humanet3
the third attempt at a human-centered internet

MAX PLANCK INSTITUTE
FOR HUMAN DEVELOPMENT

Ethics for Computational Social Scientists

Chaewon Yun

Humanet3 group, Center for Humans and Machines
Max Planck Institute for Human Development
yun@mpib-berlin.mpg.de

About myself

- Pursuing PhD at Center for Humans and Machines on Human-centered AI
- MSc. **Computational Social Systems** with specialization of **Ethics of Human-Technology Interaction**
  - Chair of Applied Ethics (where I tutored *Ethics, Technology, and Data* for data science, computer science, CSS students)
- Software engineer (Samsung, developing voice assistant Bixby)
- BA. Political Science and International Relations

Have you taken any ethics course?

Scientific Integrity

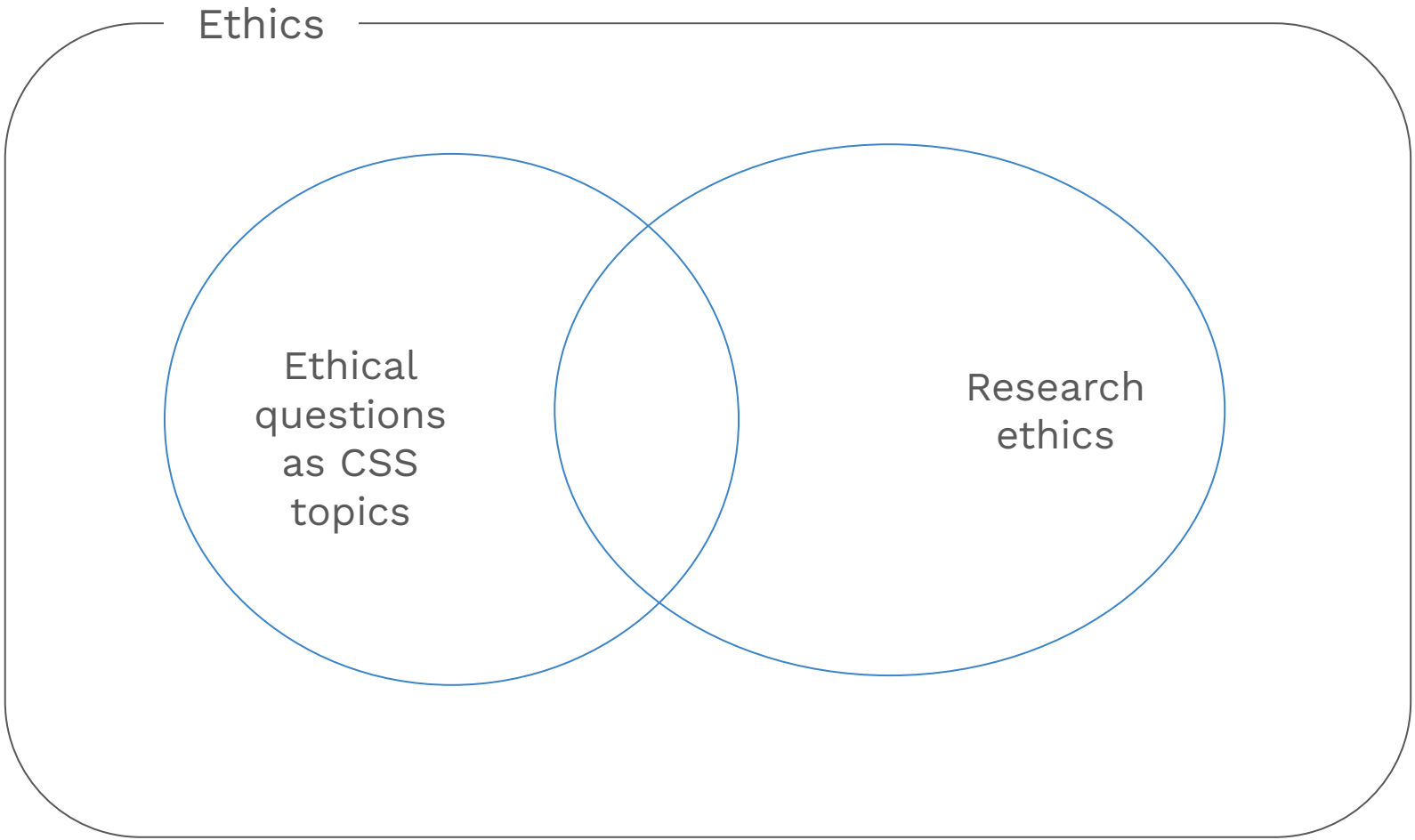
E.g. Some University makes it mandatory for every students to take before registering your thesis

