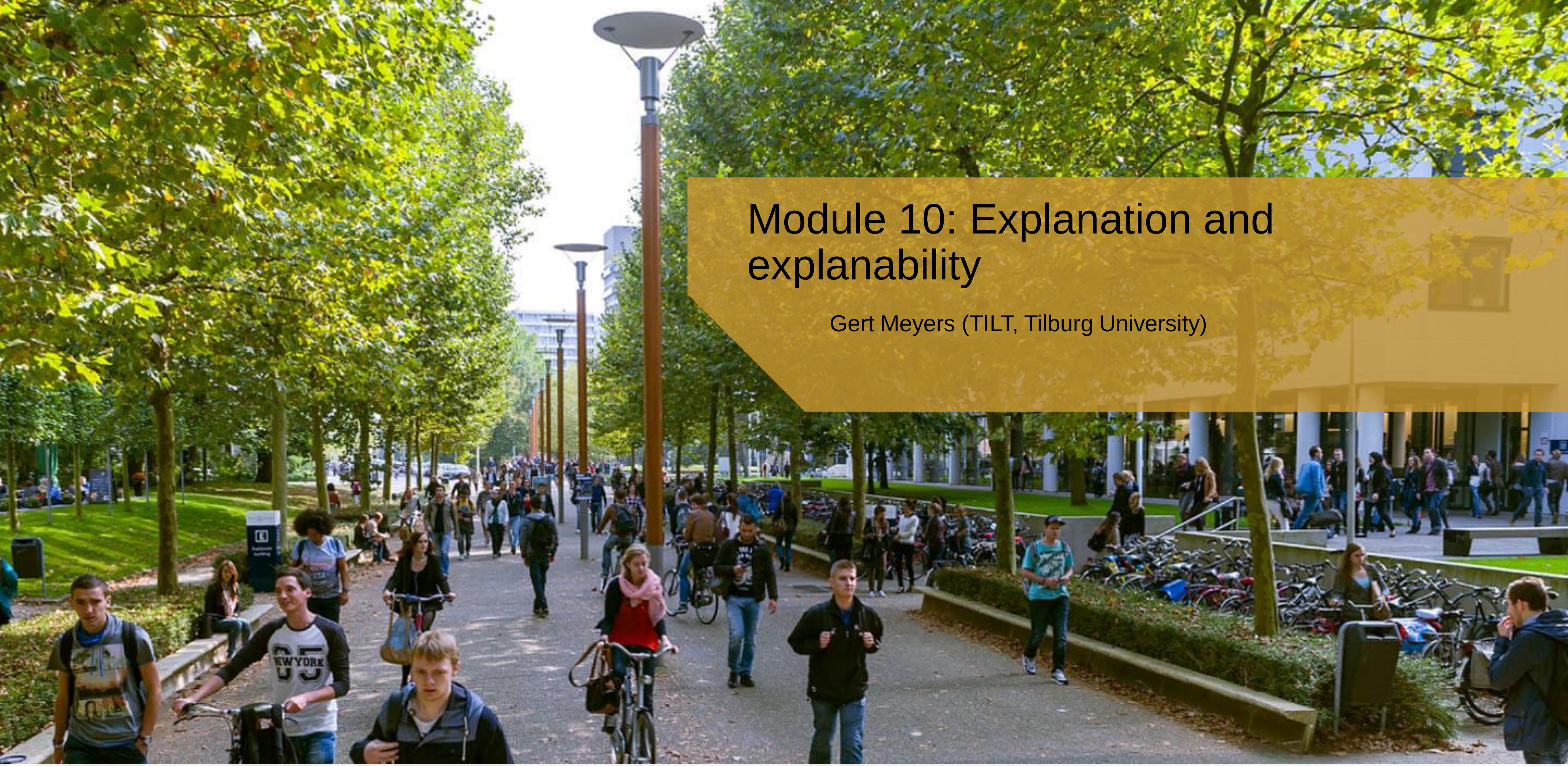


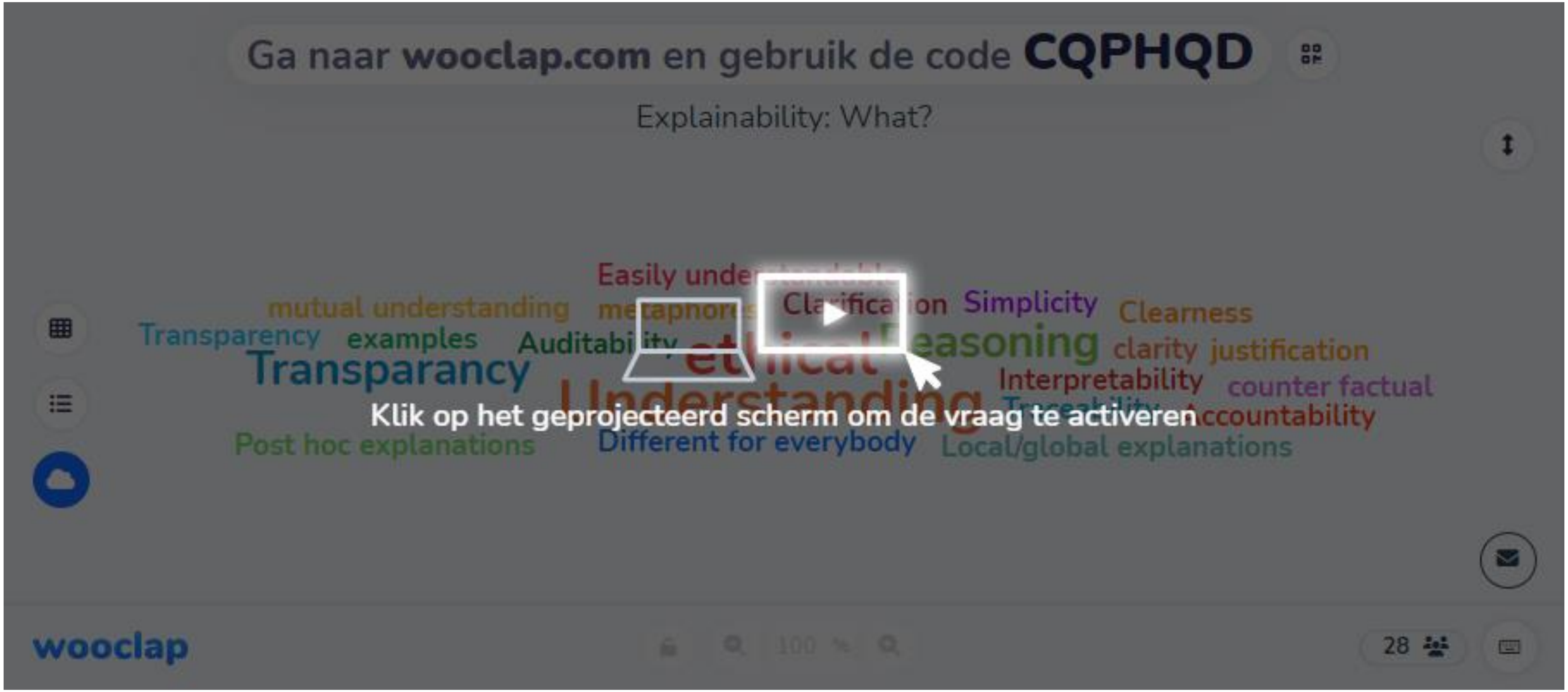
Module 10: Explanation and explanability

Gert Meyers (TILT, Tilburg University)

This week

- Explanation and explainability
- Third debating session!

