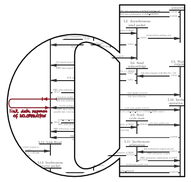
Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics C5.1
Societal and Environmental Issues

Work and Society



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

Swing Riots (1830 – 1833)

- an uprising by agricultural workers
- destruction of threshing machines
- three targets:
  - the tithe system (i.e. you had to give away 10% of what you earned as a tax)
  - the Poor Law guardians (those who decided who got poor aid)
  - and the rich farmers who lowered wages after introducing agricultural machinery

https://en.wikipedia.org/wiki/File:Batteuse_1881.jpg

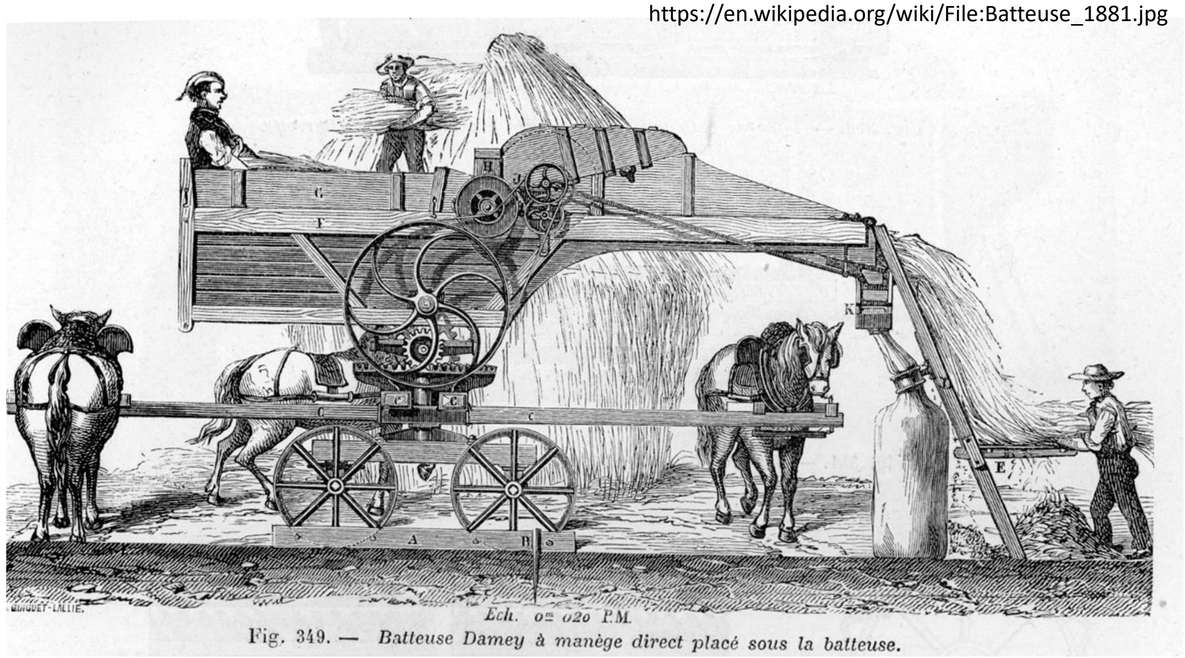
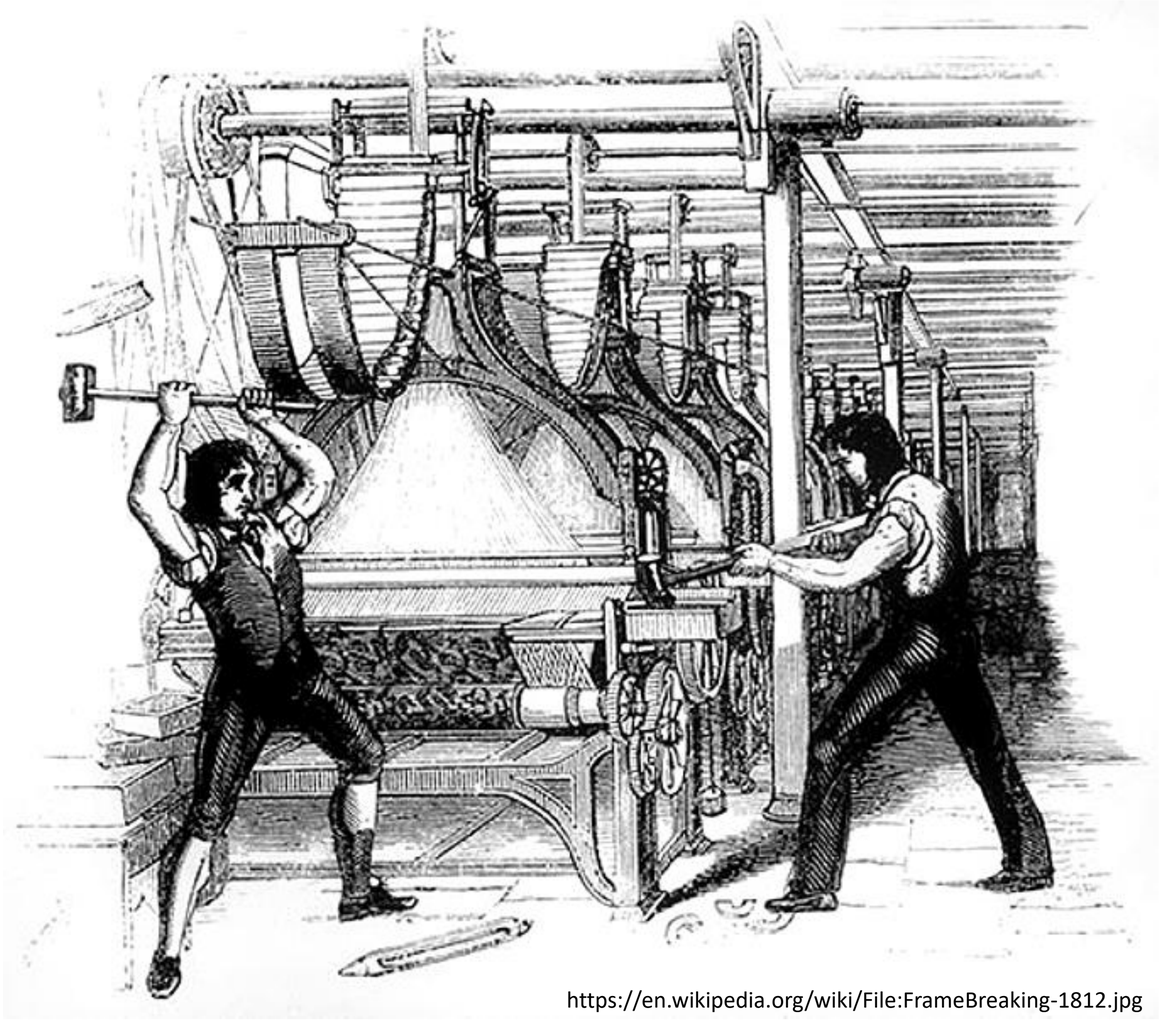