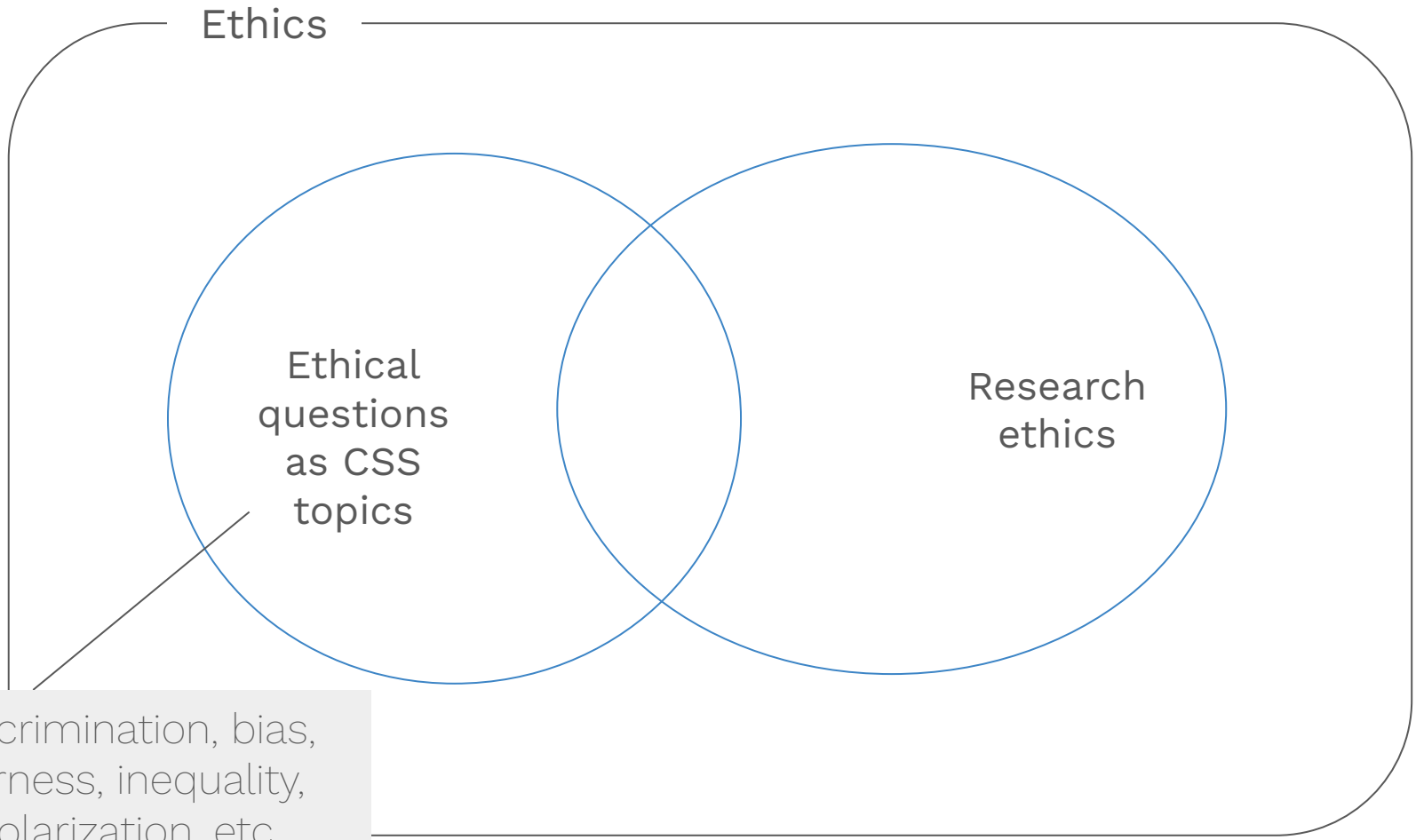
Course Description

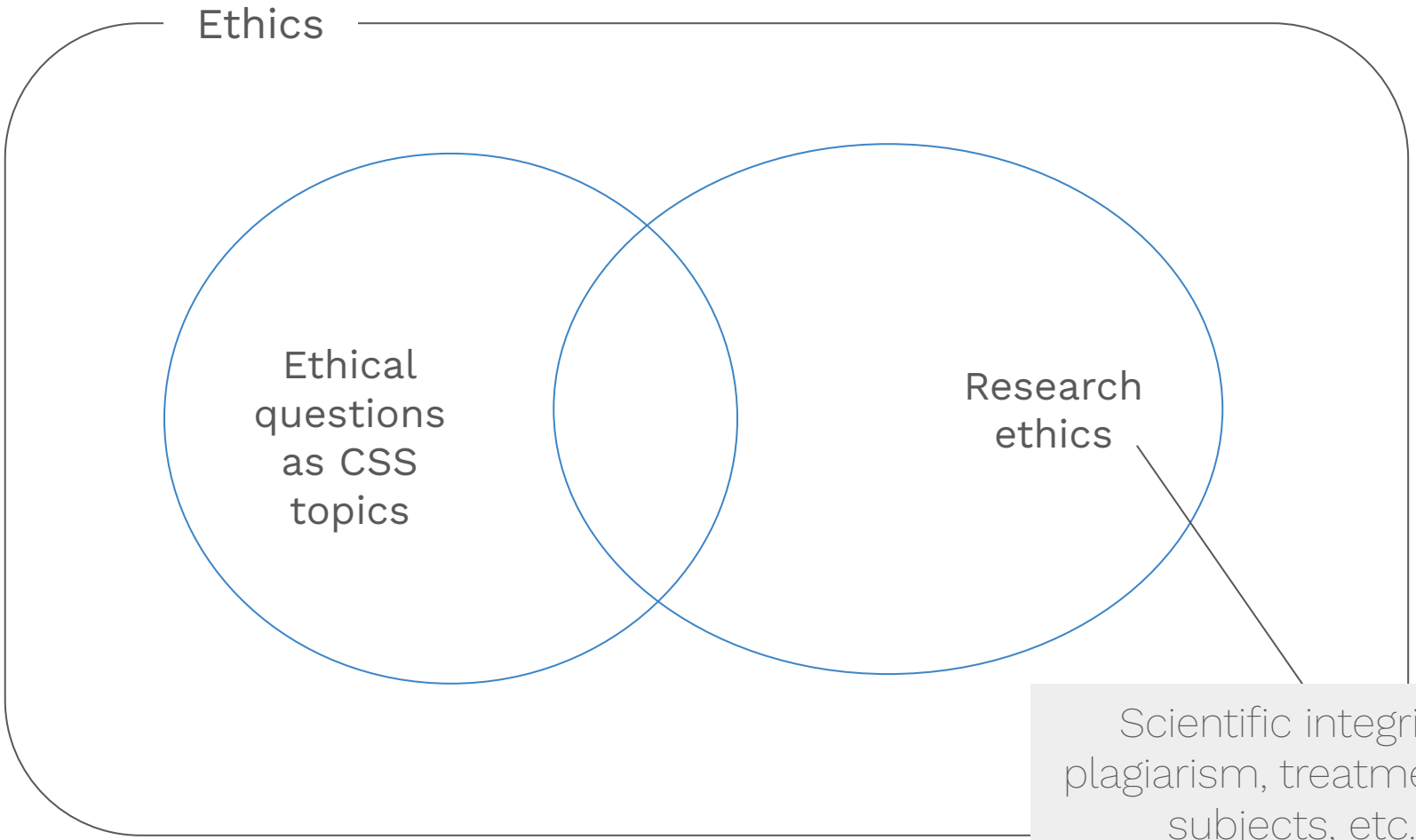
Scientific work is based on the honesty of scientists towards themselves and others. This is both an ethical norm and the basis for good scientific practice. Ensuring the validity and application of these rules in practice is therefore a core task of science and universities.

In the context of the Scientific Integrity Online Course, interdisciplinary relevant topics are presented with regard to good scientific practice. Based on the definition and principles of scientific integrity, the course covers aspects such as honesty and quality