UNIVERSITY OF TRENTO - Italy
Dipartimento di Ingegneria e Scienza dell'Informazione

Know dive

Diversity, Bias, online social relations and related issues

Fausto Giunchiglia

19.04.2021

WENET INTERNET OF US

CY. center for algorithmic transparency

The Web in perspective (now)

- The Web has transformed our lives by allowing us to transcend temporal, geographical and cultural borders.
- We get exposed daily to a seemingly unbound amount of (people) diversity
  - *Diversity as a feature*: Web as a source of opportunities, see the emergence of global digital platforms (largely unnoticed, almost a given).
  - *Diversity as a bug*: Web as a source of undesired/ unexpected/ unknown opportunities, see the emergence of online fake news, bias and unethical behavior (largely noticed).
- BUT: We have the same skills to deal with diversity as our grandparents 50 years ago.
- In positive as well in negative, we need to deal with the corresponding *social challenges* in a rapidly evolving scenario (think of, e.g. COVID).

Diversity? Similarity?

- Merriam Webster defines *diversity* (quote) *“...the condition of having or being composed of differing elements.”*
- To common human experience any entity appears distinct and different from any other entity.
- Still we all talk of *diversity* / *similarity*. Why similarity? Which similarity?
- Diversity / similarity are *social* properties.
- The social property of “being diverse or similar” is contextual:
    - What is the reference neutral point?
    - What are the dimensions and metrics of evaluation?
- “Being diverse” is the opposite of “being similar” only when context is fixed

Two types of Diversity

- *Inessential (social sciences: observable / surface level):* religion, race, ethnicity, national origin, …, demographics, culture, age, …
- *Essential (social sciences: less observable / deep level):* Knowledge of medicine, Knowledge of how to fix your plumbing, competence, skills, national origin, …, demographics, culture, age, …

Note:

- Essential diversity takes longer to recognize, … where often the Web often does not allow for prolonged interactions.
- Essential diversity is what we exploit (most often unconsciously)
- Inessential diversity is what we sometimes escape (e.g., if it is an unknown unknown)

Diversity in the Web – a unifying perspective

1. *Inessential diversity*: Diversity and Bias (*) (***)
2. *Essential diversity*: Diversity aware social relations (**)
3. *Empowering diversity (**)(***)*

CY. center for algorithmic transparency

WENET INTERNET OF US

(*) Giunchiglia, F.; Otterbacher,J.; Kleanthous, S.; Batsuren, K.; Bogin, V.; Kuflik,T.; Tal, A.S. *Towards Algorithmic Transparency: A Diversity Perspective.* ArXiv (2021)

(**) Schelenz, L.; Bison, I.; Busso, M.; de Götzen, A.; Gatica-Perez, D.; Giunchiglia, F.; Meegahapola, Ruiz-Correa,L.S. *The Theory, Practice, and Ethical Challenges of Designing a Diversity-Aware Platform for Social Relations*. 4th AAAI/ACM AIES (2021)

(***) Giunchiglia, F.; Kleanthous; S.; Otterbacher,J. *Transparency Paths - Documenting the Diversity of User Perceptions*. Submitted to ACM FairUMAP (2021)

Diversity and Bias

Bias

… Quoting Wikipedia

***Bias** is a **disproportionate** weight in favor of, or against, an idea or thing, usually in a way that is **closed-minded**, **prejudicial**, or **unfair**. People may develop biases for or against an individual, a group or a belief.*

At a first sight very clear, almost commonsensical

BUT

- **Who decides** what is closed-minded, prejudicial, or unfair?
- **Who decides** what is disproportionate?
- What is the level of agreement? And among whom?

How is Bias constructed?

1. Presence of an **observer**, always associated with a **reference** point of view and a set of evaluation criteria as well as measures for those criteria, which performs the following steps,
2. Selection of a target **phenomenon**,
3. An **observation** of the selected phenomenon,
4. An **evaluation** of what is being **observed**, with respect to the chosen reference point,
5. A **value judgement** of the results of this **evaluation**, whose result determines whether the phenomenon being observed is **biased**.

How is Bias constructed? - A first example

1. **Observer** = South European Vegan developer
2. **Phenomenon =** ML algorithm trained on a dataset coming from Mongolia
3. **Observation** = ML algorithm tested on a South European dataset
4. **Evaluation** = Unexpected, far from reference point (Mongolian cuisine is meat intensive)
5. **Value judgement** = ML algorithm is **biased**

NOTE: In process above change “ML algorithm trained on a dataset” with “local expert”

How is Bias constructed? - A second example

1. **Observer** = a Mongolian developer
2. **Phenomenon =** ML algorithm trained on a dataset coming from South European
3. **Observation** = ML algorithm tested on a South European dataset
4. **Evaluation** = Unexpected, far from reference point
5. **Value judgement** = What's the point of the South European Dataset? (because of weather, very limited availability of vegetables)

NOTE (again): In process above change "ML algorithm trained on a dataset" with "local expert"

How is Bias constructed? - example summing up

- Both the South European Vegan developer and the Mongolian meat-eater developer are well-minded,
- Both are with good motivations and good intentions,
- Making them understand each other's perspective would shft their perception from **inessential diversity essential diversity/similarity**,
- The perception of bias would most likely transform itself in a perception of diversity (which becomes known from being previously unknown),
- Then, both of them could independently decide on their interest in the other's **essential diversity/ similarity,**

QUESTION: Can we implement this process in the Web?

