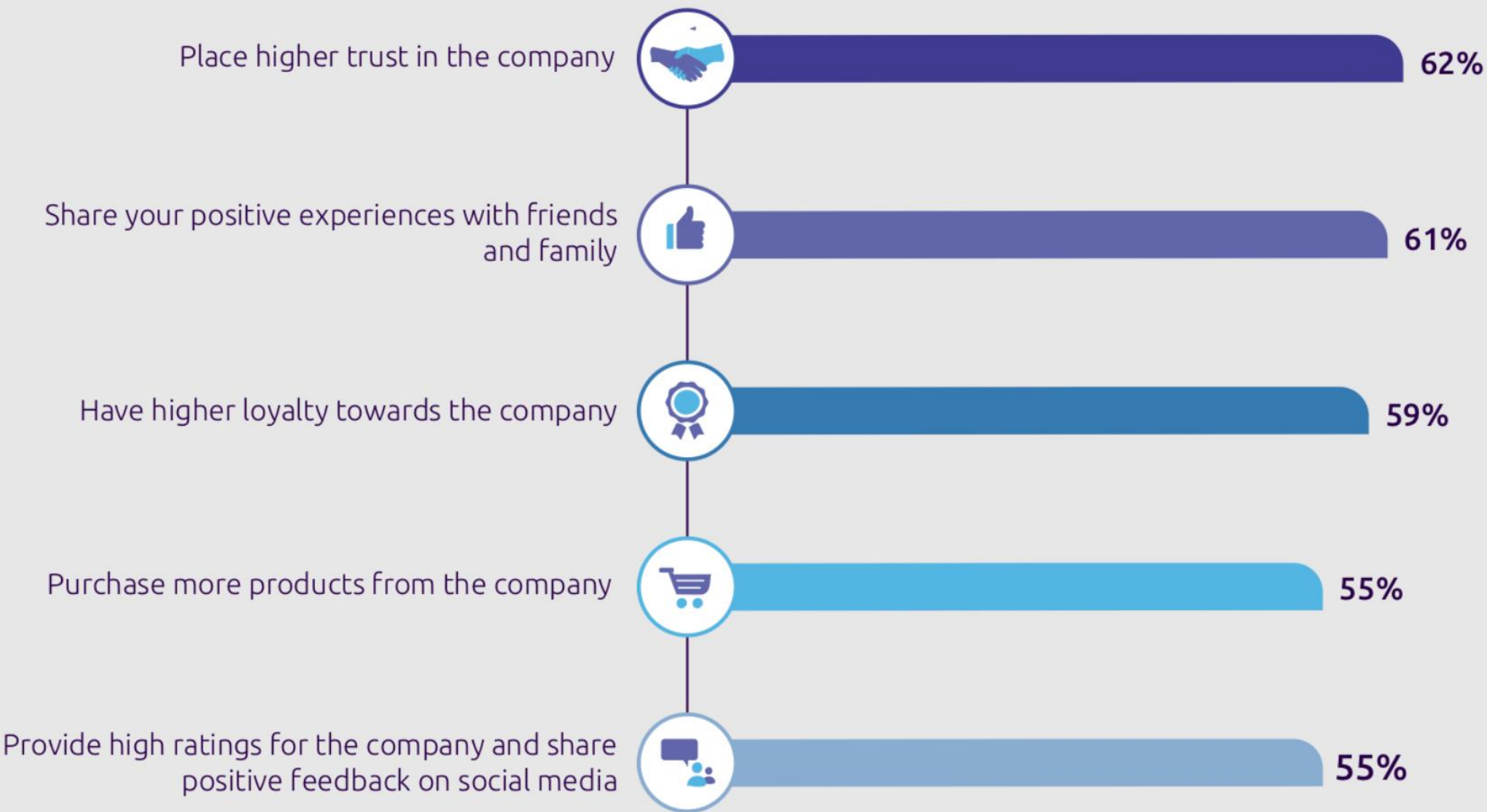


Disagree Maybe/don't know Agree

Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

Ethical AI interactions drive consumer trust and satisfaction

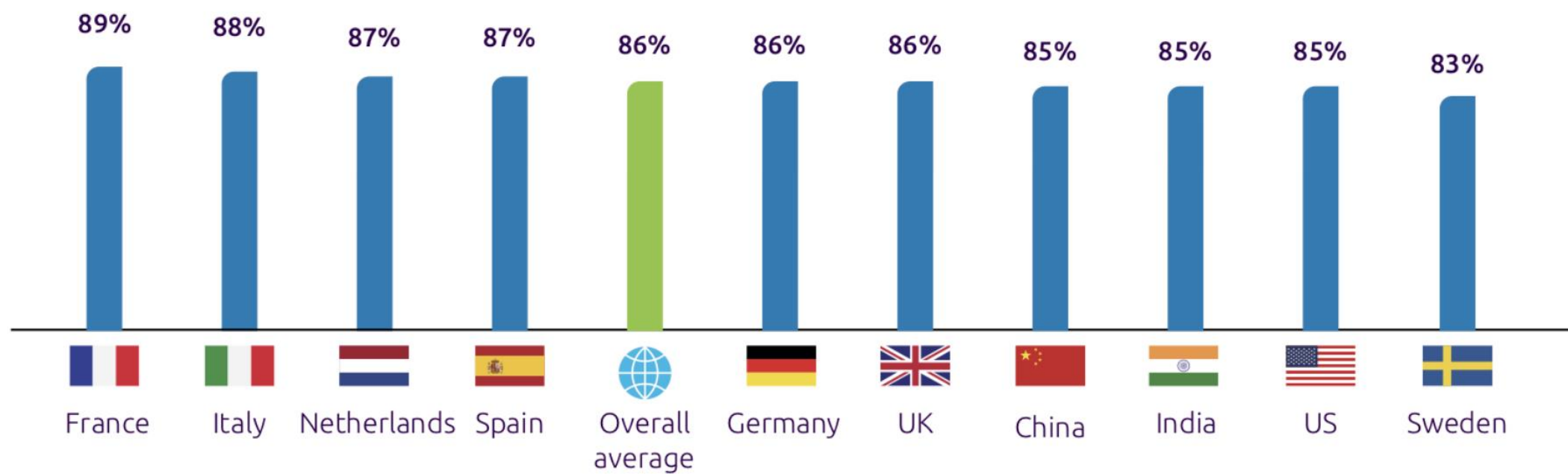
What would you do if you perceive the AI-enabled interaction to be ethical?



Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

Nearly nine in ten organizations across countries have encountered ethical issues resulting from the use of AI

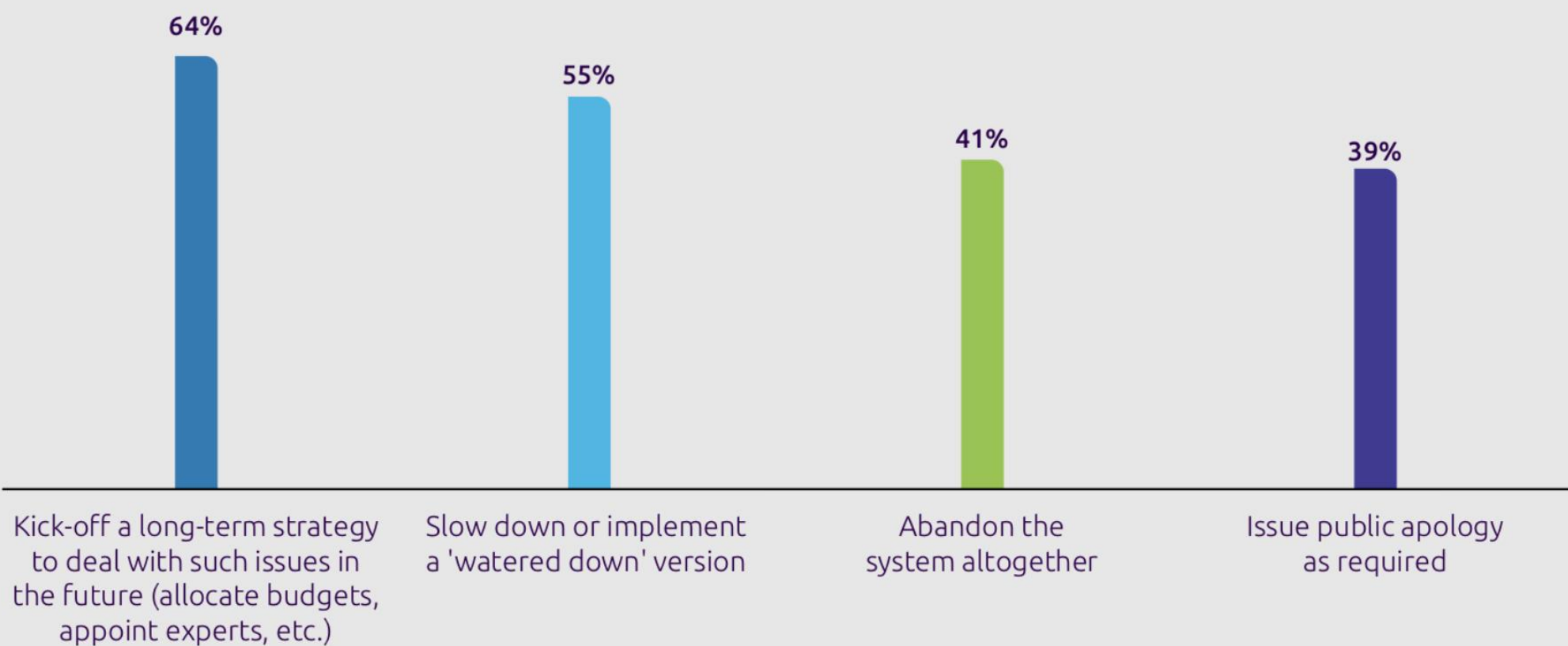
In the last 2-3 years, have the below issues resulting from the use and implementation of AI systems, been brought to your attention? (percentage of executives, by country)