criteria for good scientific research, ensuring good scientific practice, scientific misconduct, research ethics, research data management, responsibility for supervision, and conflicts of interest as well as diversity and cooperation in science.

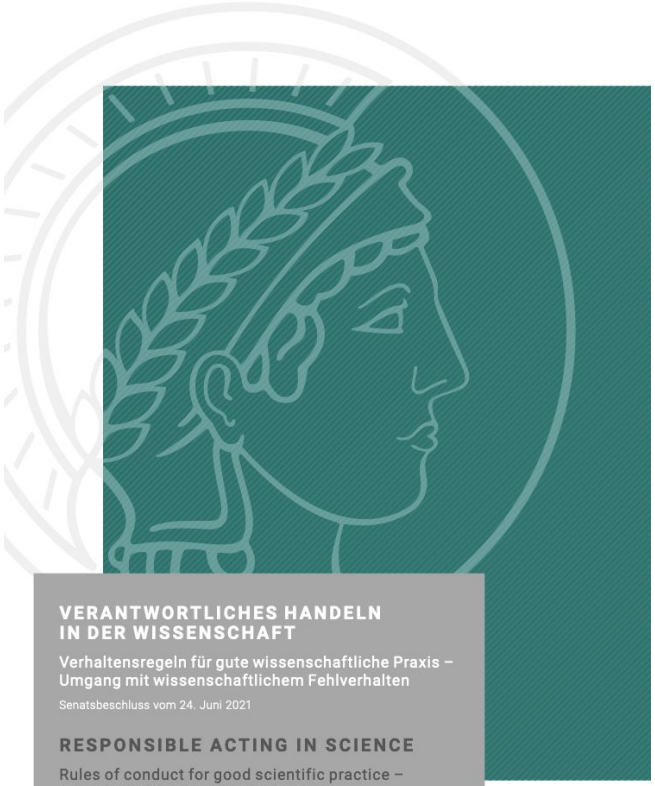
What do you think this talk (ethics for CSS) would be about?


Ethics
Ethical questions as CSS topics
Research ethics


Ethics
Ethical questions as CSS topics
Research ethics
Discrimination, bias, fairness, inequality, polarization, etc.


Ethics
Ethical questions as CSS topics
Research ethics
Scientific integrity; plagiarism, treatment of subjects, etc.