Bias? Or just Diversity? (… as from before)

Bias – a contextual social property

- **Who decides** what is closed-minded, prejudicial, or unfair?
- **Who decides** what is disproportionate?

Diversity – a contextual social property

- **What is** the reference neutral point?
- **What are** the dimensions and metrics of evaluation?

NOTE:

- Diversity seems correlated to Bias while being somehow more objective
- It seems possible to formalize the process of perception diversity (and similarity)

Putting Bias in perspective

1. Define a *Diversity space.* In the above example,
   - the diversity dimensions would be the type of food (e.g., vegetables, meat, noodle, etc.) and various other characteristics.
   - The origins, assumed to define of what is expected, of the South-European and of the Mongolian would be quite far apart.
2. The distance from the origin of any phenomenon is a measure of deviation from the expected behavior.
3. Within this diversity space, a certain volume could be defined as being the *Bias space*
4. Any observed phenomenon could then be defined as being (un)biased depending on whether it is (out)inside the bias space.

Transparency paths

1. *Diversity transparency path*: the steps by which diversity is identified:
    1. Select phenomenon to be observed
    2. Select properties/ features
    3. Select metrics
    4. Select origin (= normality)
    5. Measure phenomenon

2. *Bias Transparency path:*
    - 1-5: Diversity transparency path
    - *6: value judgement*

NOTE:

- Transparency paths: first step towards self-reflection/ assessment.
- Transparency paths: from inessential to essential diversity/similarity.
- Final decision still up to the observer.
- Value judgement external to phenomenon being observed

A first experiment (FairUMAP – submitted)

Question: *Can the inherent diversity in a dataset used to represent a phenomenon in the world become a source of bias?*

Overall Answer: *yes*, referring to the need for *awareness of the alternative views*, and bias being (caused by) the *absence of diversity*.

Two participants: referred to the unequal *distribution of representations* of cases in the dataset.

One participant: referred to the definition of bias as being *context dependent*

A first experiment (cont) (FairUMAP – submitted)

Task: *Overall Reflection in this Process by the participants*

With respect to constructing their own datasets from two different sources on two different phenomena, participants thought that we should not have used an individual process. They stressed the importance of different points of view when such tasks are taking place – "... Bias is never individual, even if it is executed individually, our choices are socially constructed (our focus from a professional background, our geographical location, etc) ..."

"I realized that from the beginning we decided on how to select photos based on some assumptions. Then when we were characterizing the images we "saw" only those characteristics that we were looking for and not the complete picture. This is how probably biased datasets are built."

A different reflection is that the exercise could be a task for understanding how bias and diversity might be transferred in AI enabled systems through datasets – "I think of this exercise as a self-reflection exercise. It enabled us to reflect on concepts like diversity and bias, and also become more aware of the fact that any dataset aimed at visual depictions of world phenomena that we would create would somehow be biased."

Empowering diversity-aware social relations

(WeNet – The Internet of Us)