Fig. 349. — *Batteuse Damey à manège direct placé sous la batteuse.*

Luddites (1811 – 1816)

- group of English textile workers and weavers who destroyed weaving machinery as a form of protest feared replacement by machines, since those could be operated by less qualified workers who got a lower wage
- did not hate technology, but only wanted to have a better bargaining position
- ended when some supporters of the movement where sentenced to death

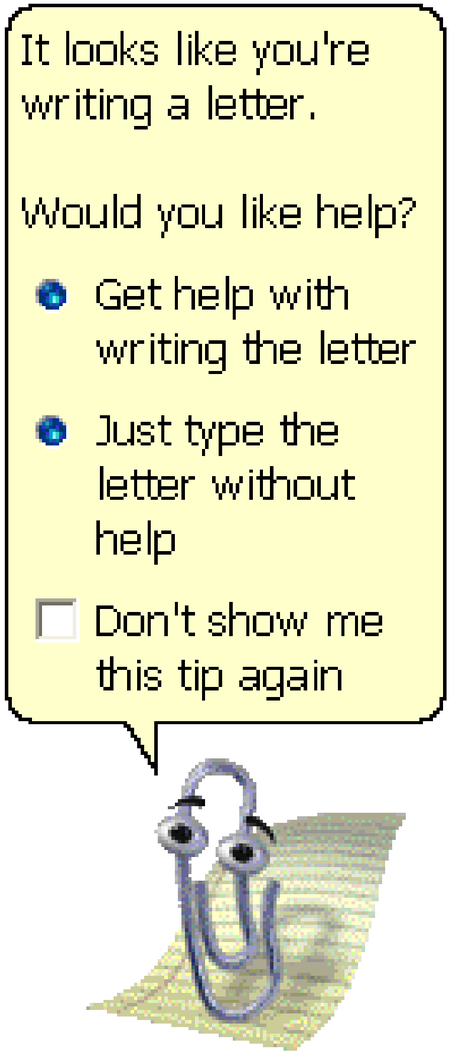


https://en.wikipedia.org/wiki/File:FrameBreaking-1812.jpg

Automation is not new

https://www.flickr.com/photos/seattlemunicipalarchives/2710933334

http://pngimg.com/download/6670



https://en.wikipedia.org/wiki/Office_Assistant

The Uber Example

UBER

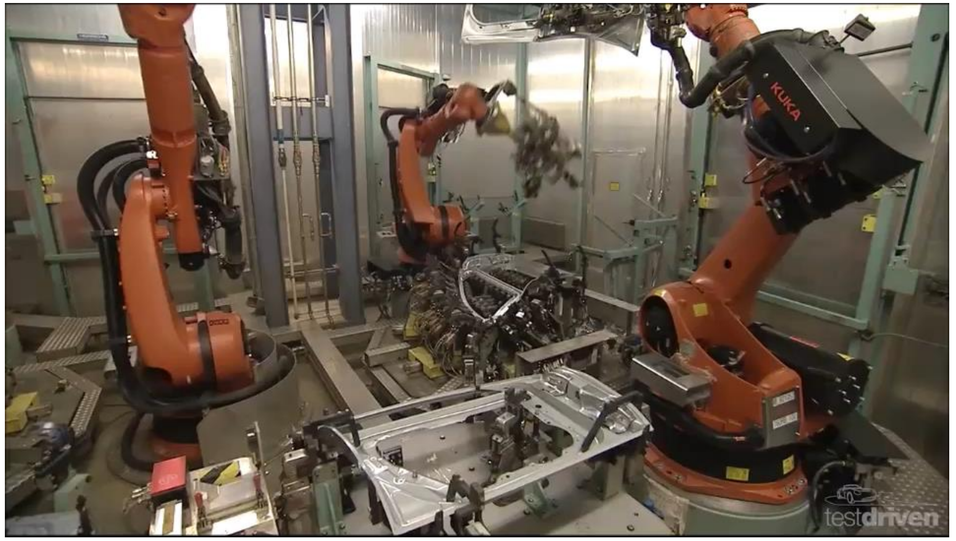
human drivers VS **autonomous cars**

- estimated 2 million drivers worldwide

- company does research on self-driving cars
- goal is to eventually substitute all cars by self-driving cars

→ Is this kind of automation desirable?

AUTOMATION AND SOCIETY



https://youtu.be/VreG1iC65Lc

https://youtu.be/8gy5tYVR-28

https://youtu.be/MkBhXx7OKto

https://youtu.be/a-9A0ar70pI

Two common reactions:

1. people *fear* automation (and modern technology in general)
2. people *do not understand* automation (and modern technology in general)

AUTOMATION AND SOCIETY

Neo-Luddites (today)

unaffiliated groups of people who are strongly against most modern technology

Example: group that attacked a fab lab in Grenoble, France

> “City administrators oblige greedy start-up companies and tendentially geek masses by opening fablabs in trendy neighborhoods. Though seemingly heterogeneous, these facilities all aim at accelerating acceptance and social use of technologies in our doomed era. (...) Because they are a nuisance we have just destroyed one of them.”
> (https://www.ecsite.eu/node/24376)

https://www.ledauphine.com/faits-divers/2017/11/25/incendie-de-la-casemate-la-mouvance-libertaire-revendique-hsya

https://www.ecsite.eu/members/members-share/news/fire-la-casemate

DIGITAL GAP

Jim Sevier, “Bridging the Digital Divide” at TEDxGreenville
https://www.youtube.com/watch?v=fzokRz1pgb0

Computer scientists understand digital technologies very well, but not everybody does so

https://c1.staticflickr.com/1/41/88353546_9d99e4b8ed_b.jpg

Digital Gap (or Digital Divide)

≈ difference in the access, usage and understanding of digital technologies, e.g. due to

- low income
- little education, including bad literacy skills
- place of residency
- age
- gender*
- …

*their gender per se does not inhibit anyone from using and understanding digital technologies, but rather cultural contexts are often such that women are disadvantaged when it comes to digital technologies

DIGITAL GAP

Jim Sevier, “Bridging the Digital Divide” at TEDxGreenville
https://www.youtube.com/watch?v=fzokRz1pgb0

Computer scientists understand digital technologies very well, but not everybody does so

https://c1.staticflickr.com/1/41/88353546_9d99e4b8ed_b.jpg

Digital Gap (or Digital Divide)

The ACM Code of Ethics and Professional Conduct tells professionals:

2.7 Improve public understanding of computing and its consequences.
Computing professionals have a responsibility to share technical knowledge with the public by encouraging understanding of computing, including the impacts of computer systems and their limitations. This imperative implies an obligation to counter any false views related to computing.