- The blogpost



Ga naar **wooclap.com** en gebruik de code **CQPHQD**

Explainability: What?

Klik op het geprojecteerd scherm om de vraag te activeren

Ga naar **wooclap.com** en gebruik de code **CQPHQD**

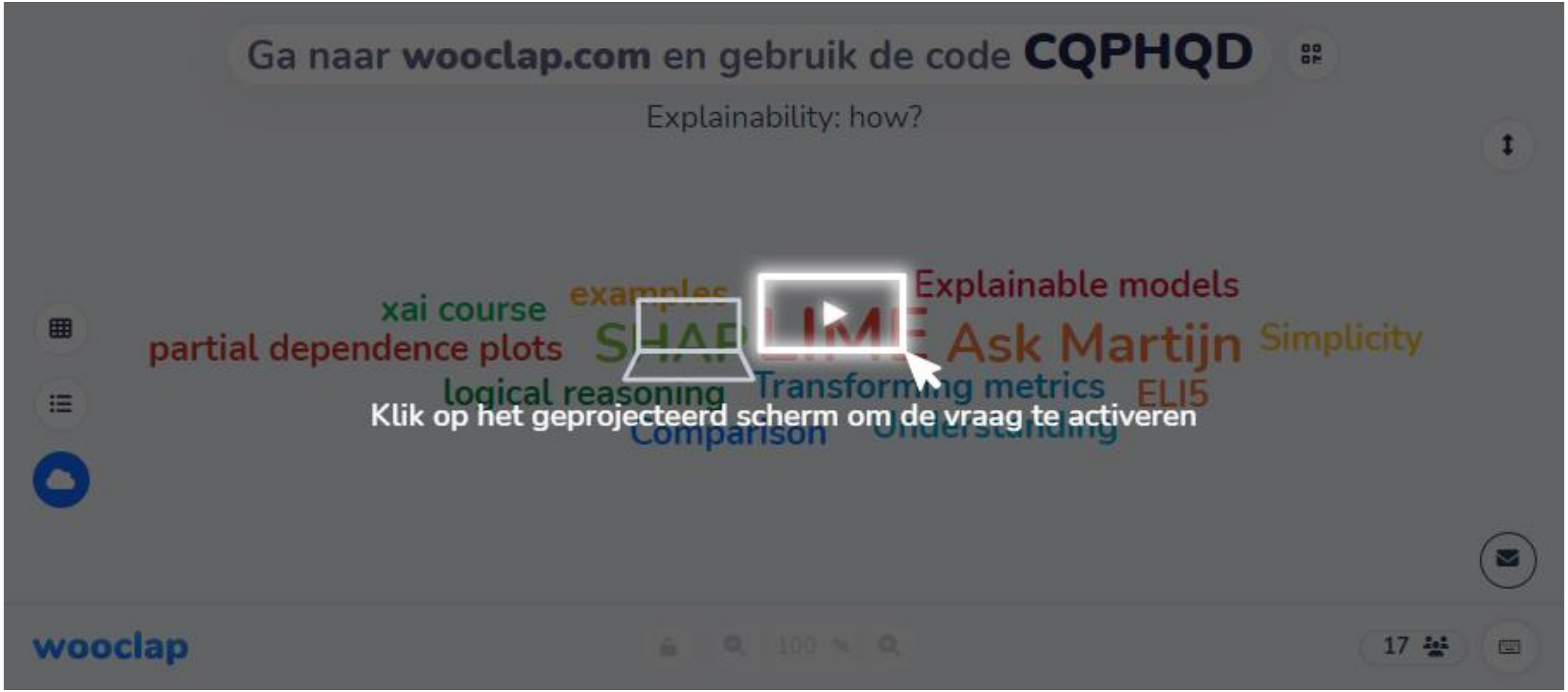
Explainability: Why?

Tijd om te stemmen!

30

antwoorden gekregen

Klik op het geprojecteerd scherm om de vraag te activeren



Ga naar **wooclap.com** en gebruik de code **CQPHQD**

Explainability: how?

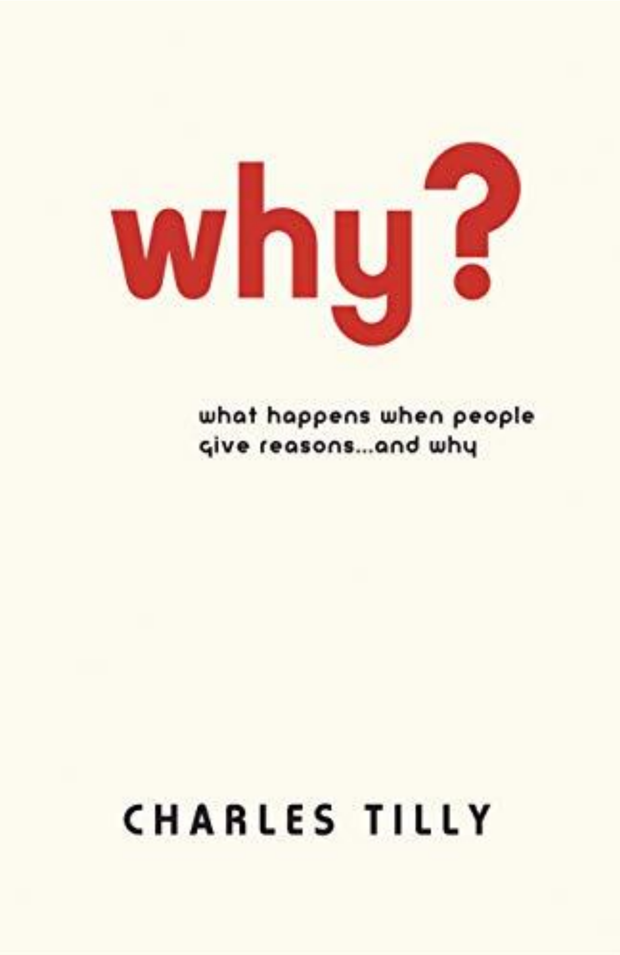
Klik op het geprojecteerd scherm om de vraag te activeren