We presented over 40 cases where ethical issues could arise from the use of AI, to executives across sectors. We asked them whether they encountered these issues in the last 2-3 years.
Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations

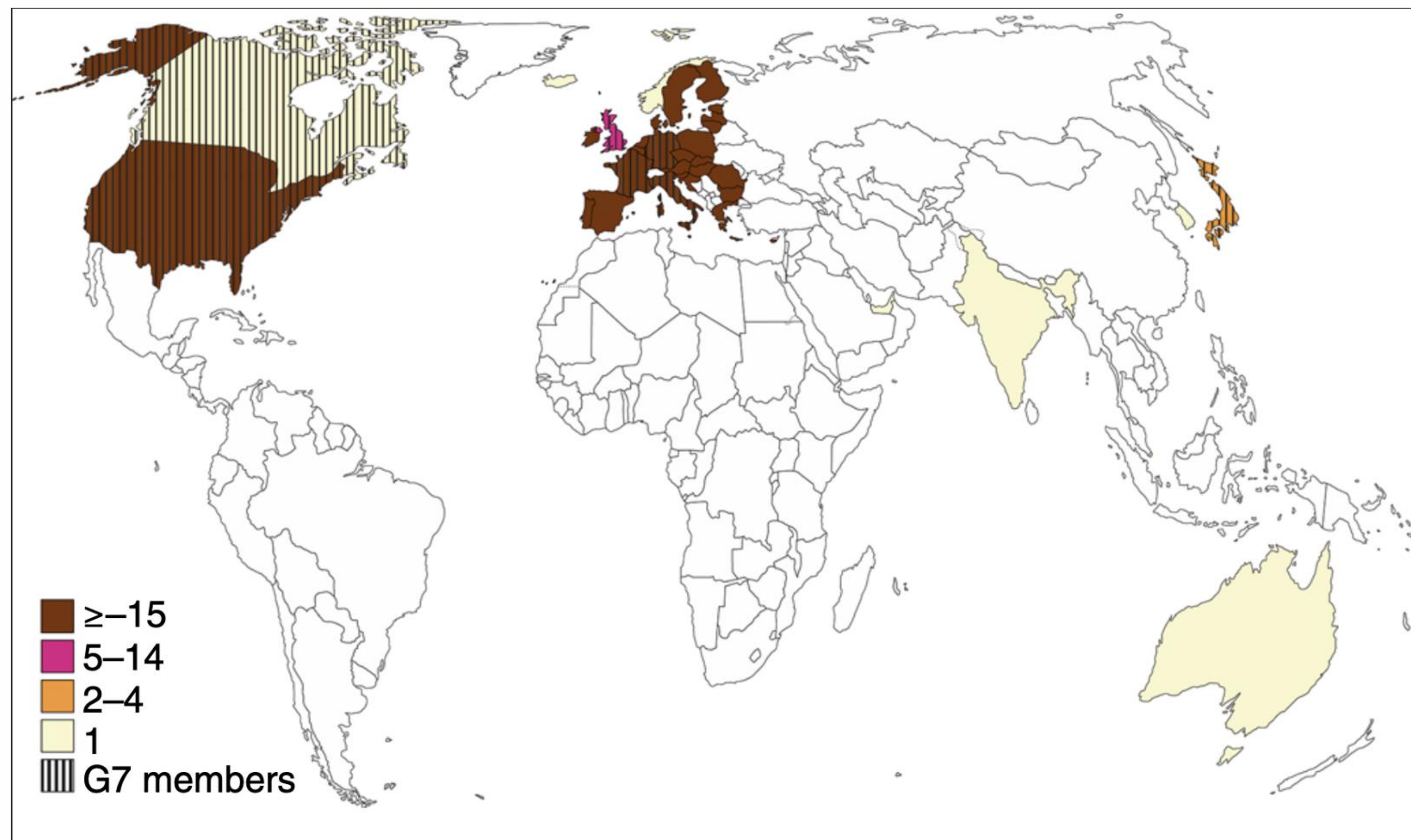
Executives are starting to realize the importance of ethical AI and are taking action when ethical issues are raised

What did your organization do/likely to do with the AI system when ethics related concerns were/are raised/brought to your notice?



Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations

Geografia standardów etycznych AI



Krajobraz standardów etycznych AI

84 dokumenty zawierające wskazówki etyczne dla AI

88% dokumentów ogłoszonych po 2016 r

- USA: 21
- UE: 19
- UK: 13
- Japonia: 4
- Kanada, Islandia, Norwegia, ZEA, Indie, Singapur, Korea Południowa, Australia: 1

Krajobraz standardów etycznych AI

Autorstwo dokumentów:

- firmy prywatne: **19** (22,6%)
- agencje rządowe: **18** (21,4%)
- instytucje akademickie i badawcze: **9** (10,7%)
- organizacje międzyrządowe i ponadnarodowe: **8** (9,5%)
- organizacje non-profit, zawodowe i naukowe: **7** (8,3%)
- organizacje sektora prywatnego: **4** (4,8%)
- zrzeszenia badawcze: **1** (1,2%)
- fundacje naukowe: **1** (1,2%)
- partie polityczne: **1** (1,2%)

Krajobraz standardów etycznych AI

Adresaci standardów etycznych

- wiele grup interesariuszy: **27** (32.1%)
- podmiot ogłaszający standardy: **24** (28.6%)
- sektor publiczny: **10** (11.9%)
- sektor prywatny: **5** (6.0%)
- sprecyzowane grupy interesariuszy np. organizacje deweloperów: **3** (3.6%)
- „organizacje”: **1** (1.2%)
- badacze: **1** (1.2%)
- adresaci niesprecyzowani: **13** (15.5%)