MAX PLANCK GESELLSCHAFT



VERANTWORTLICHES HANDELN IN DER WISSENSCHAFT

Verhaltensregeln für gute wissenschaftliche Praxis – Umgang mit wissenschaftlichem Fehlverhalten

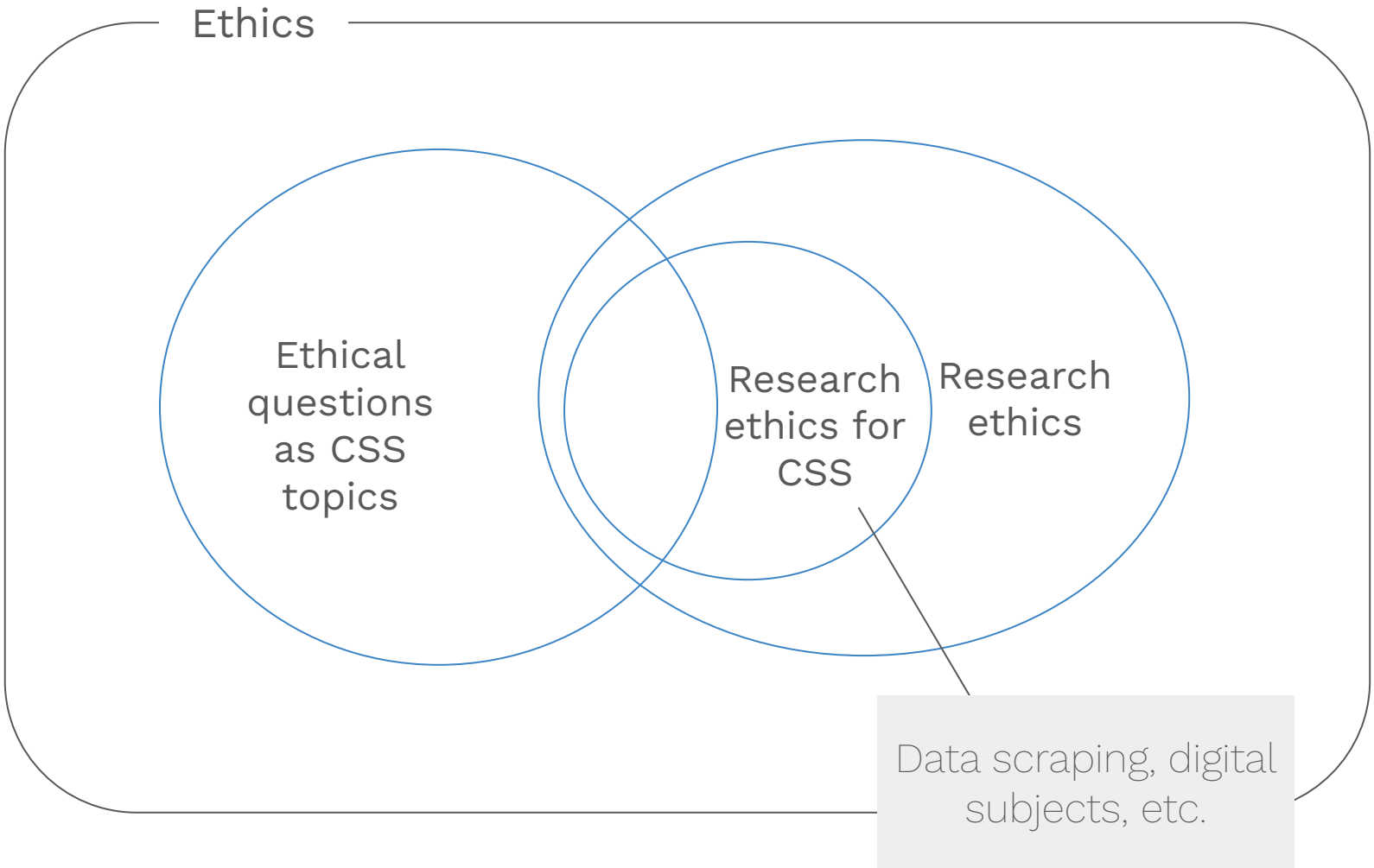
Senatsbeschluss vom 24. Juni 2021

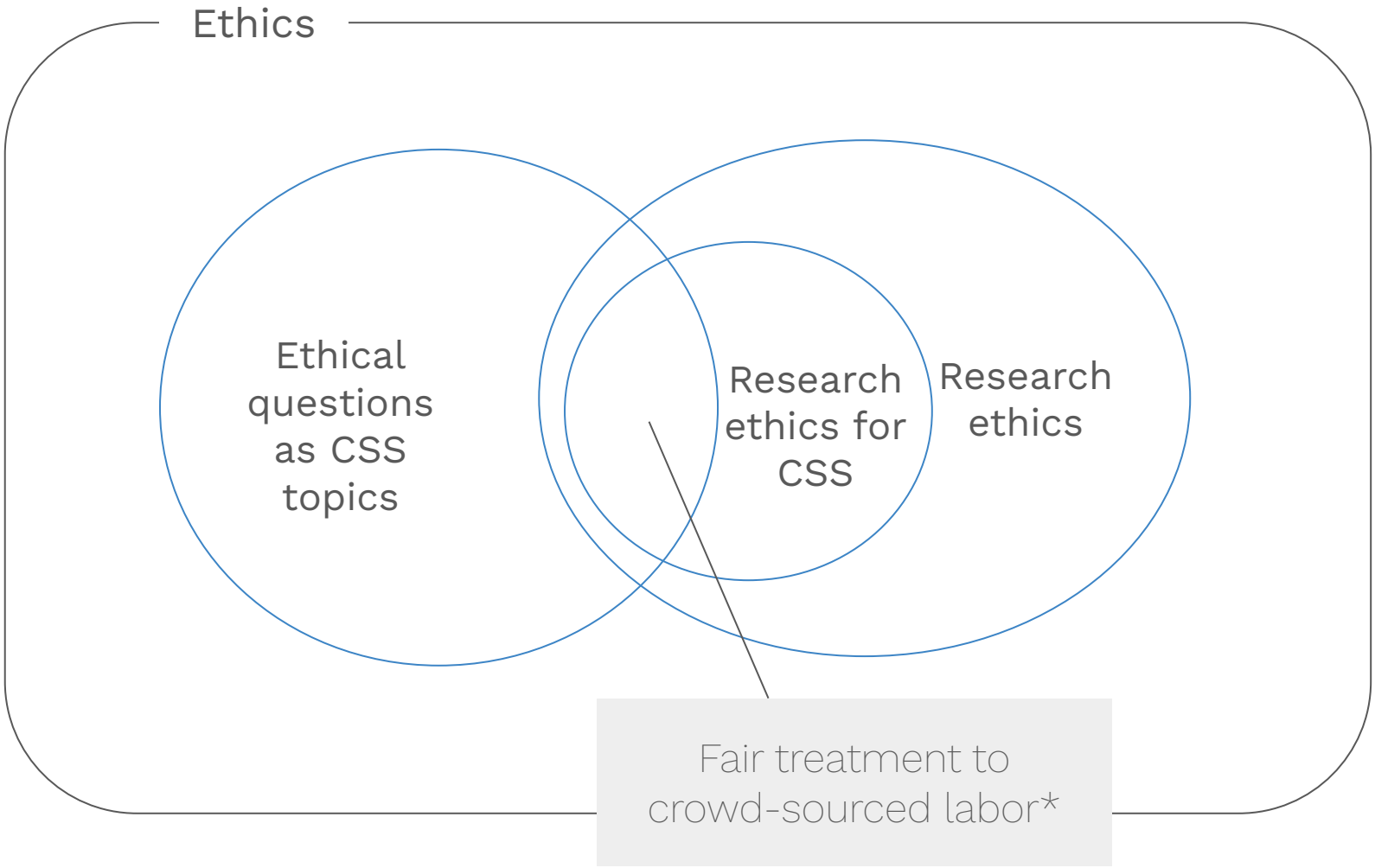
RESPONSIBLE ACTING IN SCIENCE

Rules of conduct for good scientific practice – How to handle scientific misconduct

Senatsbeschluss vom 24. Juni 2021

CONTENT


Ethics
Ethical questions as CSS topics
Research ethics for CSS
Research ethics
Data scraping, digital subjects, etc.