MODULE 8: TRUST AND TRUSTWORTHINESS

MODULE 9: AUTONOMY

MODULE 10: EXPLANATION AND EXPLAINABILITY

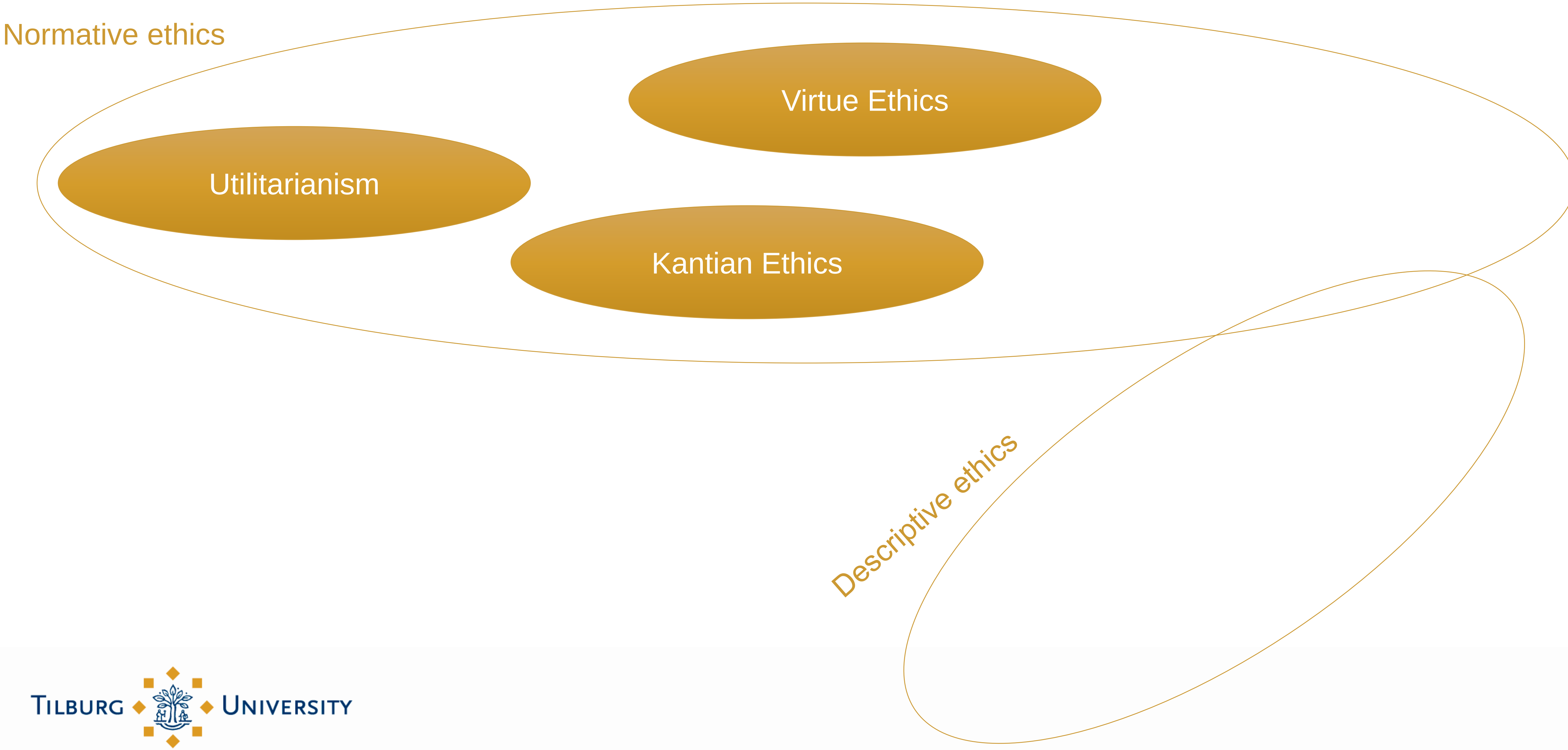
- Explanation
- Transparancy
- Accountability
- Justification



why?

what happens when people give reasons...and why

CHARLES TILLY



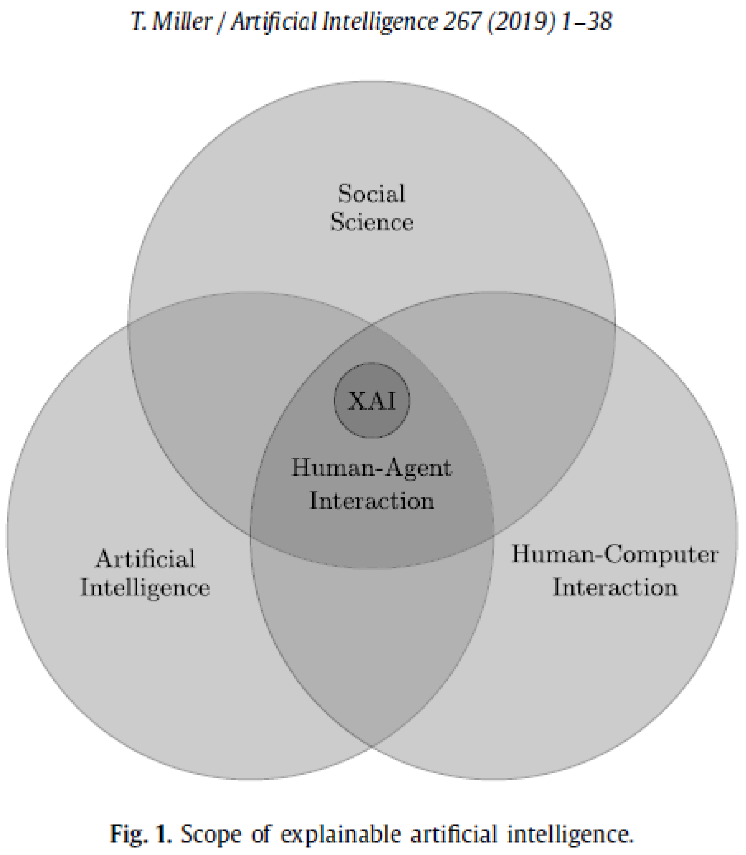
Normative ethics

Virtue Ethics

Utilitarianism

Kantian Ethics

Descriptive ethics



T. Miller / Artificial Intelligence 267 (2019) 1–38

Fig. 1. Scope of explainable artificial intelligence.

Note: ‘Everyday explanation’
vs ‘general explanation’

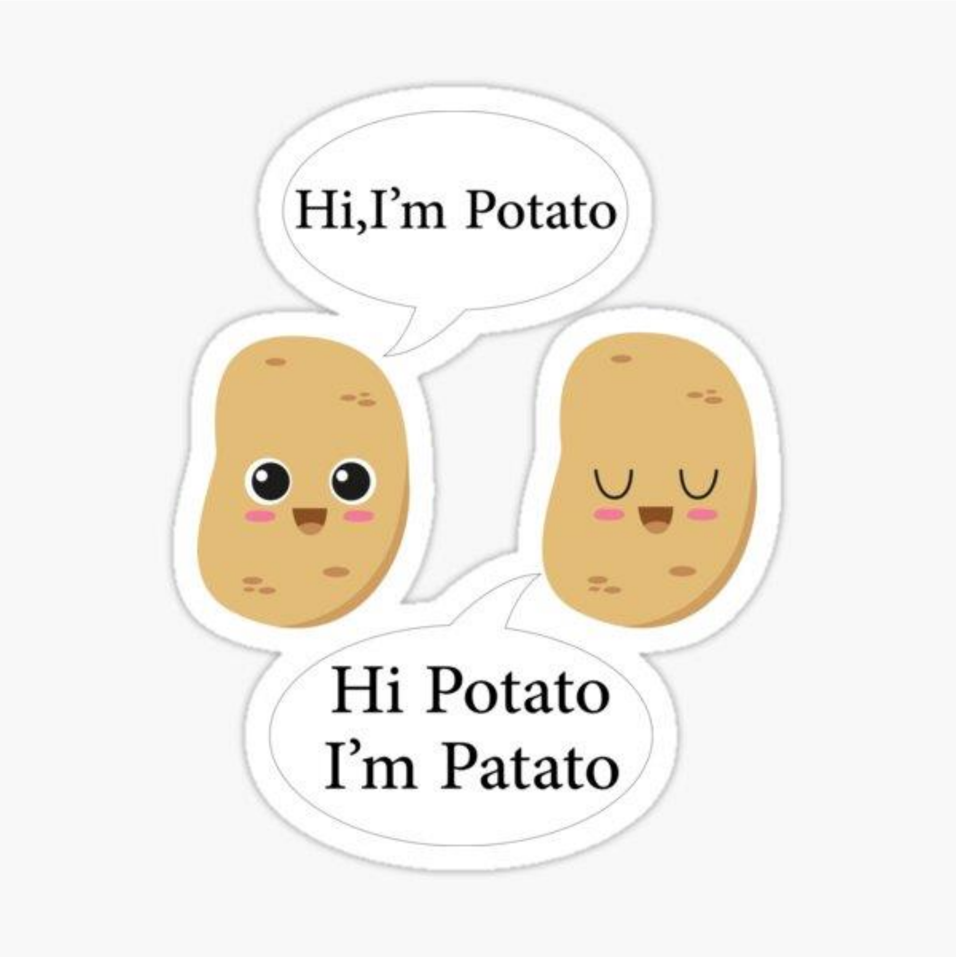
Note: decisions

Happy Birthday!