Global Ethics Code

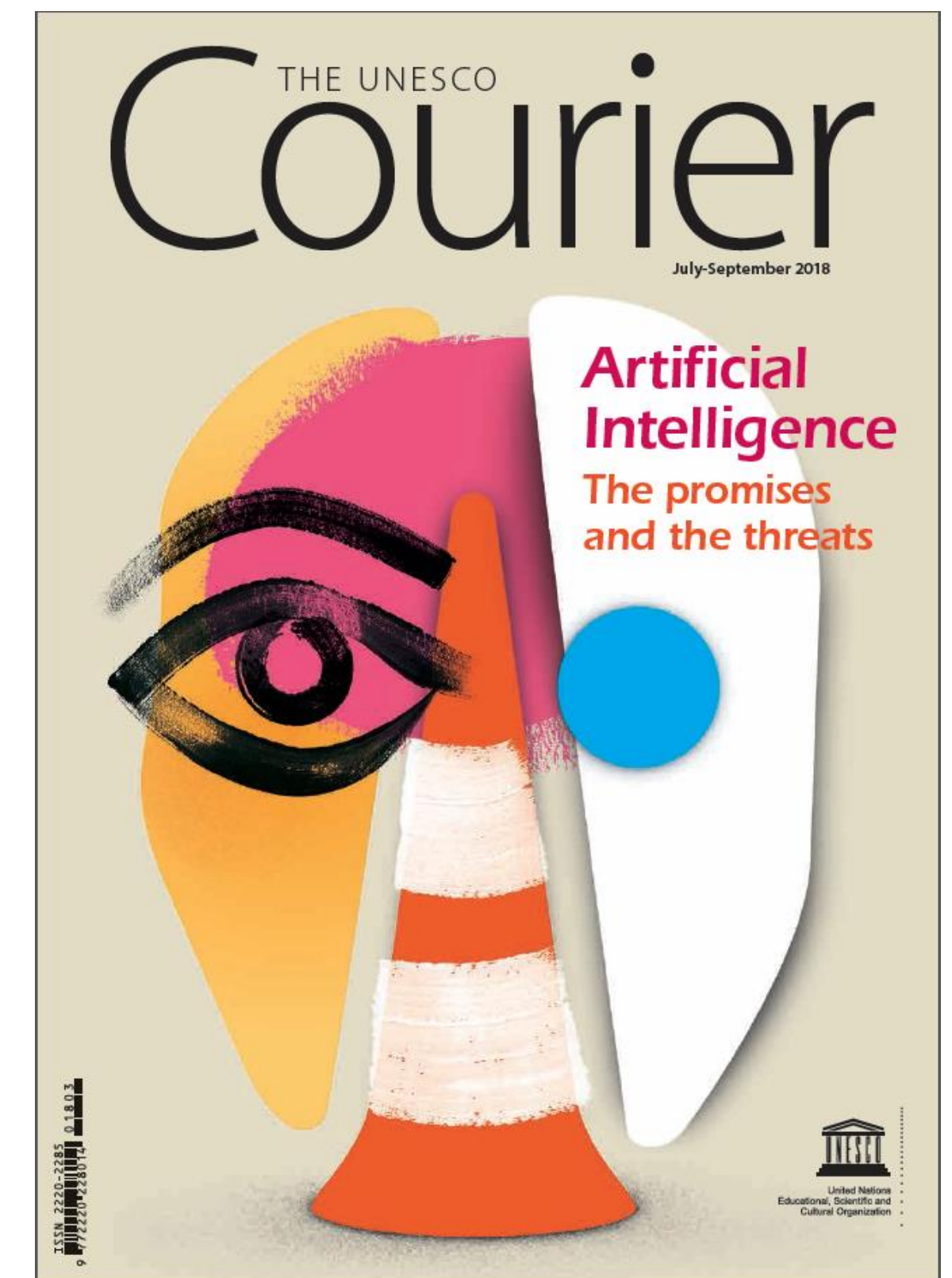
An international regulatory instrument is essential for the responsible development of AI, a task that UNESCO is in the process of undertaking.

„It is our responsibility to lead a universal and enlightened debate in order to enter this new era with our eyes wide open, without sacrificing our values, and to make it possible to establish a common global foundation of ethical principles.”

Director-General Audrey Azoulay



United Nations
Educational, Scientific and
Cultural Organization



OECD - AI GO

Sekcja 1: Zasady odpowiedzialnego zarządzania AI

- Rozwój AI dla dobrobytu i zrównoważonego rozwoju
- Koncentracja na wartościach i człowieku
- Transparentność
- Bezpieczeństwo
- Odpowiedzialność

Sekcja 2: Zasady dla decydentów na poziomie policy

- Inwestowanie w badania i rozwój
- Wspieranie ekosystemu dla rozwoju AI
- Kształtowanie środowiska politycznego
- Budowa kompetencji i przygotowanie rynku pracy
- Współpraca międzynarodowa



Standardy etyczne dla AI: przykłady

AI Now

- 5. **Work with professional organizations such as The Association for the Advancement of Artificial Intelligence (AAAI), The Association of Computing Machinery (ACM), and The Institute of Electrical and Electronics Engineers (IEEE) to update (or create) professional codes of ethics that better reflect the complexity of deploying AI and automated systems within social and economic domains.** Mirror these changes in education, making courses in civil rights, civil liberties, and ethics required training for anyone wishing to major in Computer Science. Similarly, update professional codes of ethics governing the professions into which AI systems are being introduced, such as those applied to doctors and hospital workers.



Standardy etyczne dla AI: przykłady

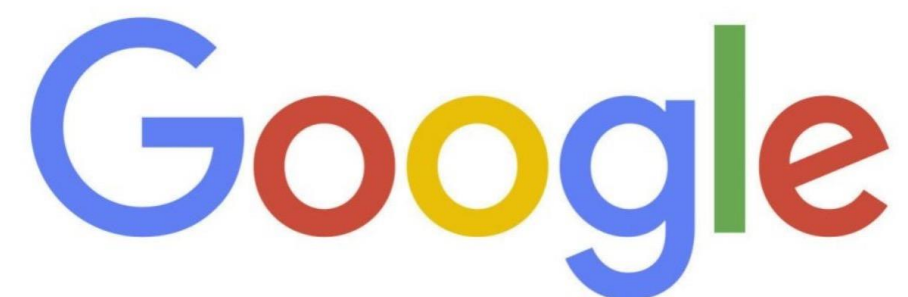
AI Now

Kiedy opublikowane: 2016 (aktualne)

Kto sformułował: zespół badawczy powołany przez AI Now

Adresat: „stakeholders engaged at various points in the production, use, governance, and assessment of AI systems”

Przez kogo i jak egzekwowane: wskazane wprost w zaleceniach



CENEO.pl

Standardy etyczne dla AI: przykłady

Google

We will assess AI applications in view of the following objectives.
We believe that AI should:

1. Be socially beneficial.
2. Avoid creating or reinforcing unfair bias.
3. Be built and tested for safety.
4. Be accountable to people.
5. Incorporate privacy design principles.
6. Uphold high standards of scientific excellence.
7. Be made available for uses that accord with these principles.

Standardy etyczne dla AI: przykłady

Microsoft

1. Fairness: AI systems should treat all people fairly
2. Inclusiveness: AI systems should empower everyone and engage people
3. Reliability & Safety: AI systems should perform reliably and safely
4. Transparency: AI systems should be understandable
5. Privacy & Security: AI systems should be secure and respect privacy
6. Accountability: AI systems should have algorithmic accountability

Standardy etyczne dla AI: przykłady

EPSRC

**Engineering and Physical Sciences
Research Council**

Engineering and Physical Sciences Research Council (EPSRC)

Standardy etyczne dla AI: przykłady

Engineering and Physical Sciences Research Council (EPSRC)

1. Robots are multi-use tools.

- Robots should not be designed solely or primarily to kill or harm humans, except in the interests of national security.

Kiedy opublikowane: 2010 (aktualne)

Kto sformułował: wskazani z imienia i nazwiska delegaci EPSRC

Adresat: „designers, builders and users of robots”

Przez kogo i jak egzekwowane: przez EPSRC przy przyznawaniu funduszy na badania

2. Humans, not robots, are responsible agents

- Robots should be designed; operated as far as is practicable to comply with existing laws & fundamental rights & freedoms, including privacy.

Standardy etyczne dla AI: przykłady

Engineering and Physical Sciences Research Council (EPSRC)

4	Robots are manufactured artefacts. They should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be transparent.	Robots are manufactured artefacts: the illusion of emotions and intent should not be used to exploit vulnerable users.	One of the great promises of robotics is that robot toys may give pleasure, comfort and even a form of companionship to people who are not able to care for pets, whether due to rules of their homes, physical capacity, time or money. However, once a user becomes attached to such a toy, it would be possible for manufacturers to claim the robot has needs or desires that could unfairly cost the owners or their families more money. The legal version of this rule was designed to say that although it is permissible and even sometimes desirable for a robot to sometimes give the impression of real intelligence, anyone who owns or interacts with a robot should be able to find out what it really is and perhaps what it was really manufactured to do. Robot intelligence is artificial, and we thought that the best way to protect consumers was to remind them of that by guaranteeing a way for them to "lift the curtain" (to use the metaphor from The Wizard of Oz). This was the most difficult law to express clearly and we spent a great deal of time debating the phrasing used. Achieving it in practice will need still more thought. Should all robots have visible bar-codes or similar? Should the user or owner (e.g. a parent who buys a robot for a child) always be able to look up a database or register where the robot's functionality is specified? See also rule 5 below.
---	--	--	---



Engineering and Physical Sciences
Research Council

Standardy etyczne dla AI: przykłady

Engineering and Physical Sciences Research Council (EPSRC)

Kiedy opublikowane: 2010 (aktualne)

Kto sformułował: wskazani z imienia i nazwiska delegaci EPSRC

Adresat: „designers, builders and users of robots”

