DIVERSITY BY DESIGN?
BALANCING THE PROTECTION AND INCLUSION OF USERS IN ONLINE SOCIAL PLATFORMS

ACM/AAAI AIES Conference Workshop 2021

Oxford, August 1-3, 2022

Paula Helm,
Technical University Munich, Germany

Loizos Michael, Open University of Cyprus/
CYENS Center of Excellence, Cyprus

Laura Schelenz, International Center for Ethics,
University of Tübingen, Germany

WWW.INTERNETOFUS.EU

How can we build interfaces and algorithms that support meaningful user interactions in online social platforms through considering the values and risks of diversity?
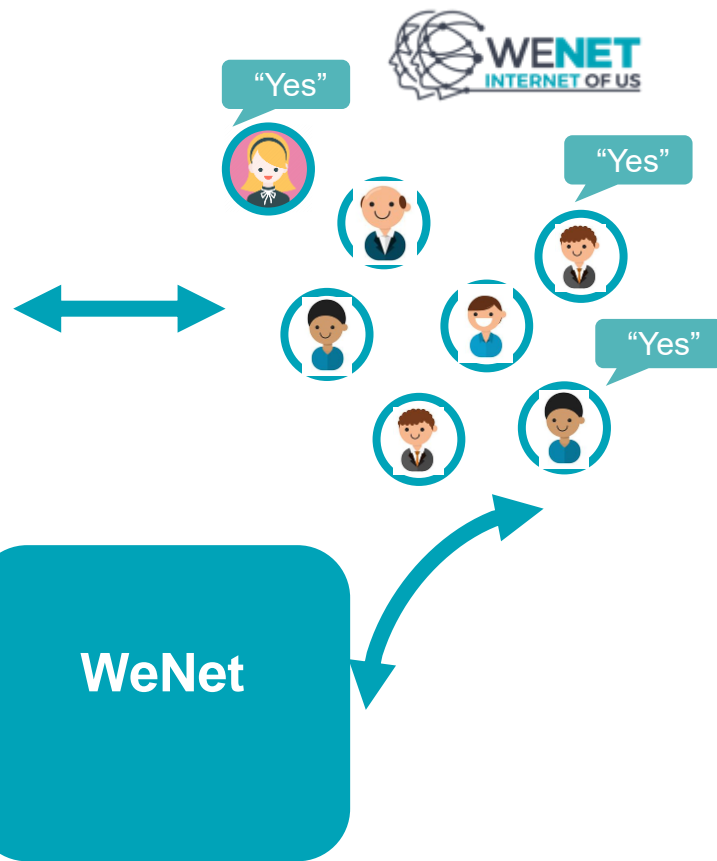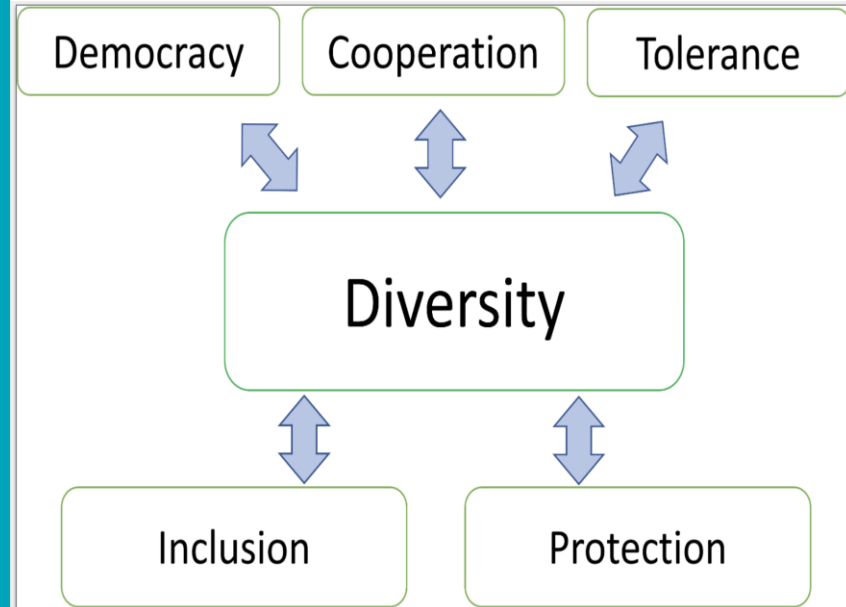
WENET
INTERNET OF US
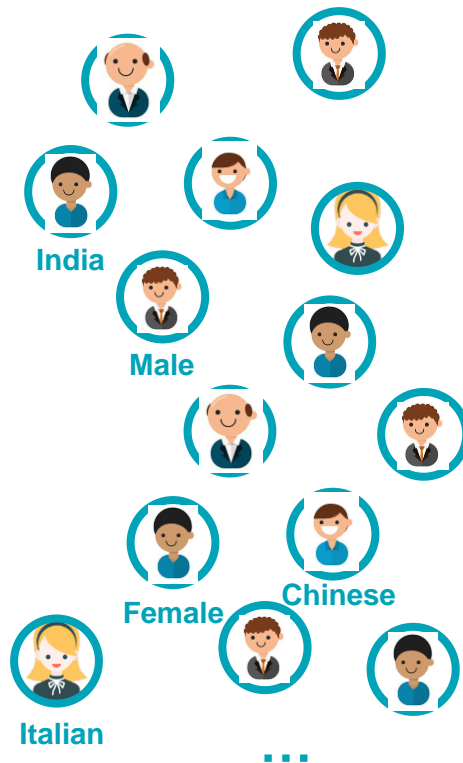
DIVERSITY