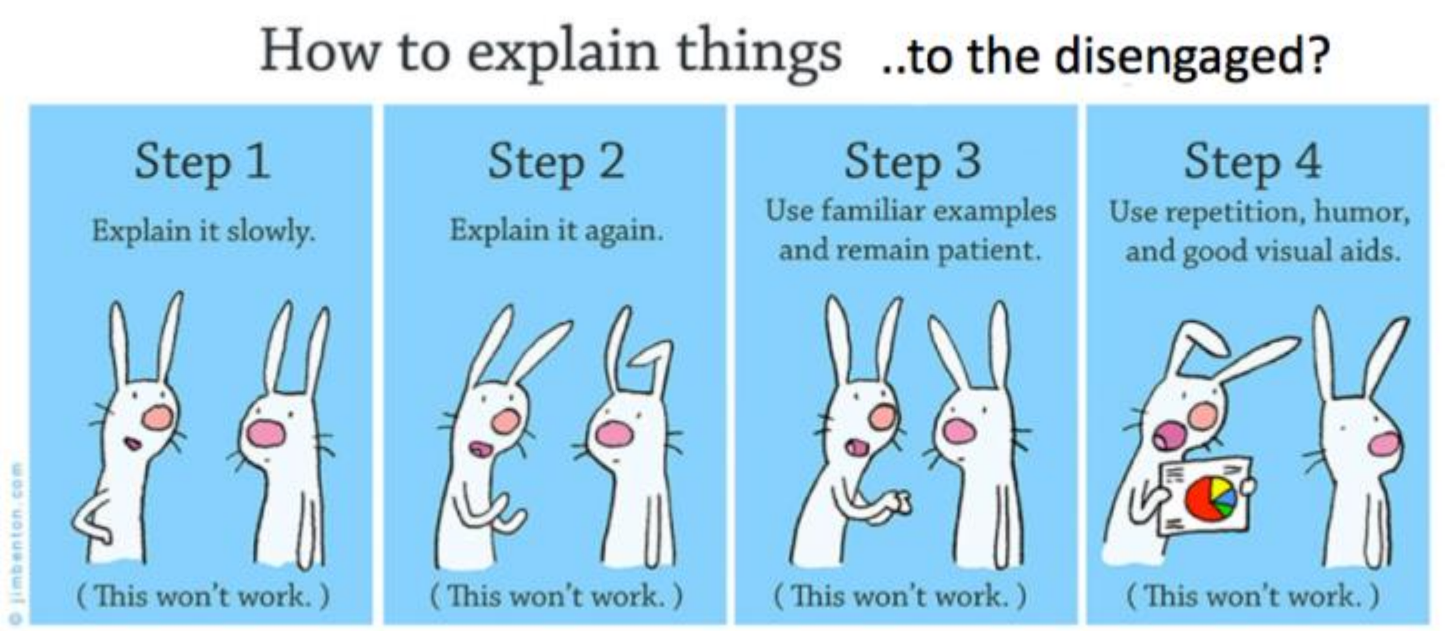
Immanuel Kant: born on 22 april 1724

Explanation as product and process (twice)

- Cognitive process: identifying the causes of a particular phenomenon
- Product: the explanation
- Social process: transferring knowledge between explainer and explainee