Przez kogo i jak egzekwowane: przez EPSRC przy przyznawaniu funduszy na badania

The 2020 AI Strategy Landscape

50 National Artificial Intelligence strategies shaping the future of humanity

[Advanced Technology](#)[Artificial Intelligence](#)[Education in 2030](#)

Krajobraz standardów etycznych AI

Wartości wskazywane w standardach etycznych

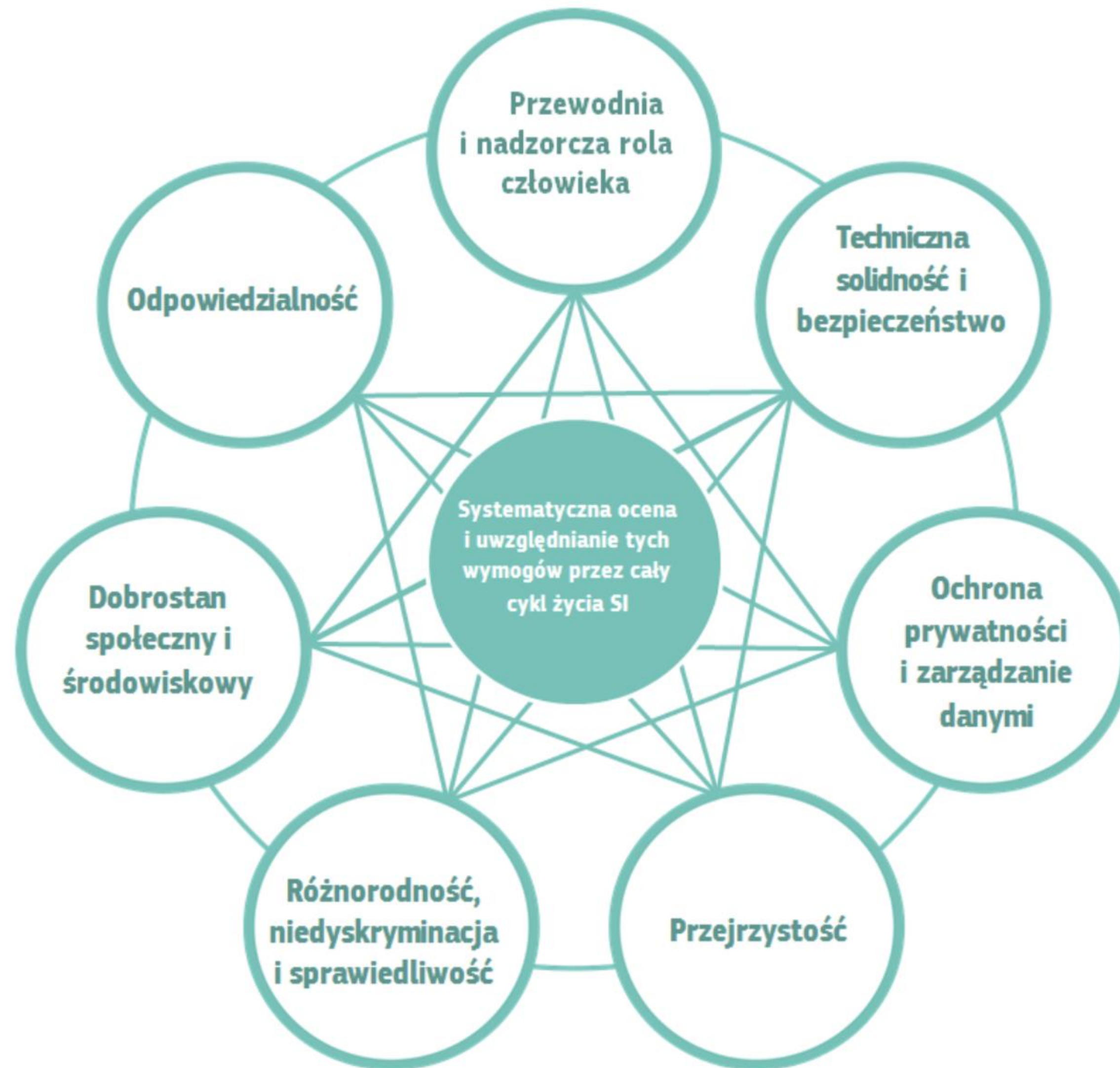
- przejrzystość (*transparency*)
- sprawiedliwość (*justice, fairness, equity*)
- nieczynienie zła (*non-maleficence*)
- odpowiedzialność i rozliczalność (*responsibility and accountability*)
- prywatność (*privacy*)
- przyczynianie się do dobra (*beneficence*)

Krajobraz standardów etycznych AI

Wartości wskazywane w standardach etycznych

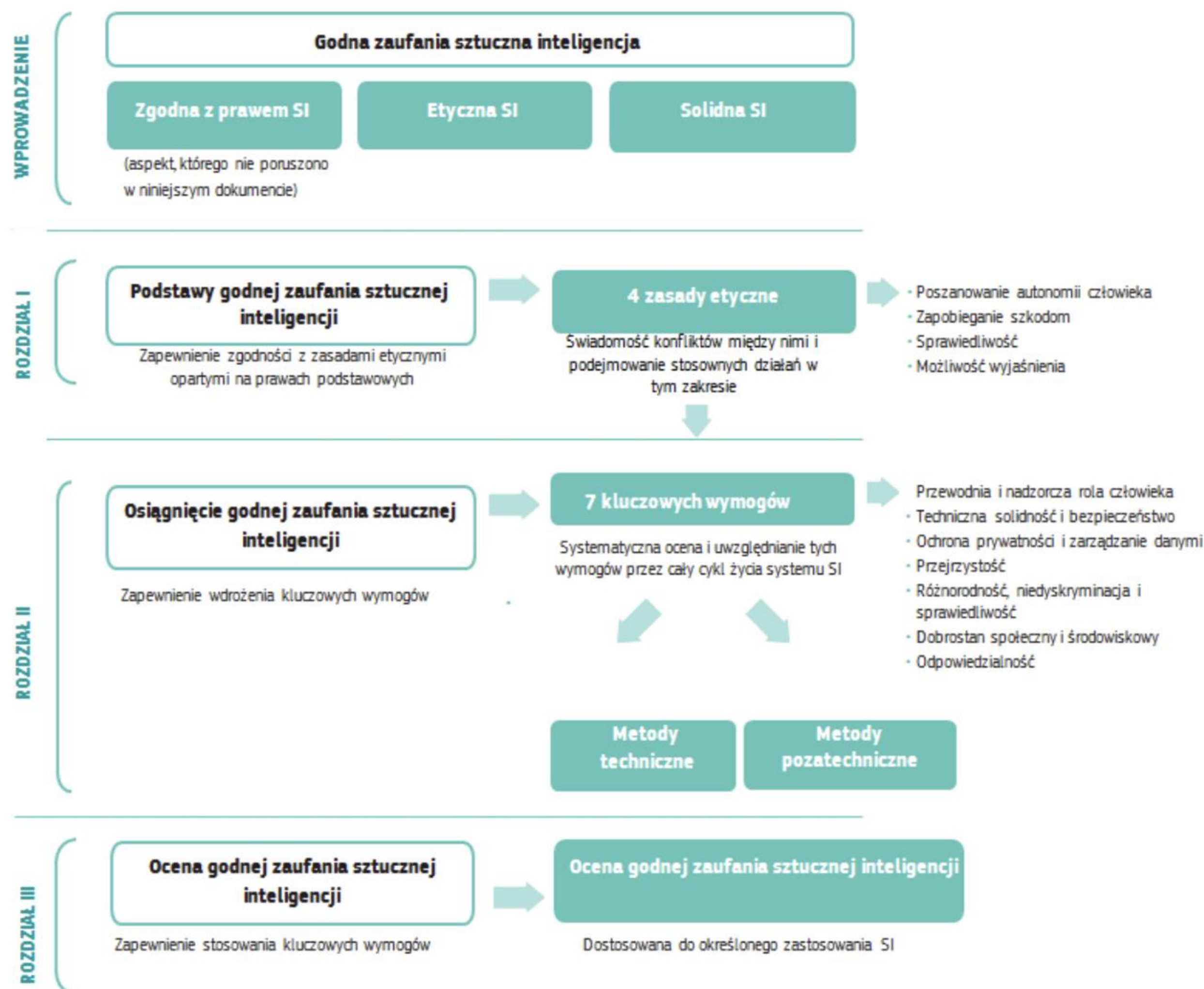
- wolność i autonomia (*freedom and autonomy*)
- zaufanie (*trust*)
- zrównoważony rozwój (*sustainability*)
- godność (*dignity*)
- solidarność (*solidarity*)