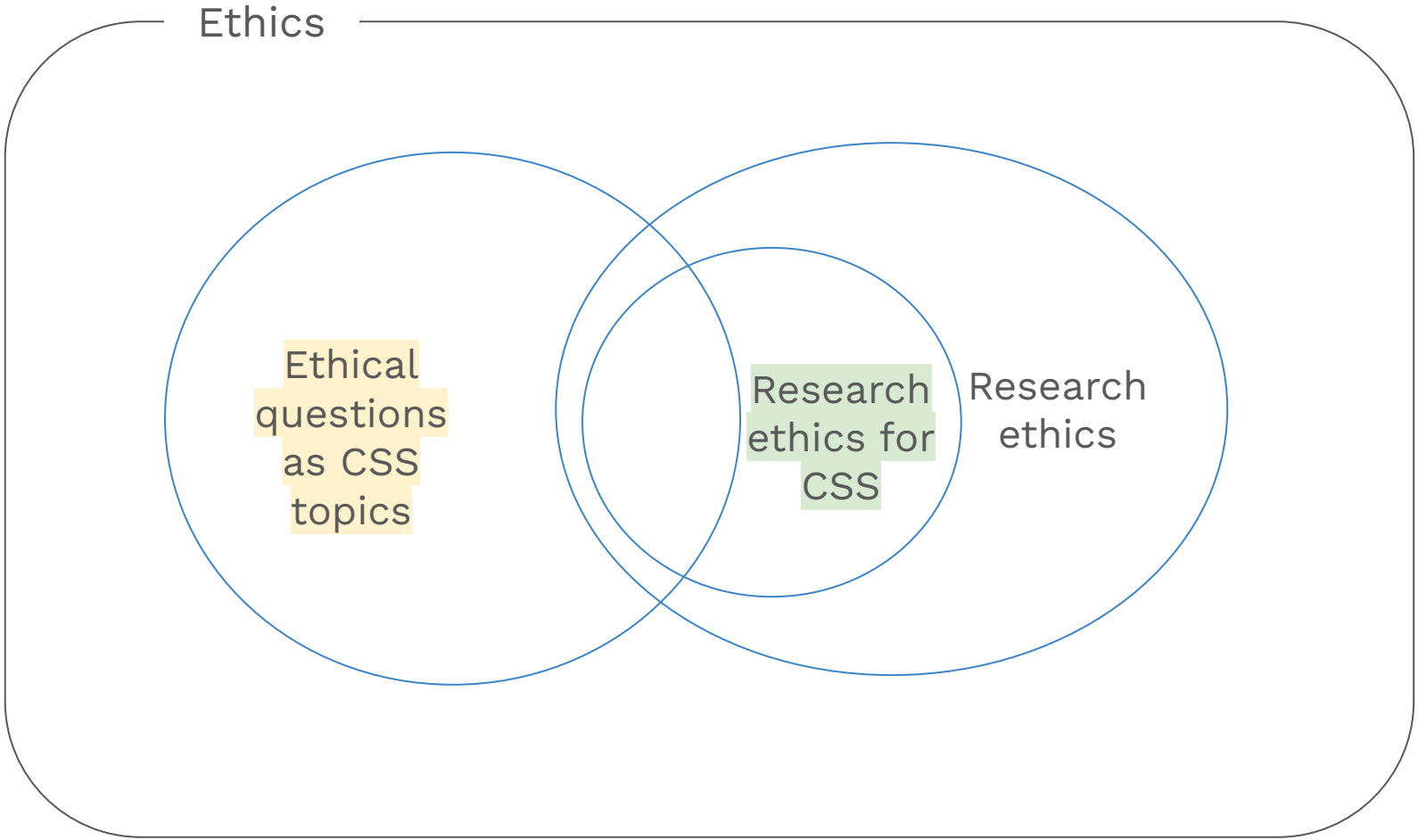

Ethics
Ethical questions as CSS topics
Research ethics for CSS
Research ethics
Fair treatment to crowd-sourced labor*

08.07.2024

17:00 Uhr - 19:00 Uhr | online

Launching the Data Workers' Inquiry

The world of AI is powered by millions of people who are precariously employed and politically dispersed. In the Data Workers Inquiry, they tell their stories.


Ethics
Ethical questions as CSS topics
Research ethics for CSS
Research ethics

Non-Topics

Legality of anything: ethical questions != legal questions

(including guidelines, white papers, policy suggestions etc.)

Ethical questions as CSS research topics

Not-so-exhaustive research topics in CSS

Nice intro: Computational Social Science: Obstacles and Opportunities

Different types of researches e.g. IC2S2 '23 program

- Qualitative studies
  - Digital ethnography e.g. TikTok and experimental digital ethnography
- Observational studies
  - Digital behavioural data e.g. Digital behavioural data to understand human behaviour
  - Social media analysis e.g. Systematic discrepancy of political advertisements in social media
- Experimental studies
  - Online experiments e.g. The moral machine experiment
  - Offline experiments
- Model-based approaches
  - Agent-based modeling
  - Modeling humans using LLMs

Social impact, fairness, discrimination, etc.

- Gender inequality (Wagner et al., 2021)
- Representation (Bayes et al., 2022)
- Social impact of NLP (Hovy and Spruit, 2016)
- Hate speech (Warner et al., 2012)

Research ethics for Computational Social Scientists

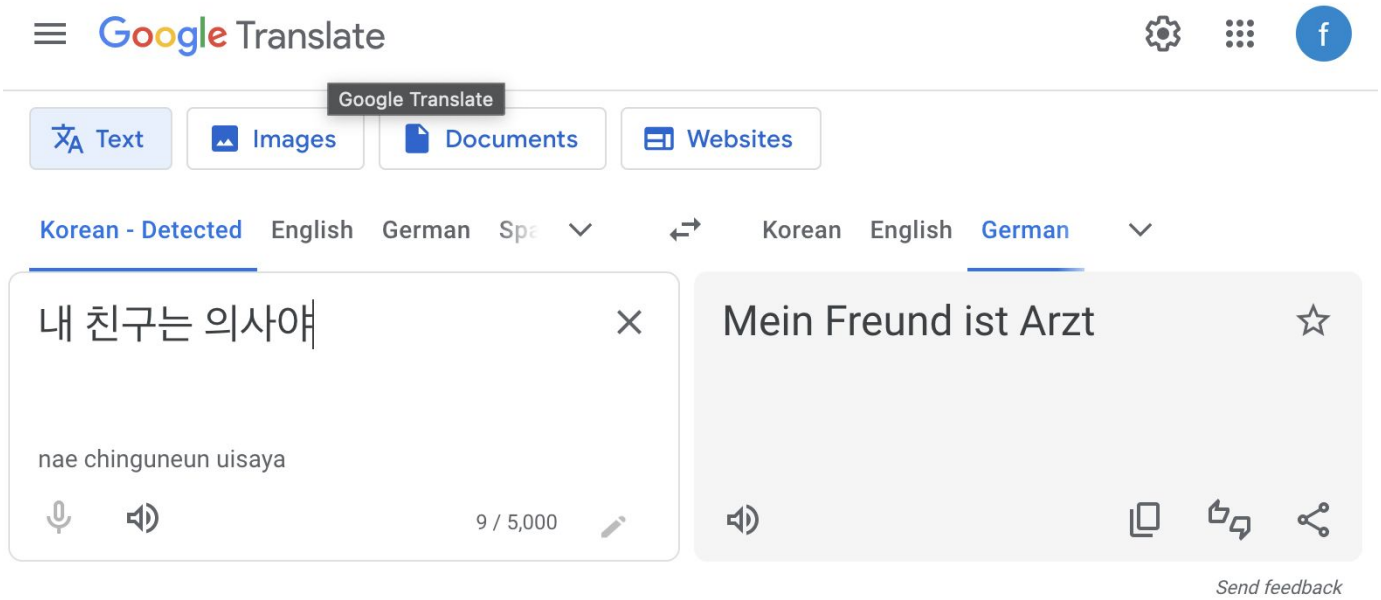
What are the (potentially) relevant ethical questions in your research?