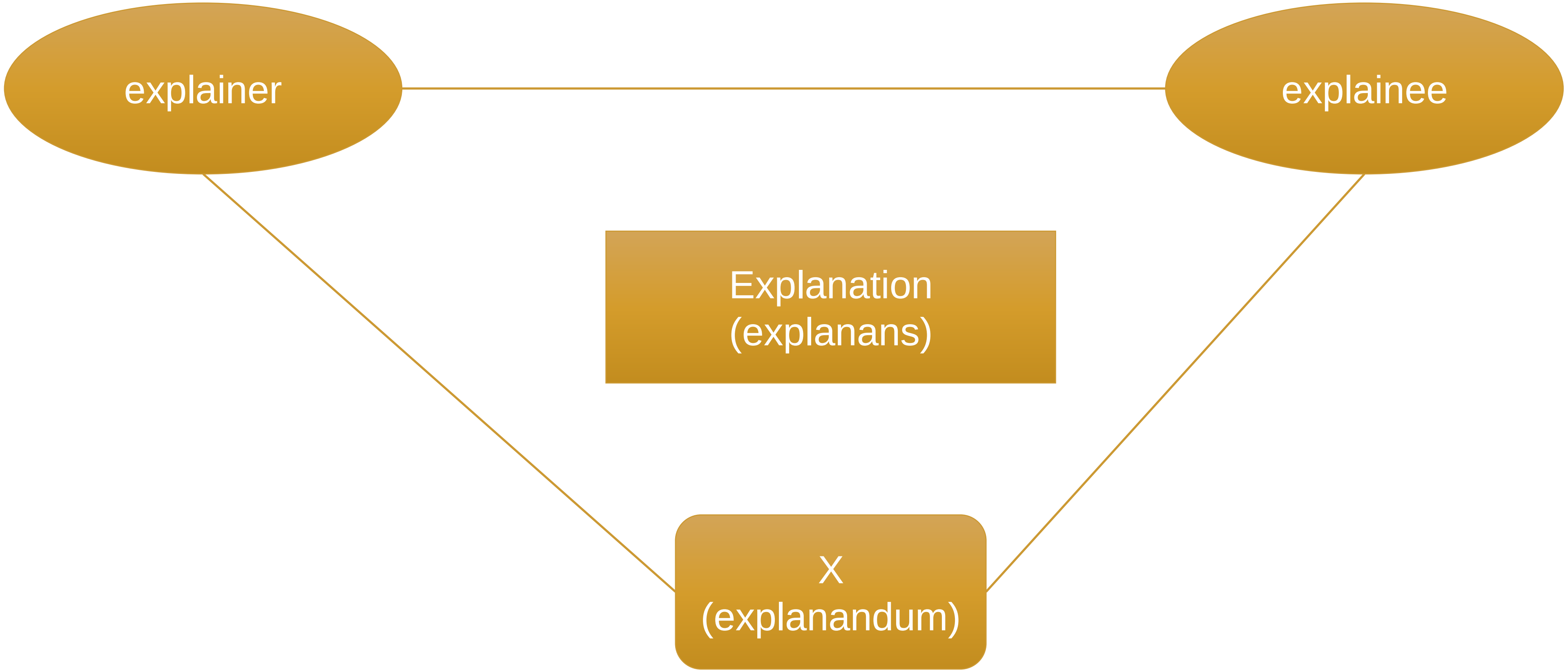


explain



e

Note: no arrows



Note: no arrows

trustee

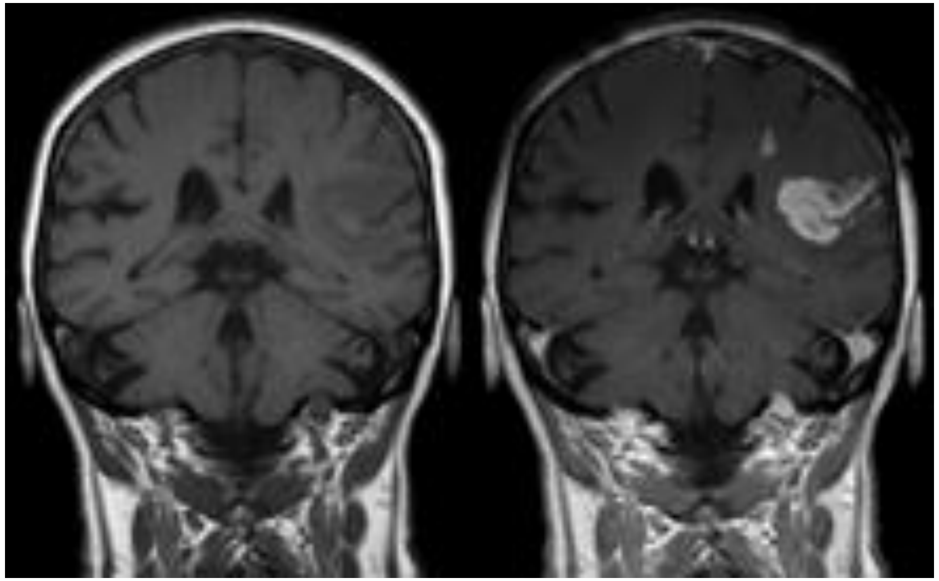
trustor

Note: no arrows

Explanations are contrastive

- Counterfactual cases
- Why this instead of that?
- Layperson will find contrastive explanations more intuitive and more valuable

- Science girl op X: 'Explain this! https://t.co/7izH4Hbivz' / X (twitter.com)



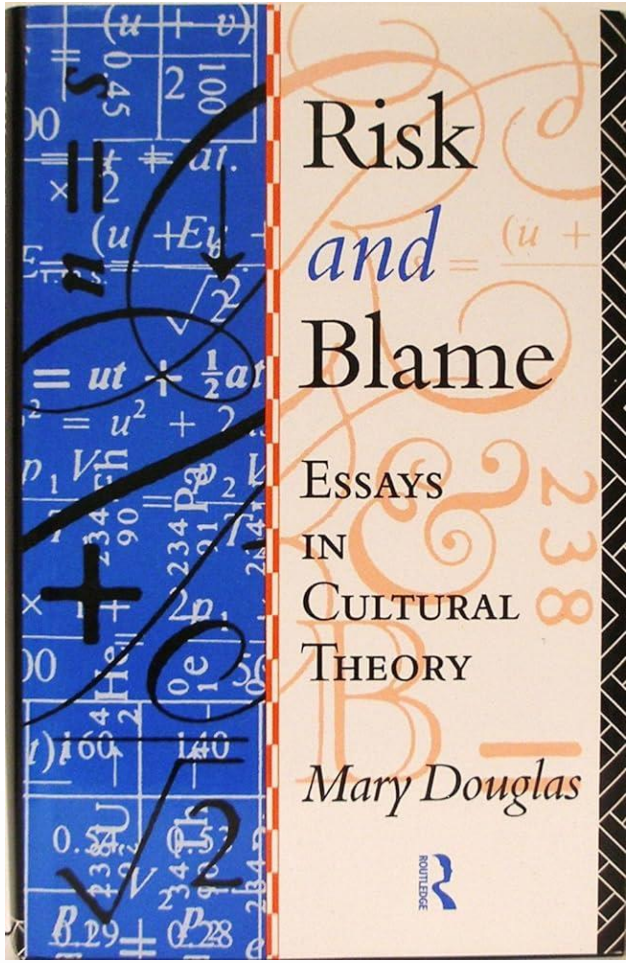
Explanations are selected

- Humans are adept at selecting one ore two causes from a sometimes infinte number of cause to be *the* explanation
- Cognitive burden of complete explanations is too great

Probabilities probably don’t matter

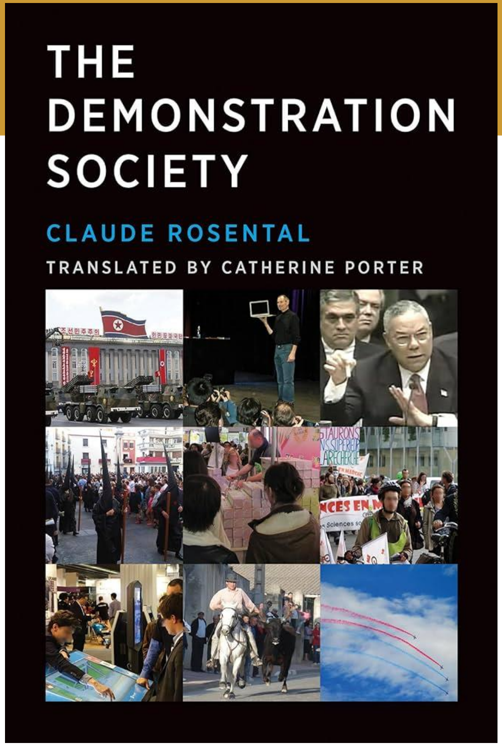
- Reffering to probabilities vs referring to causes
- The most likely explanation is not always the best explanation for a person

- Example: collapse of the soviet union



Explanations are social

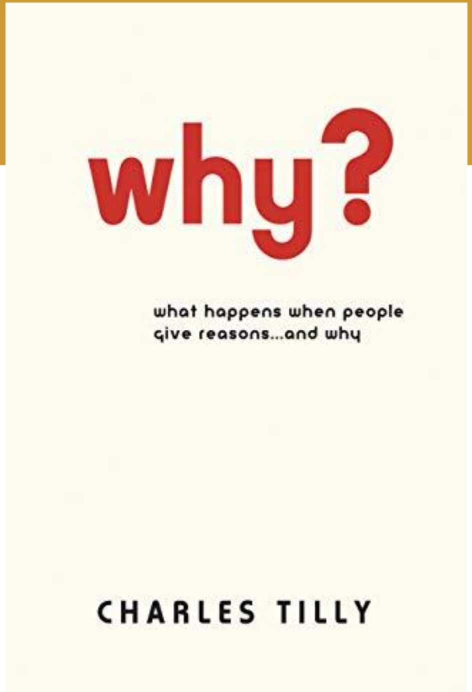
- Transfer of knowledge
- Part of an interaction
- Erving Goffman: ‘definition of the situation’



THE DEMONSTRATION SOCIETY

CLAUDE ROSENTAL

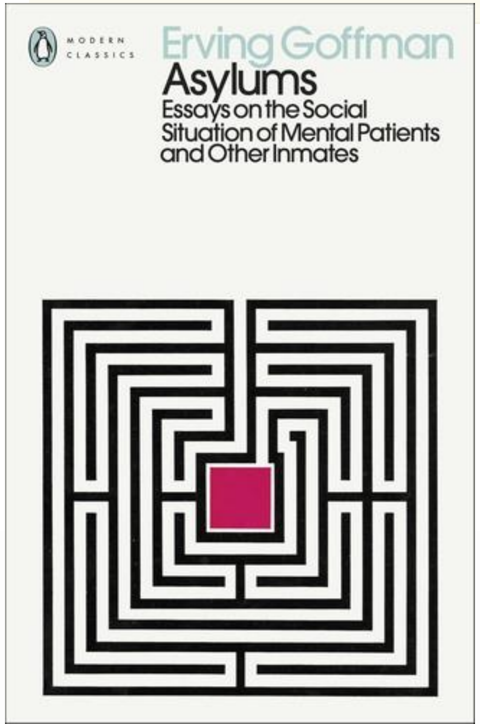
TRANSLATED BY CATHERINE PORTER



why?