Standardy etyczne dla AI w UE



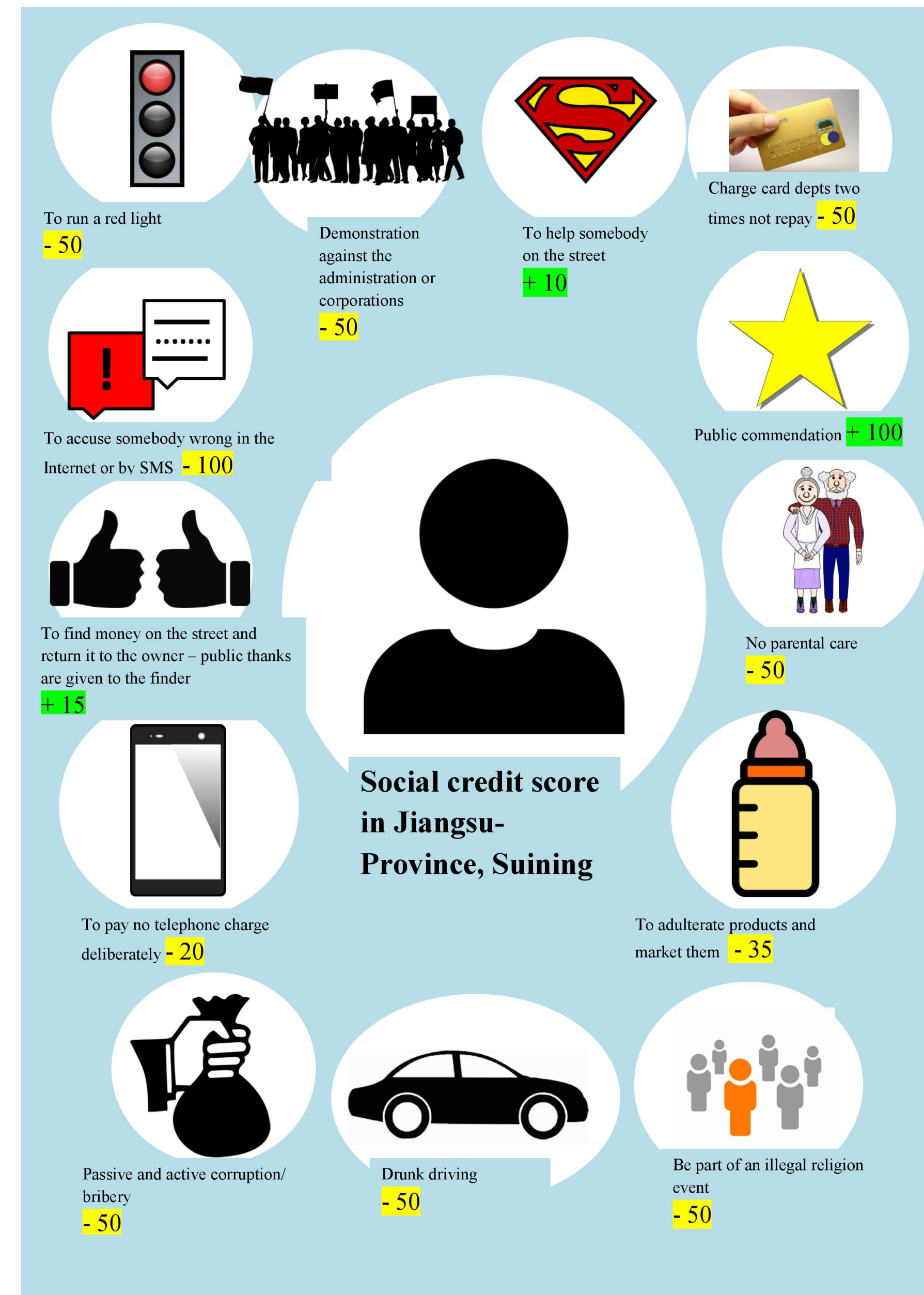
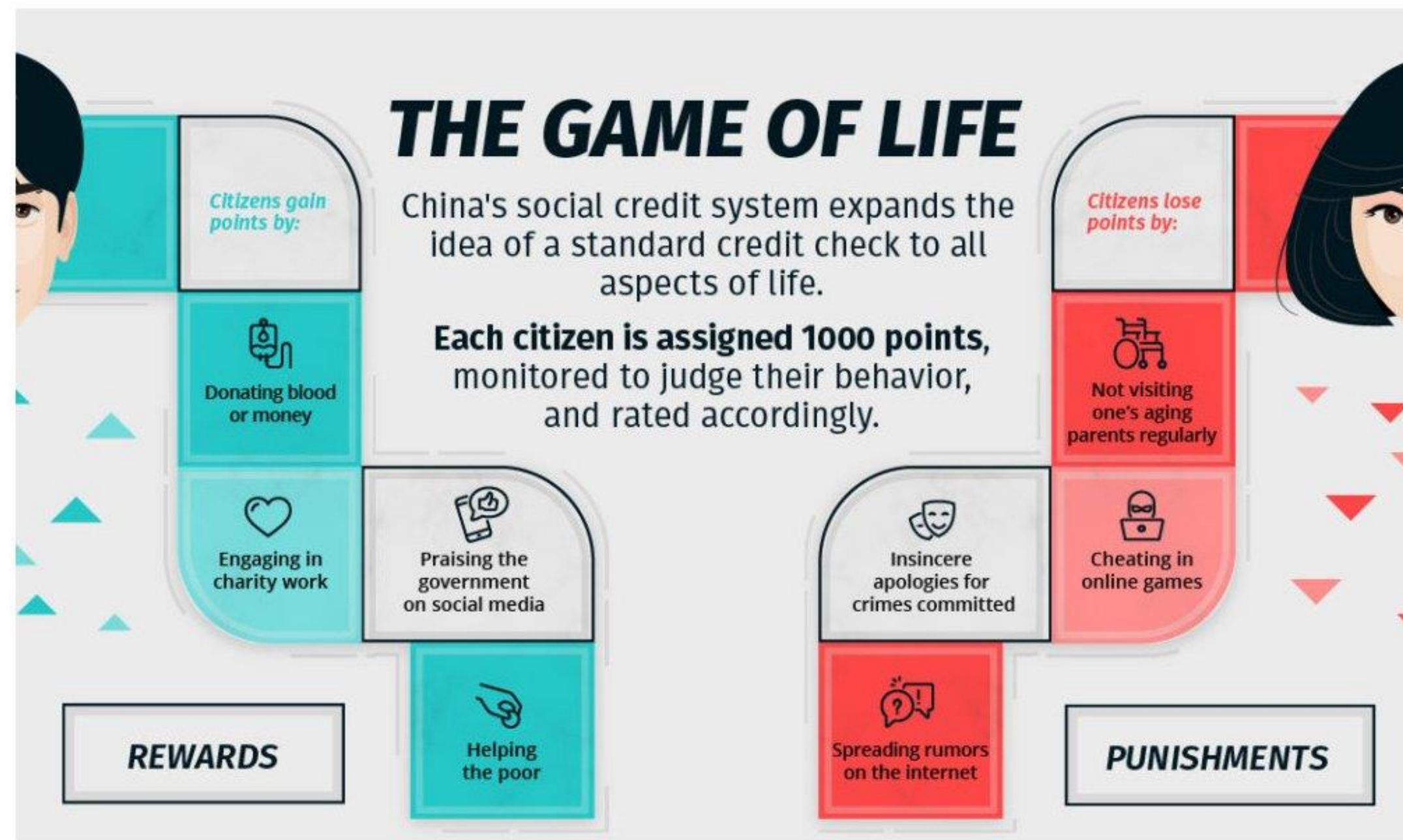
Standardy etyczne dla AI w UE

Ramy dotyczące godnej zaufania sztucznej inteligencji



Ryzyka związane z przekładem zasad na praktykę

- wybiórczość (*ethics shopping*)
- „mydlenie oczu” (*ethics bluewashing*)
- lobbing etyczny (*ethics lobbying*)
- dumping etyczny (*ethics dumping*)
- uchylanie się (*ethics shirking*)





Education
2030

UNESCO certifies that

The Project

**Improving literacy, preserving language, and
providing inclusion using AI and big data**

of

TAL Education Group

was one of the selected innovations for the **Mobile Learning Week 2020**
under the theme "Artificial Intelligence and Inclusion"

UNESCO recognizes the value of this innovation in the contribution
towards the direction of more inclusion and equity in education