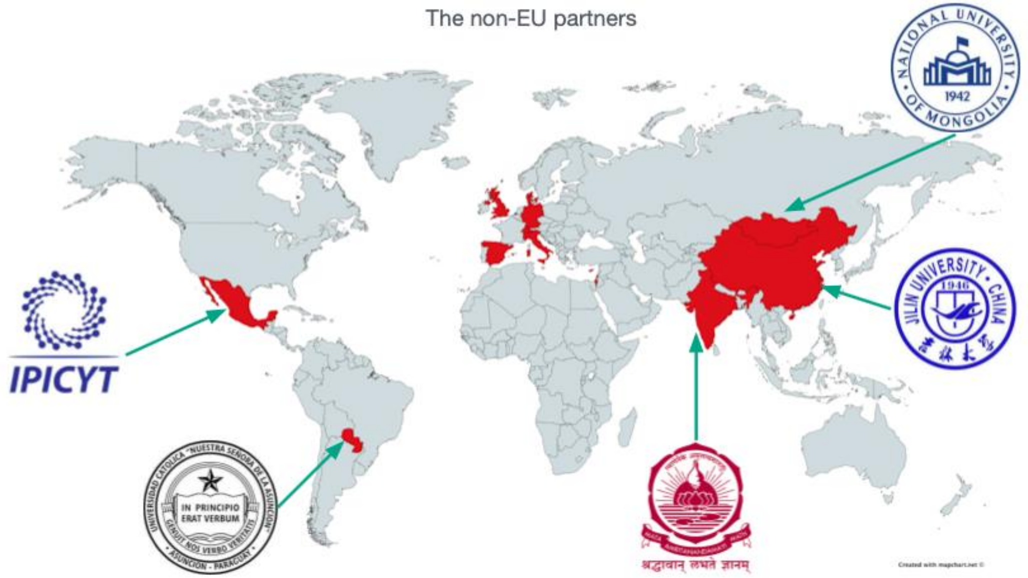
WENET INTERNET OF US

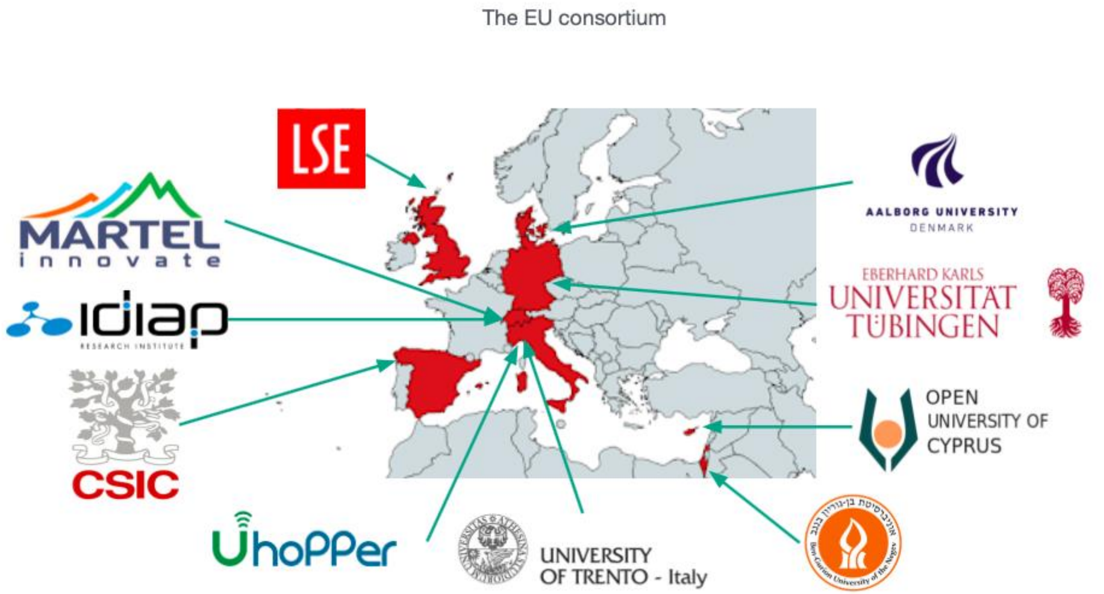
CY. center for algorithmic transparency

The WeNet Goal in a nutshell

Empowering diversity aware social relations

Create a platform connecting people, including ***unknown*** people, towards satisfying need by leveraging their (essential) diversity



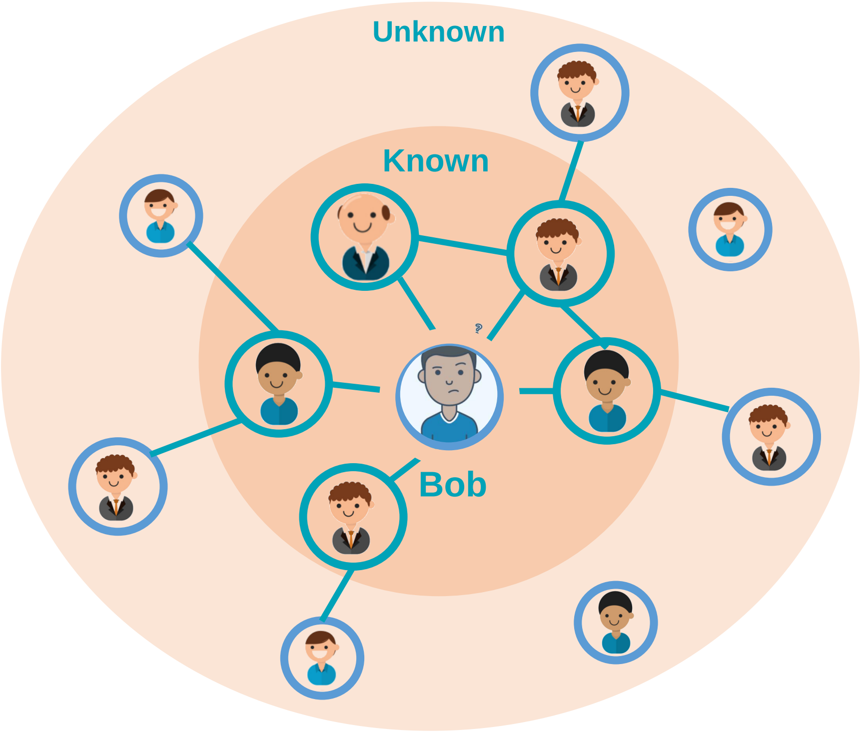


Experiment and data driven
Technological Innovation informed by *Social Innovation*