what happens when people give reasons...and why

CHARLES TILLY



MODERN CLASSICS

Erving Goffman

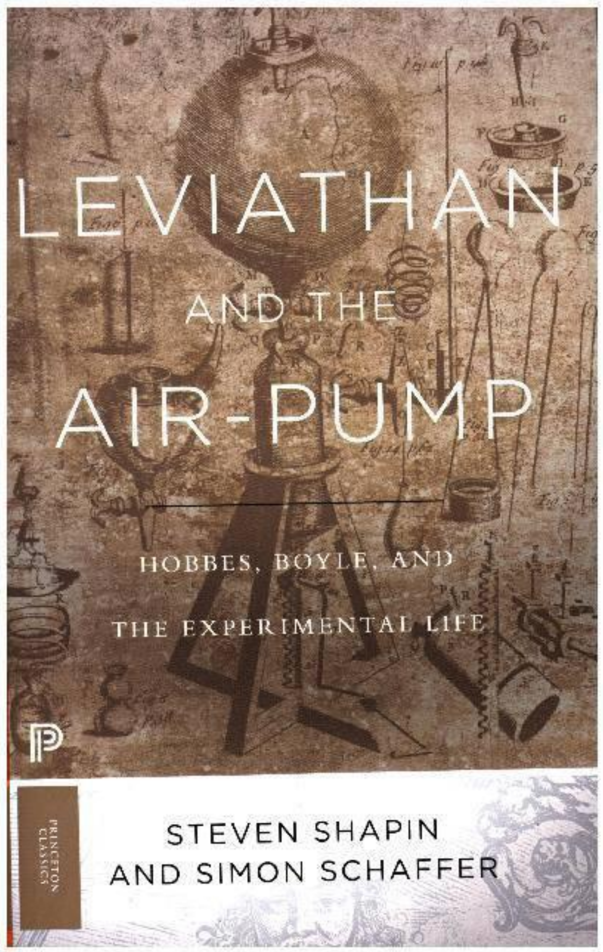
Asylums

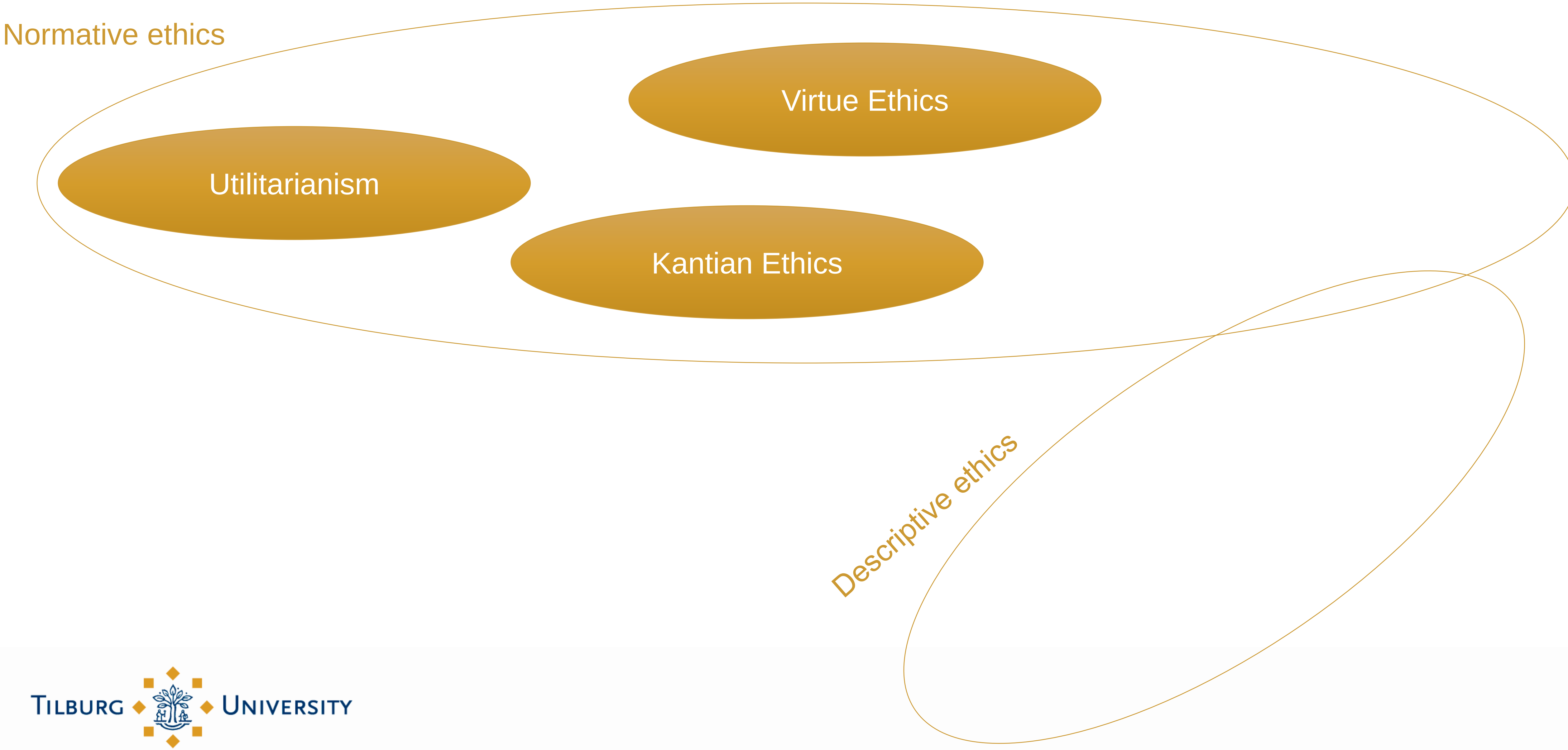
Essays on the Social Situation of Mental Patients and Other Inmates

Explanation

- Contrastive
- Selected
- Propabilities probably don’t matter
- Social

- CONTEXTUAL
  - An explanation is not a (mathemathical) proof
  - ‘the best’ explanation does not always work