Berhane Chalkoon

Director of the Division for Policies and Lifelong Learning
Systems

Paris, 8 April 2020

The Moral Machine, MIT

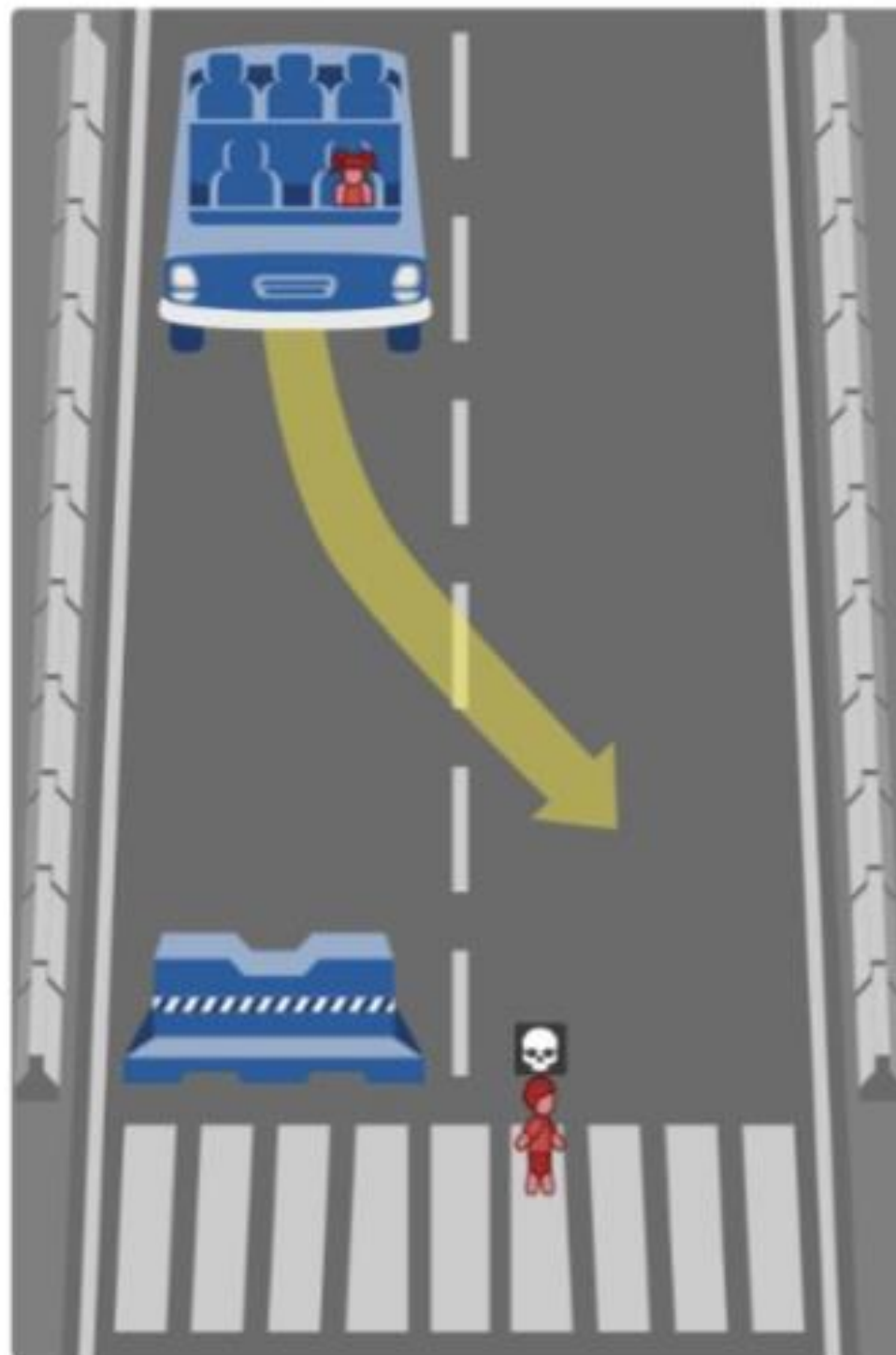


What should the self-driving car do?

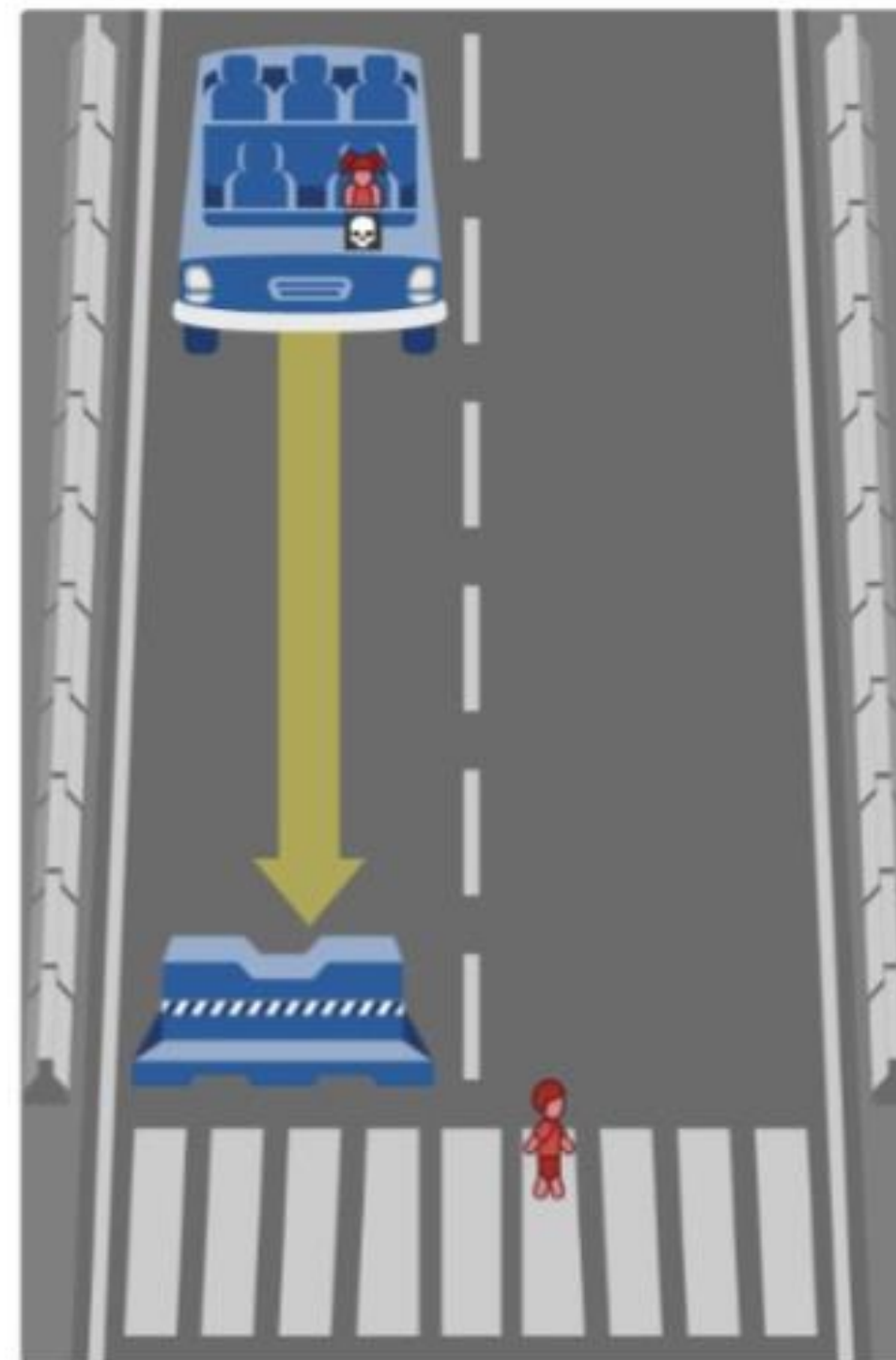
In this case, the self-driving car with sudden brake failure will swerve and drive through a pedestrian crossing in the other lane. This will result in ...

Dead:

- 1 boy



Hide Description



Hide Description

2 / 13

In this case, the self-driving car with sudden brake failure will continue ahead and crash into a concrete barrier. This will result in ...