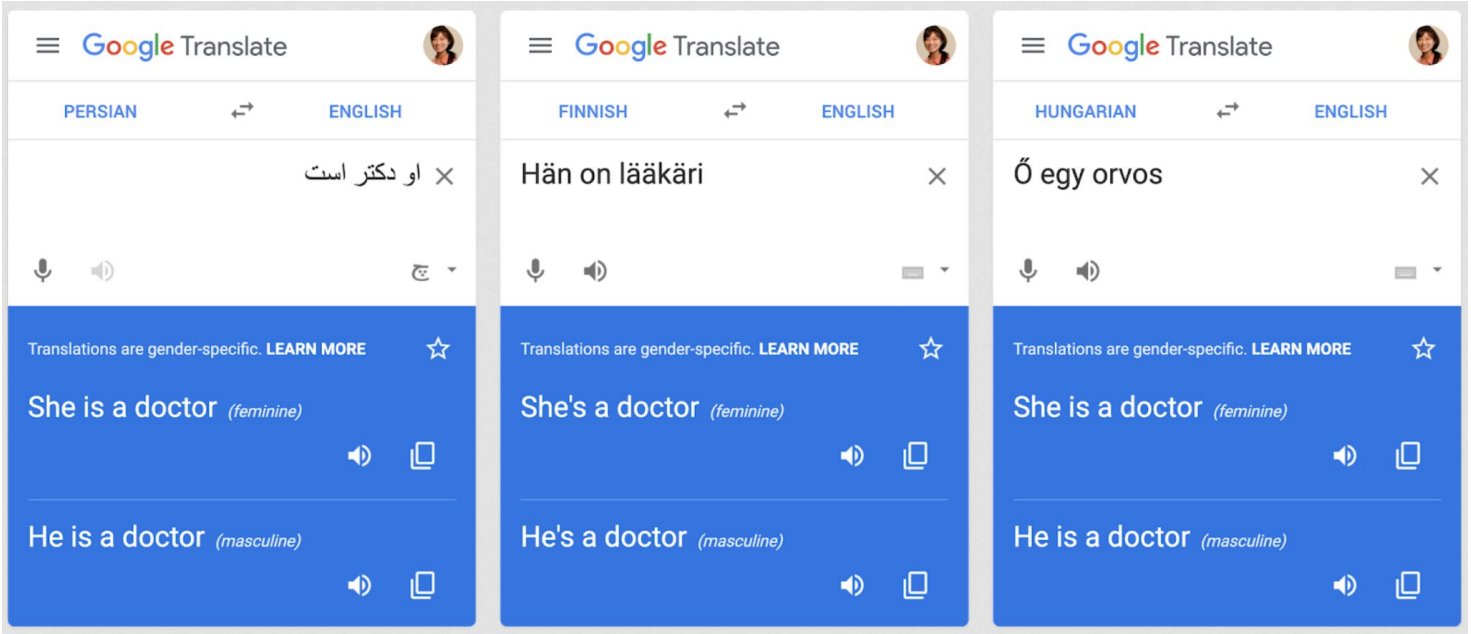
Research ethics in full research cycle

- Choosing the topic
- Conducting experiments - not only with human subjects directly but through data
- Collecting data
  - Digital behavioural data - no explicit informed consent
    - Which platform to use? Which data to collect? How to collect?
- Creating dataset
  - Annotators - whom to choose? How to pay? etc.
- Analyzing the data
  - Using ChatGPT?
- Publishing the result
  - Anonymization, open science practice, etc.