Diversity as a normative concept

From an ethical perspective, diversity helps achieve other values such as inclusion, tolerance, equality, and democracy, and vice versa.

Diversity as a descriptive category

As a descriptive category, diversity helps us define differences between users and ultimately classify elements of their personalities.
Here, diversity refers to both demographic properties and different social practices that relate to different skills, competencies, interests, sensibilities, etc.

India

Male

Female   Chinese

Italian

...

Diverse in:

Visible properties:
Gender
Nationality
Age
Role
…

Deep properties:
Personality
Culture
Competence
…

WENET
INTERNET **OF US**
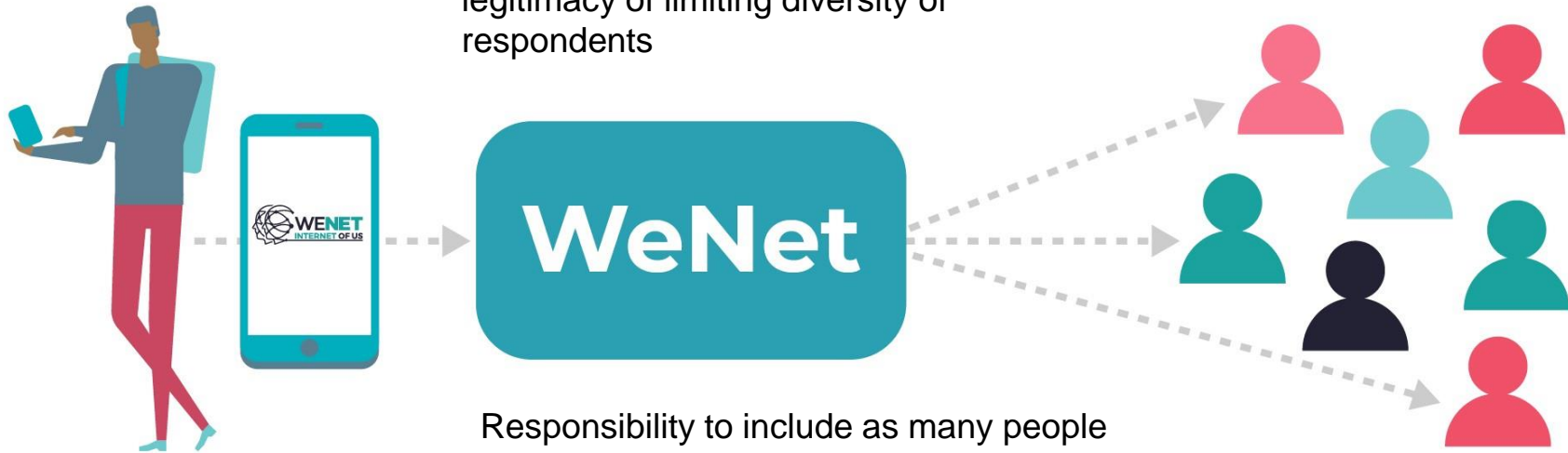
THE PROTECTION–INCLUSION–TENSION

Opening the Echo Chamber while Protecting Vulnerable Groups

Different tests/standards to determine legitimacy of limiting diversity of respondents

WeNet

Responsibility to include as many people who are different as possible → promote diversity

Methodology

Value Sensitive Design

- Departs from a widely-held understanding of technology as being objective and value-neutral.

- Proactively seeks to identify key norms to guide the inscription of values into technologies.

- Evaluates the potential range of tools that may be employed to realize this process, and it reflects on their future implications (Friedman, Hendry, 2019).