Dead:

- 1 girl

Manipulowane czynniki:
płeć
działanie vs. brak działania,
poświęcanie pasażera vs.
poświęcanie przechodniów

The Moral Machine, MIT



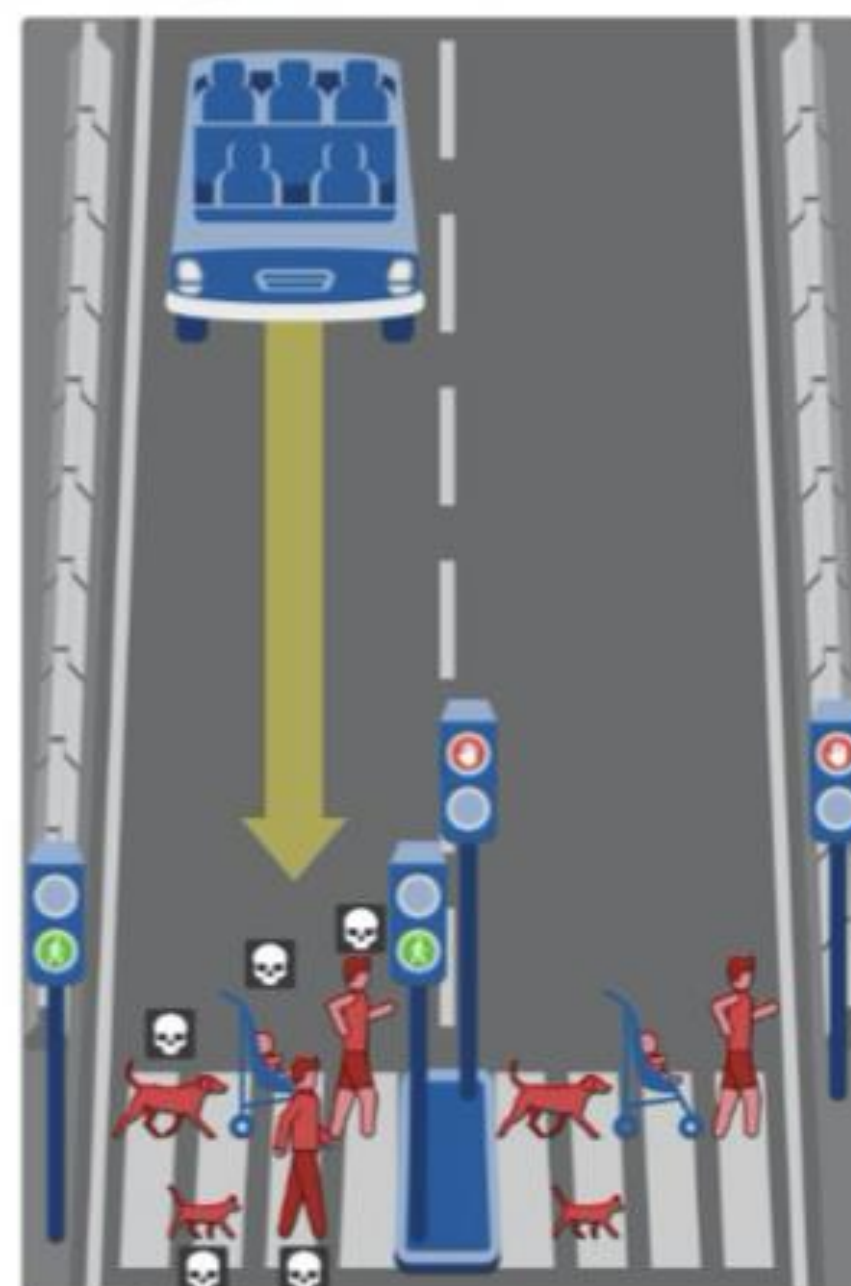
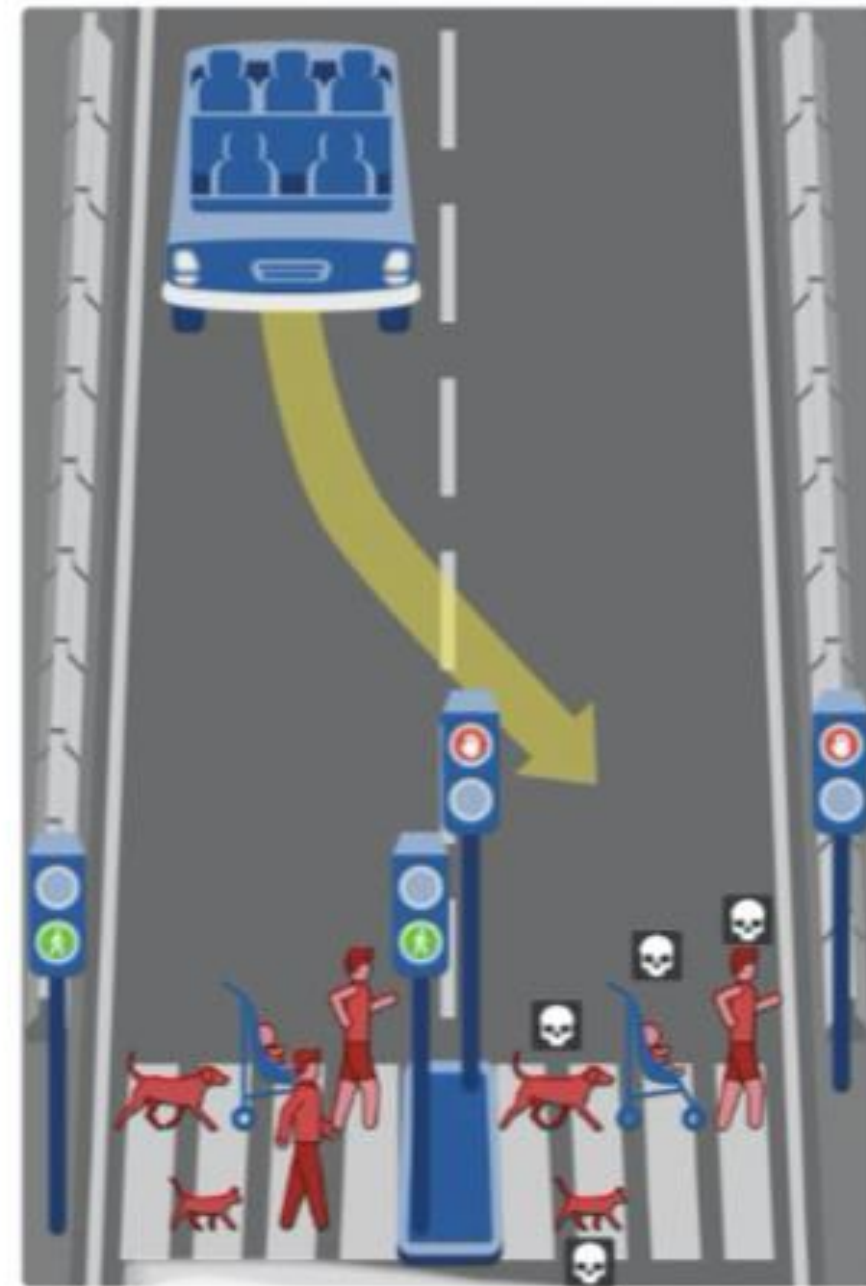
What should the self-driving car do?

In this case, the self-driving car with sudden brake failure will swerve and drive through a pedestrian crossing in the other lane. This will result in ...

Dead:

- 1 dog
- 1 baby
- 1 male athlete
- 1 cat

Note that the affected pedestrians are flouting the law by crossing on the red signal.



In this case, the self-driving car with sudden brake failure will continue ahead and drive through a pedestrian crossing ahead. This will result in ...

Dead:

- 1 dog
- 1 baby
- 1 male athlete
- 1 cat
- 1 man

Note that the affected pedestrians are abiding by the law by crossing on the green signal.