Common ethical questions

- Famous grey area - Data scraping & Terms of services
- Tragedy of the commons
- Societal impact of model produced in research
  - e.g. “Detecting” sexual minorities , facial recognition for emotion detection etc.
- Cooperation with big tech?
  - E.g. Getting dataset without being able to share, can run experiment but cannot make it reproducible
  - TikTok API - you need to be approved by the company before getting the access
- “Experimental Evidence of Massive-Scale Emotional Contagion” (Kramer et al., 2014)


Google Translate
Google Translate
Text
Images
Documents
Websites
Korean - Detected
English
German
Spa
Korean
English
German
내 친구는 의사야
nae chinguneun uisaya
9 / 5,000
Mein Freund ist Arzt
Send feedback


Google Translate
PERSIAN
ENGLISH
او دکتر است
Translations are gender-specific. LEARN MORE
She is a doctor (feminine)
He is a doctor (masculine)
Google Translate
FINNISH
ENGLISH
Hän on lääkäri
Translations are gender-specific. LEARN MORE
She's a doctor (feminine)
He's a doctor (masculine)
Google Translate
HUNGARIAN
ENGLISH
Ő egy orvos
Translations are gender-specific. LEARN MORE
She is a doctor (feminine)
He is a doctor (masculine)

Who’s responsible for
Google Translate’s gender bias?

Problem of Many Hands

- The actions of the four people together lead to a harmful outcome, but none of the individual persons can seemingly be held responsible
- Because there are different people involved, it is impossible to identify one single person that is responsible

Responsibility

Direction of responsibility

- Backward-looking responsibility: “I am responsible for having drank your coffee.”
  - The most common form invoked when considering moral blame or praise; linked to liability
- Forward-looking responsibility: “I am responsible for finishing my thesis.”
  - For moral responsibility, appear as a recognition of moral duties; one’s role often serves to assign forward-looking responsibility

Mode of Responsibility

- Passive responsibility: being held responsible by others for some outcome
  - Strongly linked to backward-looking and liability responsibility
- Active responsibility
  - Strongly linked to forward-looking and role responsibility

Problem of Dual Use

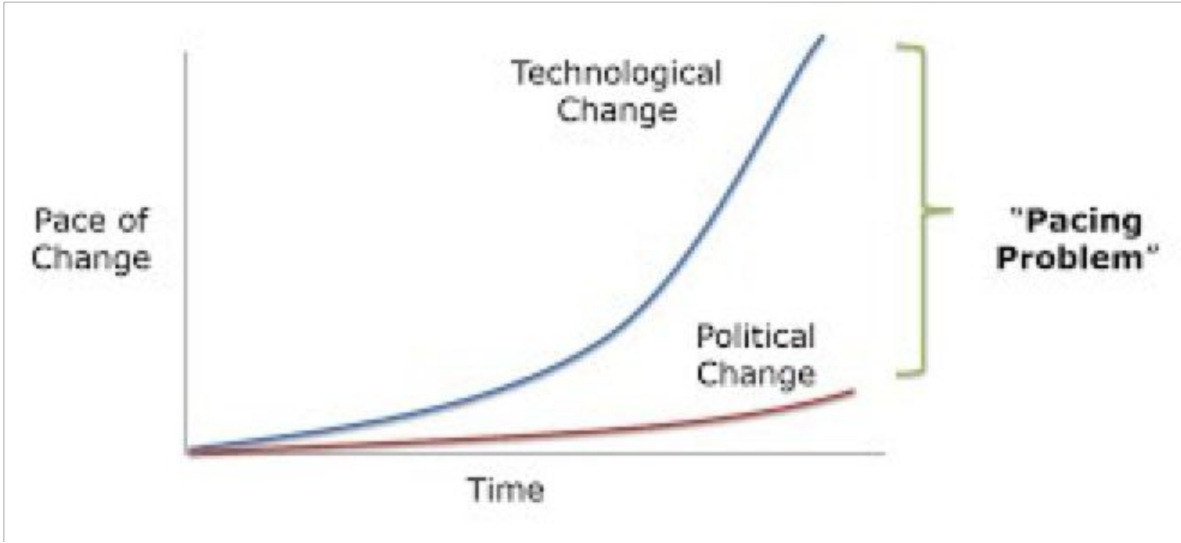
- The unintended use of a product = the dual-use
  - Hammer,
- Ethical challenge: We are responsible not only for what we intend a product to do, but also (to some degree) for the dual-uses thereof
- CSS version of dual use?

Precautionary Principle

- Strongest variant: no course should be pursued without absolute surety of safety
  - Untenable – the bar is set far too high
- Moderate variant: a course should be pursued if, on the balance of probabilities, the course will bring about sufficiently greater benefit than harm

→ How to determine the level of sufficiency, and how accurate must the measure of probabilities be?

Collingridge Dilemma



Collingridge Dilemma

- By the time we know what the risks are, changes are difficult
- Discussion of technological developments happen too early or too late
- Technology in the early phase of its development:
- little overall knowledge - in principle, a lot of room for design in the social sense
- The more technology advances (and the more knowledge we have), the smaller the scope for design in the social sense

Reading recommendations

- Introduction to Data Ethics (Shannon Vallor)
- (SEP) Internet Research Ethics
- (Open Textbook) Introduction to Philosophy: Ethics
- (SEP) Ethics of Artificial Intelligence and Robotics
- (SEP) Computing and Moral Responsibility

Thank you!

Chaewon Yun

Humanet3 group, Center for Humans and Machines
Max Planck Institute for Human Development
yun@mpib-berlin.mpg.de