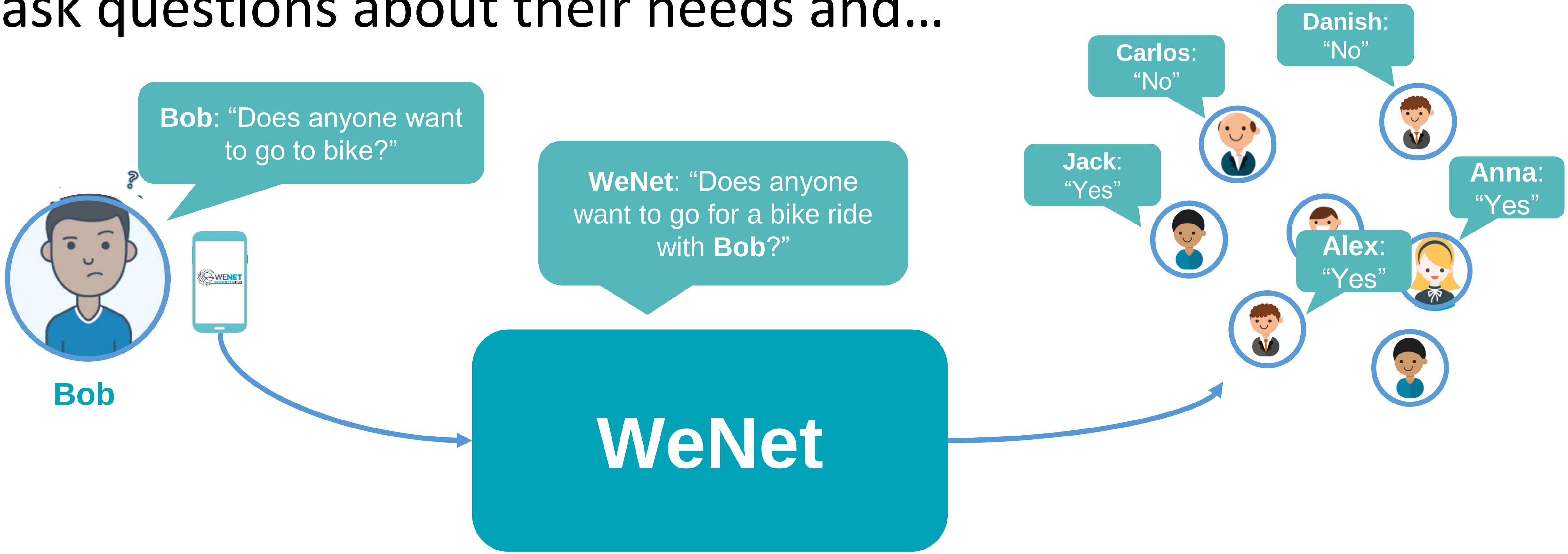
The WeNet platform

Will be used in apps for social interactions, that allows people to search for the proper person in order to satisfy their needs.