Normative ethics

Virtue Ethics

Utilitarianism

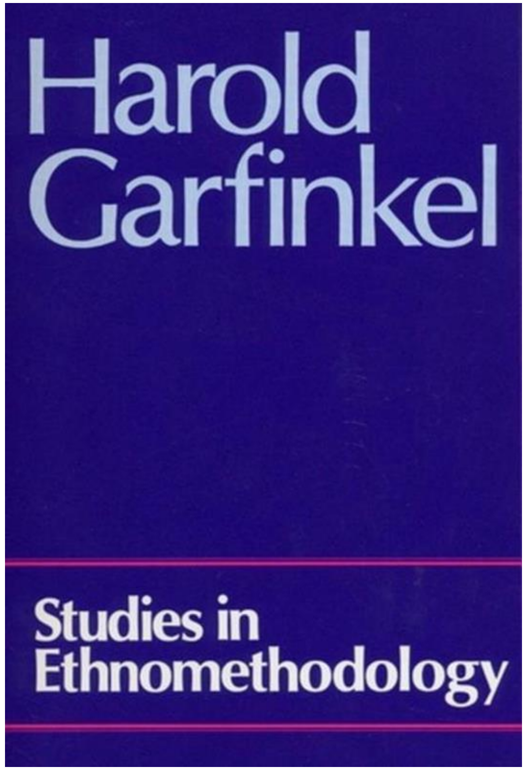
Kantian Ethics

Descriptive ethics

Providing (observable) accounts

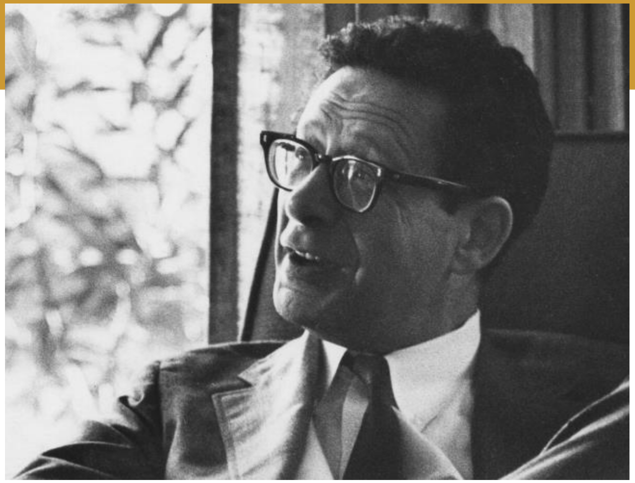
Accountability

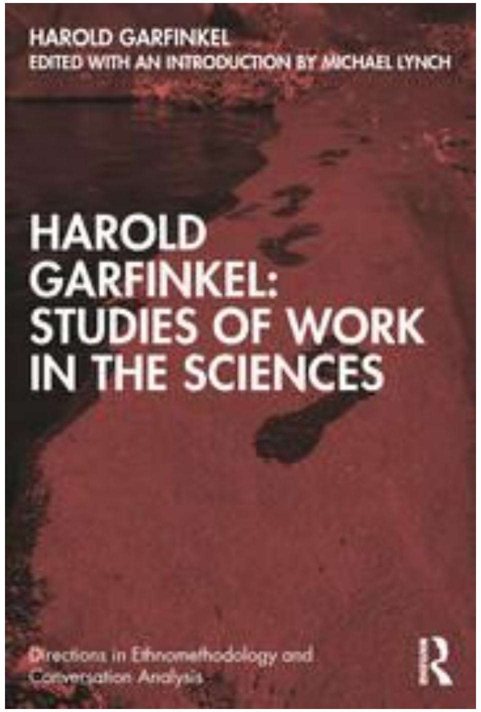
- Actors continuously provide observable accounts of what they are doing (ethnomethods)
- Often implicit
  - Et cetera clause
  - For ex. Skipping slides because lack of time
- ADM are decision that take place in social context, but opacity problem



Harold Garfinkel

Studies in Ethnomethodology





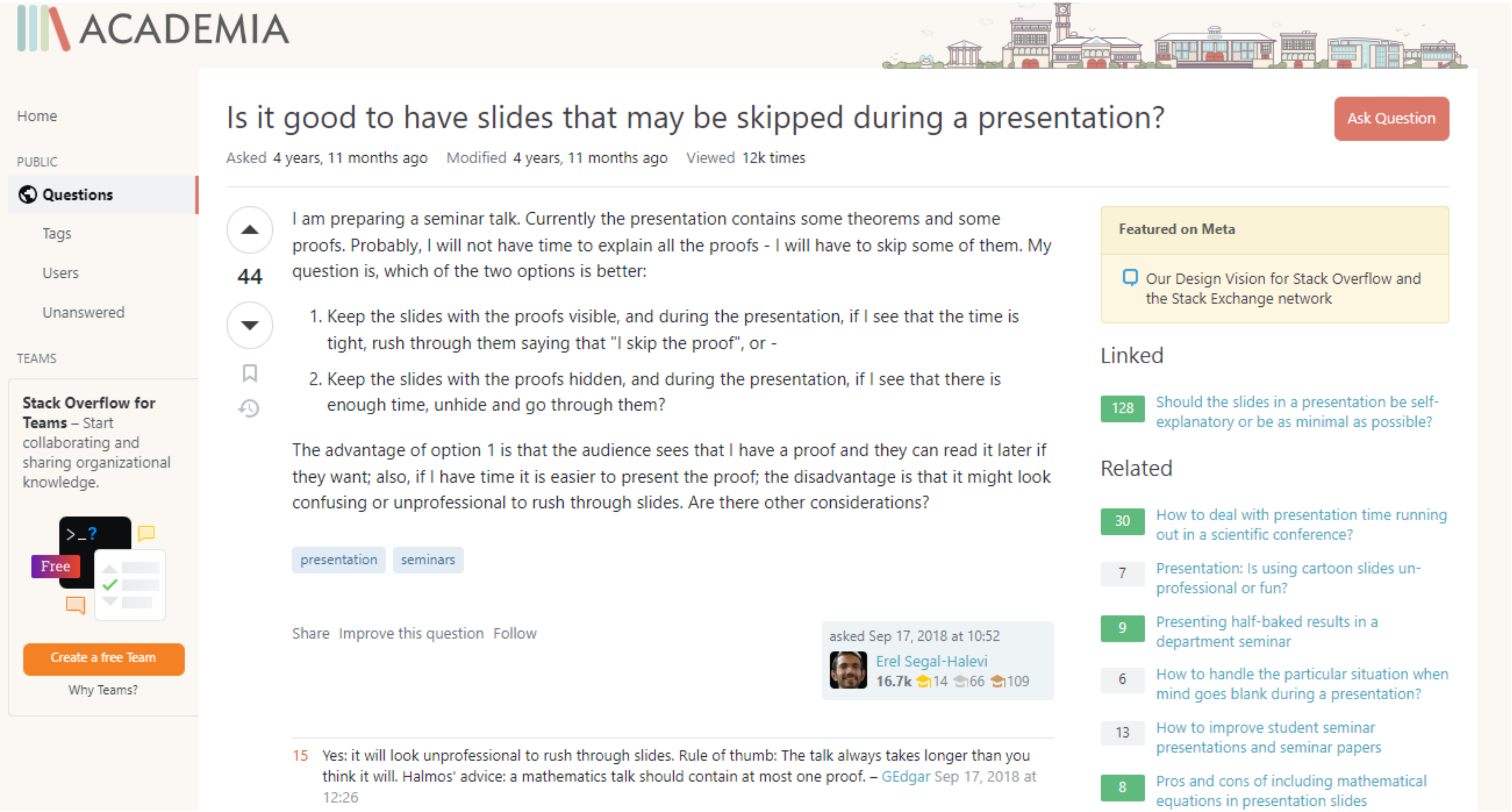
HAROLD GARFINKEL
EDITED WITH AN INTRODUCTION BY MICHAEL LYNCH

HAROLD GARFINKEL: STUDIES OF WORK IN THE SCIENCES

Directions in Ethnomethodology and Conversation Analysis

R

This slide is skipped because lack of time



Stack Overflow for Teams – Start collaborating and sharing organizational knowledge.

Is it good to have slides that may be skipped during a presentation?

Asked 4 years, 11 months ago Modified 4 years, 11 months ago Viewed 12k times

44

I am preparing a seminar talk. Currently the presentation contains some theorems and some proofs. Probably, I will not have time to explain all the proofs - I will have to skip some of them. My question is, which of the two options is better:

1. Keep the slides with the proofs visible, and during the presentation, if I see that the time is tight, rush through them saying that "I skip the proof", or -
2. Keep the slides with the proofs hidden, and during the presentation, if I see that there is enough time, unhide and go through them?