Scenario Analysis

- Applied to envision multiple possible futures and prepare and develop ways to deal with them.

- Value scenarios are narratives that tell stories about use and are intended to show the entanglement of human and technical aspects in specific contexts.

- They emphasize impacts on associated key values, indirect impacts, and similar systemic effects (Nathan, Klasnja, Friedman, 2007).

3 Steps

1. Assess the relevant values and the factors that contextualize them.

2. Empirically determine how to balance the relevant factors through a value-centered scenario analysis and identify the key ethical norms.

3. Technically implement the identified norms into an actionable policy by operationalizing the relevant factors in a formal representation.

Ask a question through the WeNet chatbot

WeNet

Technology connects to someone similar or different

WeNET
INTERNET OF US

Khulan

- Student at an European university
- Exchange student from abroad
- Suffers from test anxiety

What would you like to ask? (Users send a request via the chatbot into a community)	"Are you struggling with mental health issues that can make it difficult for you to study? And if so, how do you deal with them?"
Who should receive this request? Someone similar or different? (Users specify in the chatbot)	Someone "similar to me"
Ethical analysis: Is it legitimate to limit the scope of recipients by asking for users who are similar to oneself to answer the request?	Limiting diversity is legitimate (protection argument)
Ethical reasoning	In this scenario, it is imperative for users to be able to limit diversity in order to create a safe space.

Protection argument:
Can be generalized by creating the following test or standard:

Does the request relate to a sensitive sphere such as religion, health, or is sensitive due to cultural or social circumstances?

What appropriate interventions could follow?
Send an incentive, a nudge, etc.

Khulan

- Student at an European university
- Exchange student from abroad
- Suffers from test anxiety

What would you like to ask? (Users send a request via the chatbot into a community)	"Are you struggling with mental health issues that can make it difficult for you to study? And if so, how do you deal with them?"
Who should receive this request? Someone similar or different? (Users specify in the chatbot)	Someone "similar to me"
Ethical analysis: Is it legitimate to limit the scope of recipients by asking for users who are similar to oneself to answer the request?	Limiting diversity is legitimate (protection argument)
Ethical reasoning	In this scenario, it is imperative for users to be able to limit diversity in order to create a safe space.

test#	a	b	c	d	e
question	Does the request speak to a certain demographic, or involve specific knowledge/skills that are tied to a social practice?	Does the request relate to a sensitive sphere such as religion or health or is sensitive due to a given cultural/social contexts and thus requires protection of the requester?	Is the request related to a relevant social sphere such as economy, politics, education making it a relevant resource that should be accessible to everybody?	Is the request related to the sphere of leisure, free time, voluntary activities?	Is the request violating a fundamental ethical principle?
ethical action	limit diversity	limit diversity	expand diversity	ideally expand diversity but also ok to limit diversity	neither limit nor expand diversity, make the user aware of this violation
argument type	efficiency argument	protection argument	inclusion argument	freedom of choice argument	principle of no harm argument
ethical reasoning	If there is no concern for the safety of the person (see test b) and no fundamental resource involved (see test c), and the request is not unethical per se (see test e), the system may allow the user to limit diversity for efficiency reasons based on similar demographics or similar social practices.	If the requester poses a question that deals with a sensitive issue or speaks from a vulnerable position (such as an oppressed social status), the system should allow the user to limit diversity in order to enable a safe space for the protection of the user. This is relevant to prevent the user from exposure to hate speech and offensive comments.	If the request involves a fundamental resource such as an economic opportunity or the opportunity to participate in political processes, the system should strongly incentivize the user to NOT limit diversity because a limiting of diversity would lead to the exclusion and discrimination of users in fundamental areas of life.	If the request involves an opportunity to participate in a leisure activity, the system should softly nudge the user to open this request to everyone in order to avoid exclusion. However, limiting diversity is acceptable because there is no fundamental resource involved, and the user has the freedom to choose their companions in leisure time.	If a request violates laws or the fundamental rights of other users, the system must flag this request and communicate to the author that this request is unethical per se and cannot be send to the community as it is.
reference to scenario	scenario 3	scenarios 1 and 2	scenario 6	scenario 4	scenario 5

Ethical Challenges

TECHNICALLY IMPLEMENT THE IDENTIFIED NORMS INTO AN ACTIONABLE POLICY

- Operationalizing the Ethical Norms

- Resolving Tensions between Norms

Conclusion

Diversity by Design –

A contradiction in terms?

THANK YOU

Paula.helm@tum.de

Laura.Schelenz@uni-tuebingen.de

Loizos@ouc.ac.cy