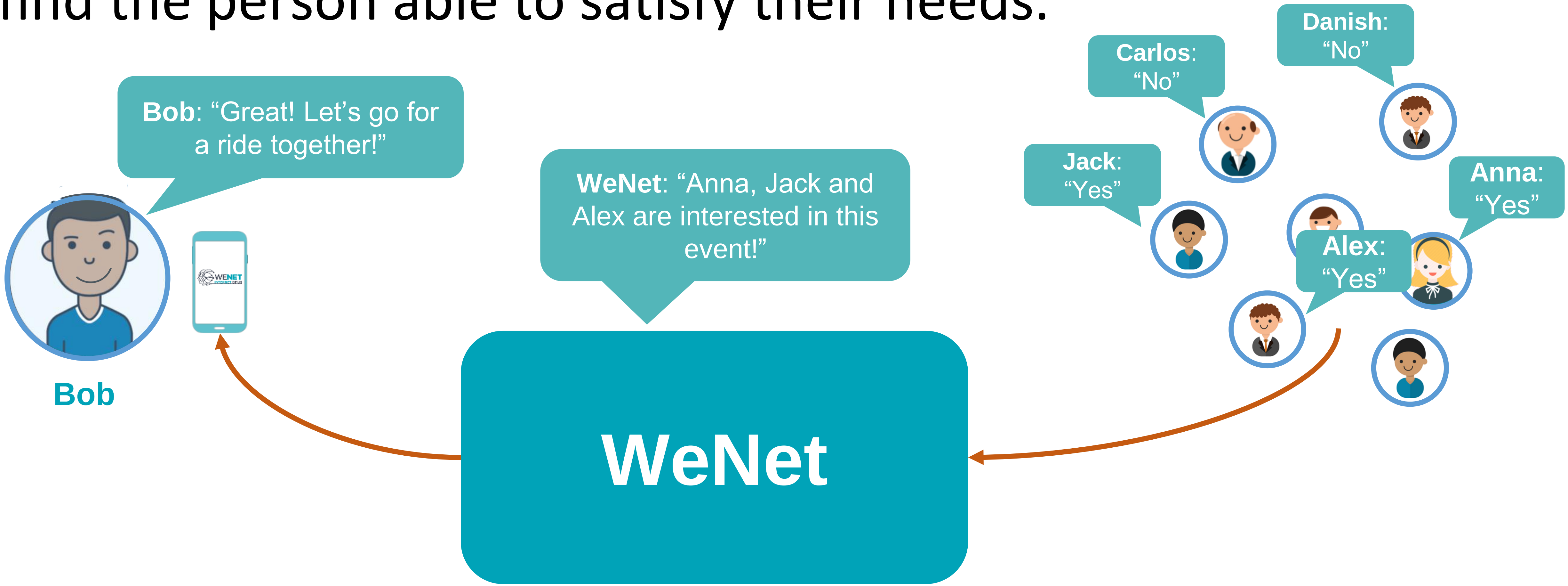
Thanks to the apps, a person could…



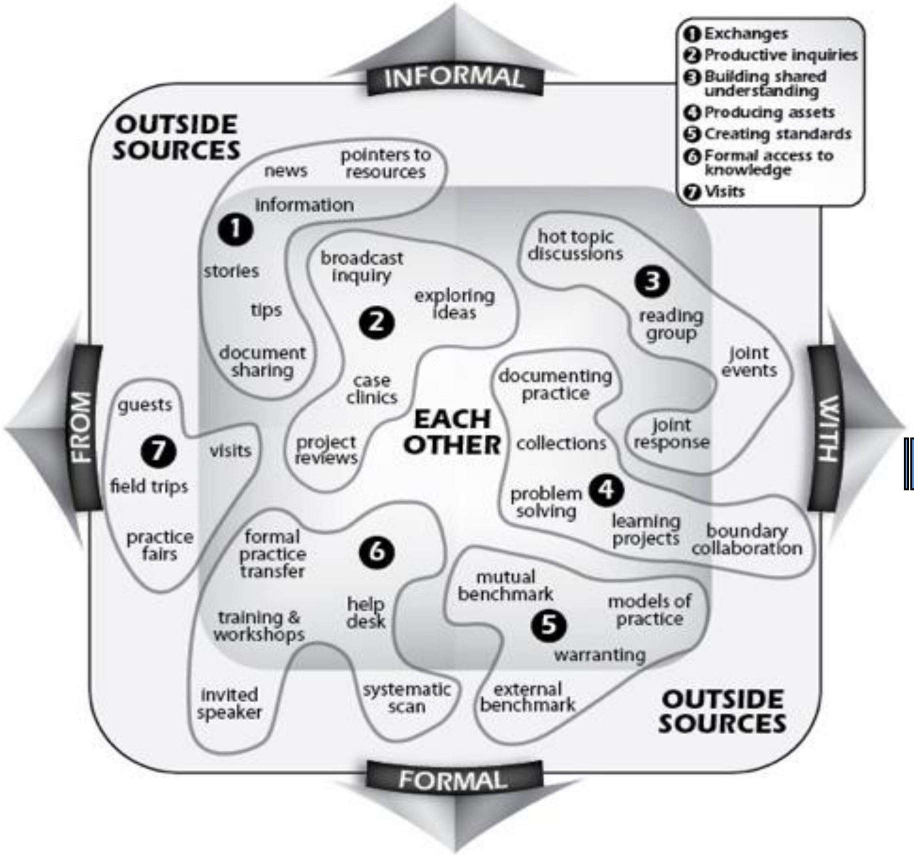
…ask questions about their needs and…



…find the person able to satisfy their needs.



Types of queries (from social sciences)