The advantage of option 1 is that the audience sees that I have a proof and they can read it later if they want; also, if I have time it is easier to present the proof; the disadvantage is that it might look confusing or unprofessional to rush through slides. Are there other considerations?

presentation seminars

asked Sep 17, 2018 at 10:52
Erel Segal-Halevi
16.7k 14 66 109

15 Yes: it will look unprofessional to rush through slides. Rule of thumb: The talk always takes longer than you think it will. Halmos' advice: a mathematics talk should contain at most one proof. – GEdgar Sep 17, 2018 at 12:26

Featured on Meta

Our Design Vision for Stack Overflow and the Stack Exchange network

Linked

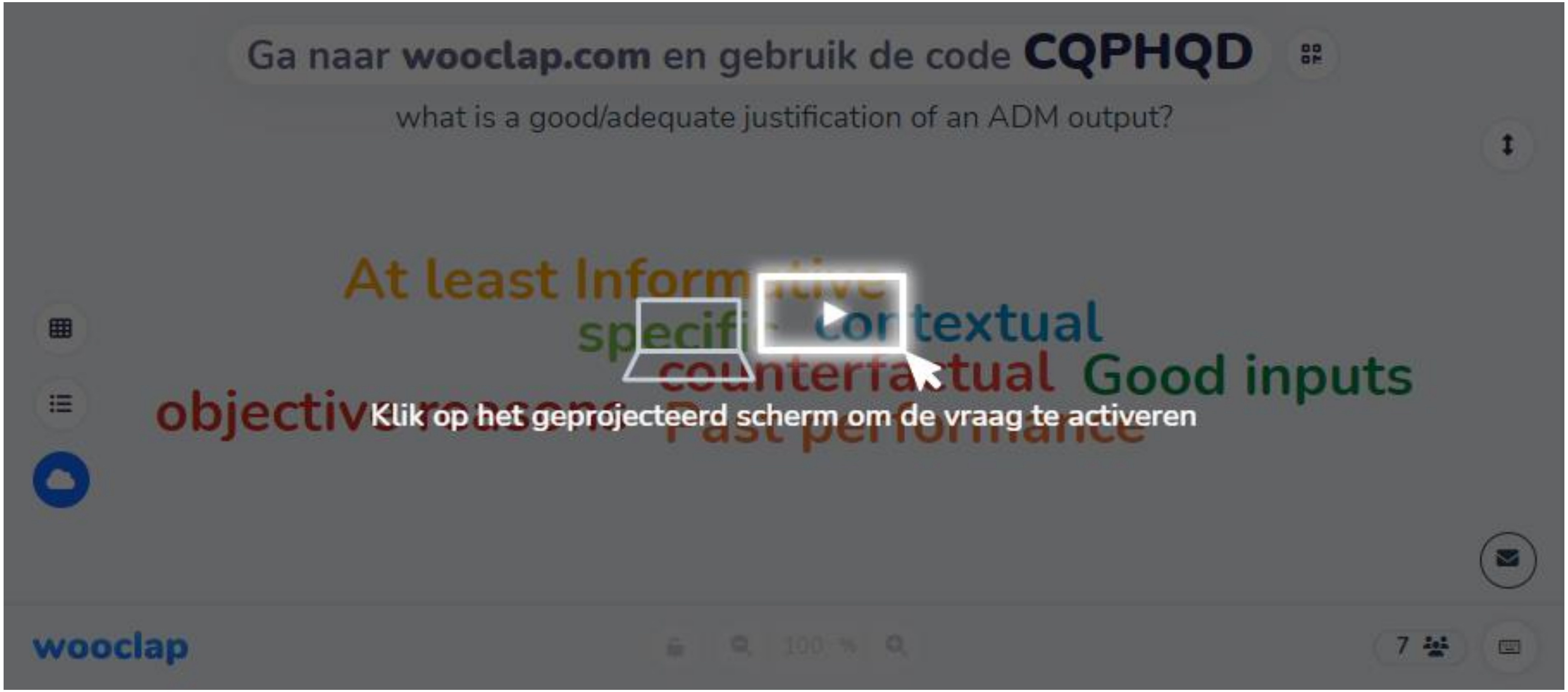
- 128 Should the slides in a presentation be self-explanatory or be as minimal as possible?

Related

- 30 How to deal with presentation time running out in a scientific conference?
- 7 Presentation: Is using cartoon slides un-professional or fun?
- 9 Presenting half-baked results in a department seminar
- 6 How to handle the particular situation when mind goes blank during a presentation?
- 13 How to improve student seminar presentations and seminar papers
- 8 Pros and cons of including mathematical equations in presentation slides

Accountability and public reason

- ‘obligation to provide justification’
- What type justifications are *good* justifications?
- ADM
  - Epistemic and normative assumptions
  - Reflected in justification
- No guarantee that explanation is acceptable to all (reasonable pluralism)
- ‘There is therefore likely to be a gap between merely providing reasons and explanations for an ADM system’s output, and providing adequate justification for them that will be acceptable to affected decision-subjects’



Ga naar **wooclap.com** en gebruik de code **CQPHQD**

what is a good/adequate justification of an ADM output?

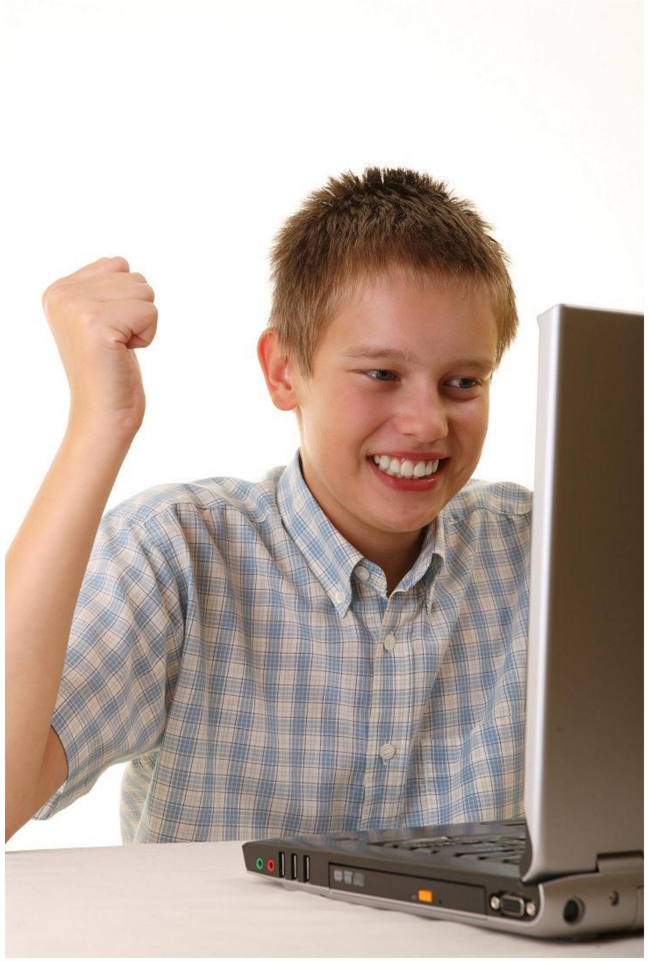
Klik op het geprojecteerd scherm om de vraag te activeren

Public reason

- ‘Public reason proposes that universal rules must be justifiable on grounds that are suitably public and shared by all reasonable people in the society, and without appeal to beliefs that are controversial.’

Next week:

- Bias, discrimination and data justice (Donovan)
- Last debating session



Questions on Module 10?

G.Meyers@tilburguniversity.edu