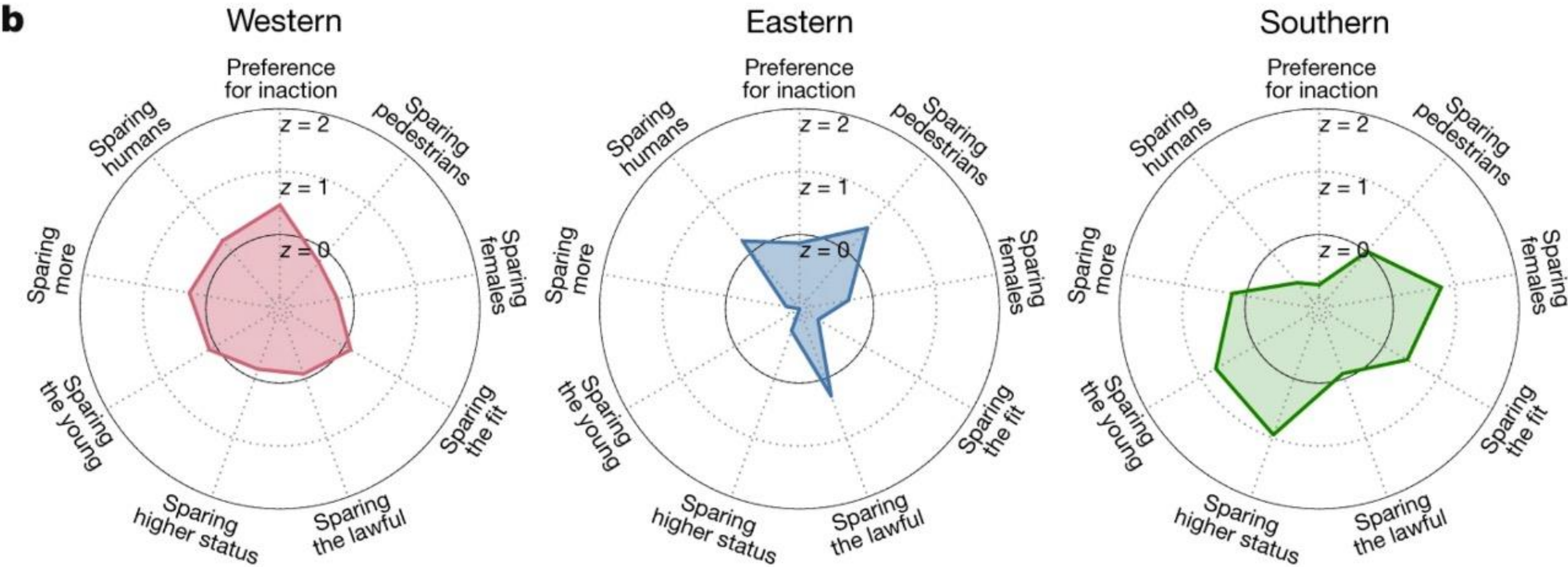
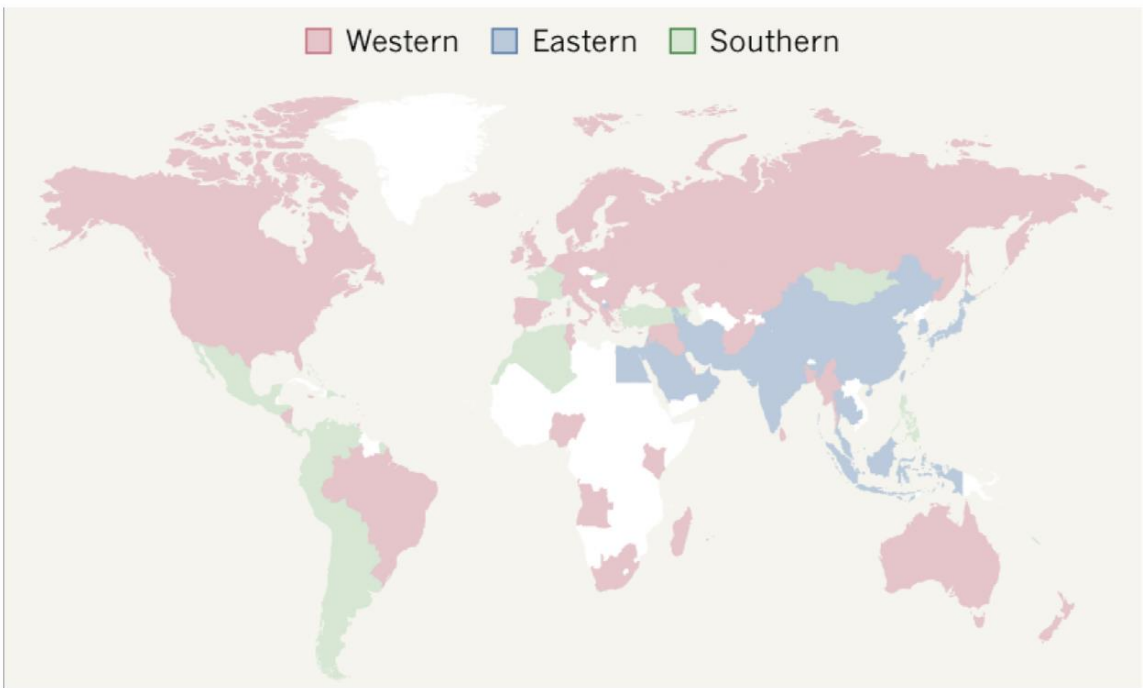
działanie vs brak działania

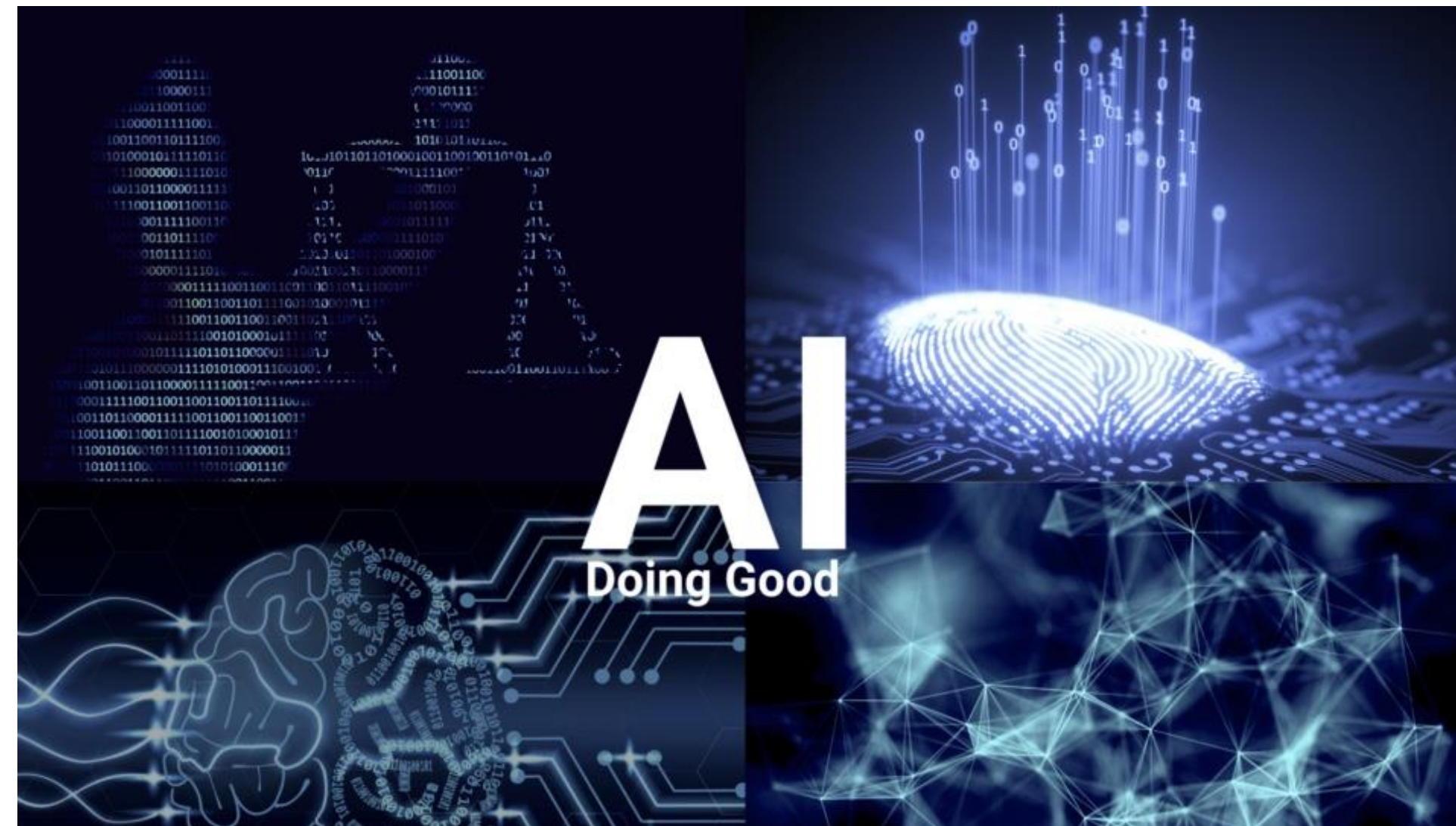
liczba zabitych osób

rodzaj zabitych istot

przestrzeganie nakazów prawnych

The Moral Machine, MIT





Dziękuję za uwagę

Ada Florentyna Pawlak

Uniwersytet Łódzki

Towarzystwo Etyki i Filozofii Techniki

adaflorentynapawlak@gmail.com