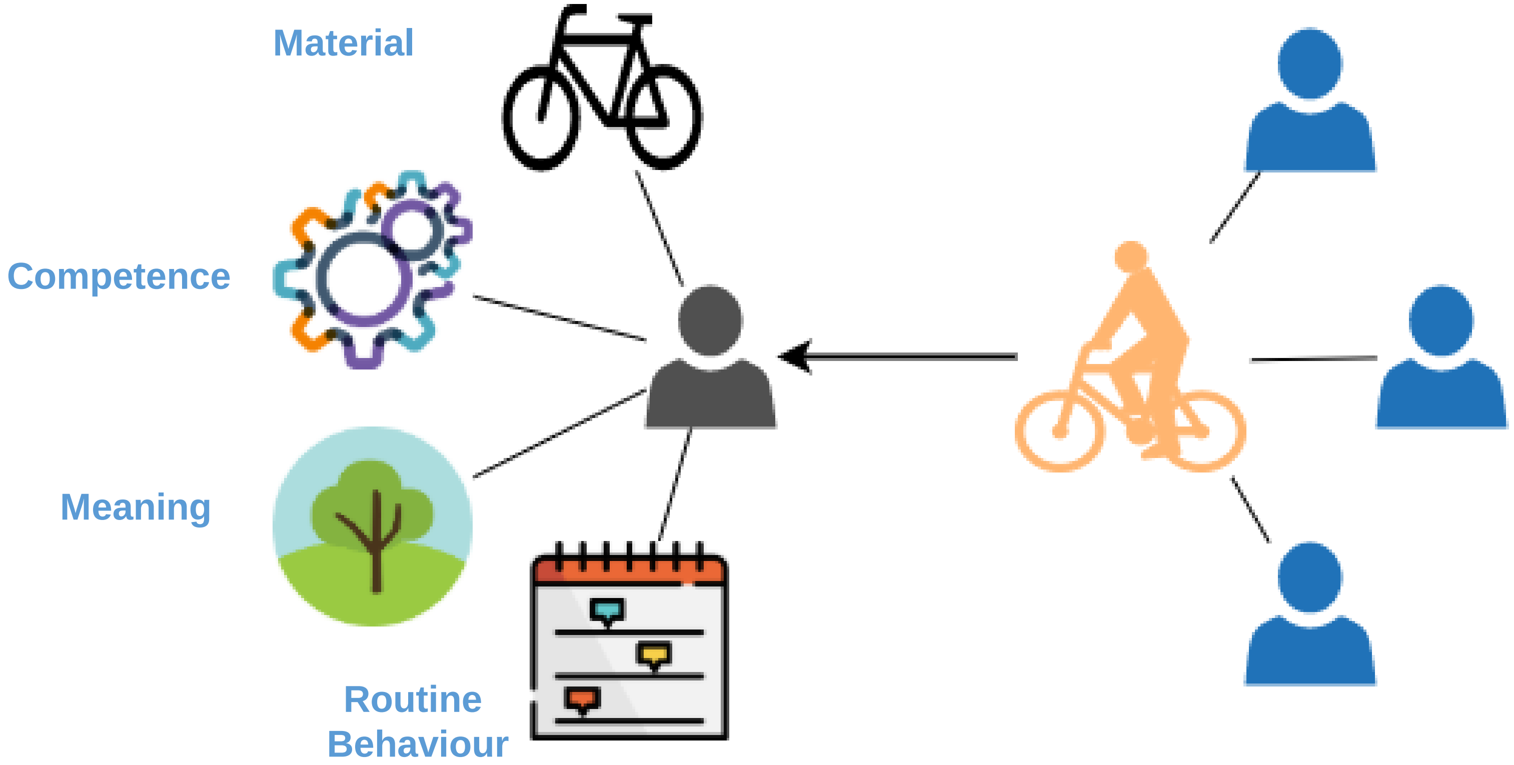
Problem solving	"Can we work on this design and brainstorm some ideas; I'm stuck."
Requests for information	"Where can I find the code to connect to the server?"
Seeking experience	"Has anyone dealt with a customer in this situation?"
Reusing assets	"I have a proposal for a local area network I wrote for a client last year. I can send it to you and you can easily tweak it for this new client."
Coordination and synergy	"Can we combine our purchases of solvent to achieve bulk discounts?"
Discussing developments	"What do you think of the new CAD system? Does it really help?"
Documentation projects	"We have faced this problem five times now. Let us write it down once and for all."
Visits	"Can we come and see your after-school program? We need to establish one in our city."
Mapping knowledge and identifying gaps	"Who knows what, and what are we missing? What other groups should we connect with?"

Which (essential) diversity?

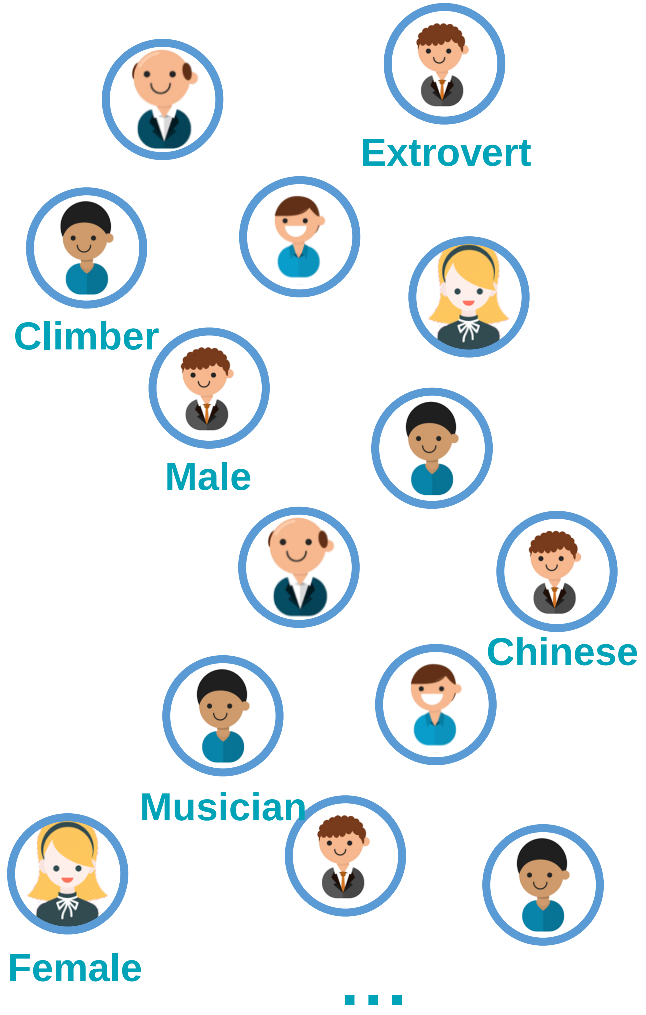
Social Practices
(Shove et al., 2012)

Habits
(Van Krieken, 2003)
Habitus
(Bourdieu, 1985)

Communities of practice
(Wenger, 2005)



In order to select the correct person, the WeNet app capture the diversity of people in both their visible and deep characteristics. In other words, the platform not only considers the gender, age or nationality traits, but also the abilities of a person, their interests, their values and beliefs, their culture.



Diverse in:

Surface
Gender
Nationality
Age
…

Deep
Personality
Culture
Competence
…

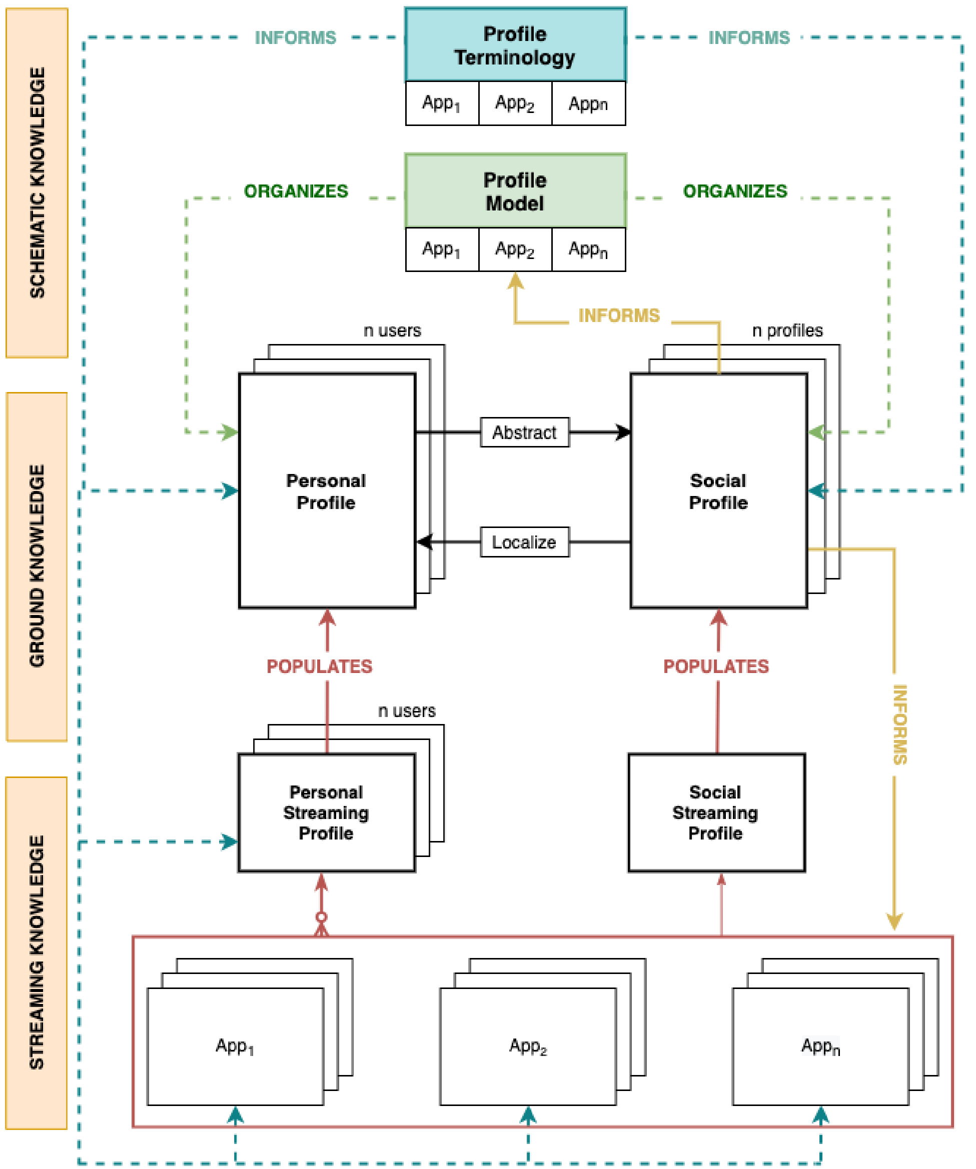
In this way the app is able to select the people who can answer the question, favouring the deep (essential) aspects of their **diversity**. In this way the app will not discriminate on the basis of non-essential characteristics.

- ✓ Ride her bike every week
- ✓ Owns a gravel
- ✓ She loves to watch the Giro d'Italia

- ✓ Go to downhill once a week
- ✓ Owns a MTB and a city bike
- ✓ Knows many bike riders

- ✓ Play basketball
- ✓ Goes to the gym once a week
- ✓ Likes to watch sports on TV

The profile


SCHEMATIC KNOWLEDGE
GROUND KNOWLEDGE
STREAMING KNOWLEDGE
Profile Terminology
App1 App2 Appn
INFORMS
INFORMS
Profile Model
App1 App2 Appn
ORGANIZES
ORGANIZES
INFORMS
n users
n profiles
Personal Profile
Abstract
Localize
Social Profile
POPULATES
POPULATES
INFORMS
n users
Personal Streaming Profile
Social Streaming Profile
App1
App2
Appn

Empowering diversity

(whats next?)

WENET INTERNET OF US

CY. center for algorithmic transparency

The Web in Perspective (to morrow)

Phase 1 – Enterprise systems (Closed world, homogeneous data)

Problem: ***Data diversity is a bug.*** Data diversity is a design and/ or implementational mistake.
Solution (Software Engineering): Design time absorption of diversity + robustness of the design.

Phase 2 – The Web of data (Open world, heterogeneous data)

Problem: ***Data diversity is unavoidable.*** Data diversity is a consequence of independent design
Solution (The Semantic Web): Design time + Run time absorption of diversity via reference ontologies.

Phase 3 – The Web of people (Everybody, everywhere, everytime, people diversity)

Problem: ***Data diversity is a feature that we do not know yet how to handle.*** It is the welcome byproduct of the inherent diversity of people
Solution (Human-centric AI): **A move from a network of computers, which in turn may be connected to people, to a diversity aware network of people, whose interactions are mediated and empowered by computers.**

My person to the moon

A common view

The ultimate success of ***Artificial Intelligence*** would be to make it possible for everybody to exploit, when trying to achieve her objectives, the most suitable ***machine*** (whose Intelligence exceeds human intelligence) with the best expertise available in the planet.

My view

The ultimate success of ***Artificial Intelligence*** would be to make it possible for everybody to exploit, when trying to achieve her objectives, the most suitable ***person*** with the best expertise available in the planet.

Each person would be augmented with the people's knowledge and skills available in the world.

UNIVERSITY OF TRENTO - Italy

Dipartimento di Ingegneria e Scienza dell'Informazione

Know dive

The need for inter-disciplinarity (WeNet)

Methodology	Concepts	Analysis	Profile	Ethics	GDPR
Weisberg, H. F. (2009). *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press.	Shove, E., Pantzar, M., & Watson, M. (2012). *The dynamics of social practice: Everyday life and how it changes*. Sage.	Abbott, A. (1995). Sequence analysis: New methods for old ideas. *Annual review of sociology*, *21*(1), 93-113.	Scekic, O., Miorandi, D., Schiavinotto, T., Diochnos, D. I., Hume, A., Chenu-Abente, R., ... & Giunchiglia, F. (2015, October). SmartSociety–A Platform for Collaborative People-Machine Computation. In *2015 IEEE 8th International Conference on Service-Oriented Computing and Applications (SOCA)* (pp. 147-154). IEEE.	Schelenz, L., Segal, A., & Gal, K. (2020). Best Practices for Transparency in Machine Generated Personalization. *arXiv preprint arXiv:2004.00935*.	General Data Protection Regulation (2016)
Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2019). *Managing and sharing research data: a guide to good practice*. SAGE Publications Limited.	Wenger, E. (2009). Communities of practice. *Communities*, *22*(5).	Giunchiglia, F., Zeni, M., & Big, E. (2018, March). Personal context recognition via reliable human-machine collaboration. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (pp. 379-384). IEEE.	Giunchiglia, F., & Fumagalli, M. (2014, November). From ER models to the entity model. In *International Conference on Knowledge Engineering and Knowledge Management* (pp. 116-119). Springer, Cham.	Schelenz, L., Reinhardt, K., & Gjuraj, D. (2019). D9. 1 DEVELOPING A CONCEPTUAL AND ETHICAL FRAMEWORK OF DIVERSITY.	WP Article 29 (2014, 2017)
Zeni, M. (2017). *Bridging Sensor Data Streams and Human Knowledge* (Doctoral dissertation, University OF Trento).	Bourdieu, P. (1990). Structures, habitus, practices. *The logic of practice*, 52-65.	Edwards, J. R. (2003). Construct validation in organizational behavior research.	Giunchiglia, F., Bignotti, E., & Zeni, M. (2018). Human-Like Context Sensing for Robot Surveillance. International Journal of Semantic Computing, 12(01), 129-148.	Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399.	Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; et al. (15 Marc 2016). "The FAIR Guiding Principles for scientific data management an stewardship". Scientific Data. 3: 160018.

	Social Sciences
	Ethics and GDPR
	Interdisciplinary
	Computer Science

Diversity, Bias, online social relations and related issues

Thank you!

Fausto Giunchiglia

19.04.2021