DIGITAL GAP

Jim Sevier, “Bridging the Digital Divide” at TEDxGreenville
https://www.youtube.com/watch?v=fzokRz1pgb0

Computer scientists understand digital technologies very well, but not everybody does so

https://c1.staticflickr.com/1/41/88353546_9d99e4b8ed_b.jpg

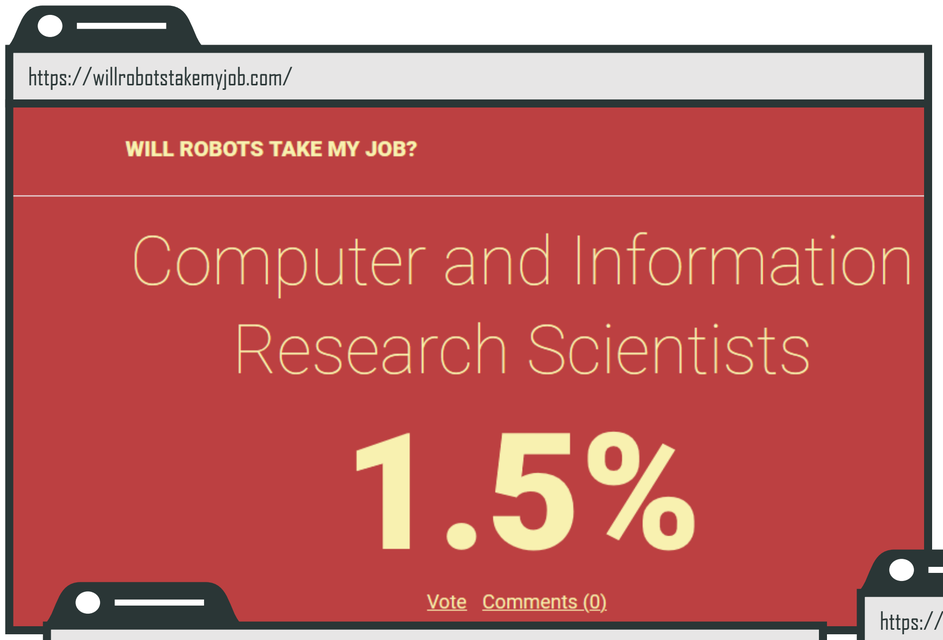
Digital Gap (or Digital Divide)

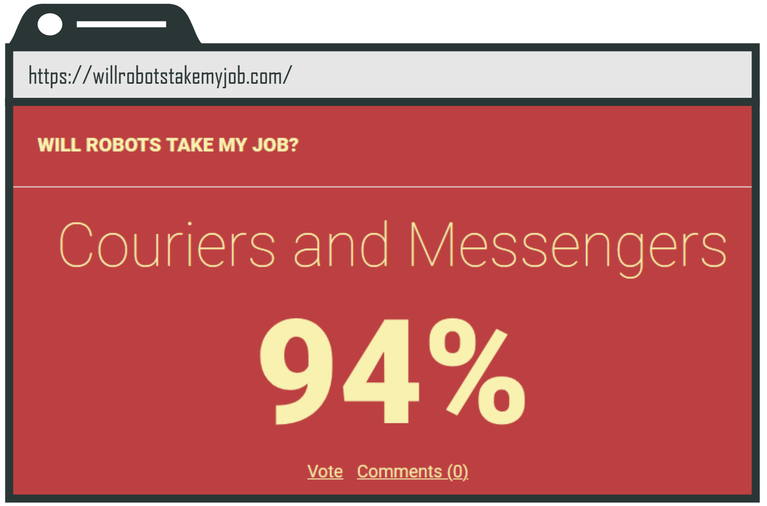
- lots of different countermeasures for the different demographic, socioeconomical, geographic groups needed
- finding countermeasures gets harder the further technology develops

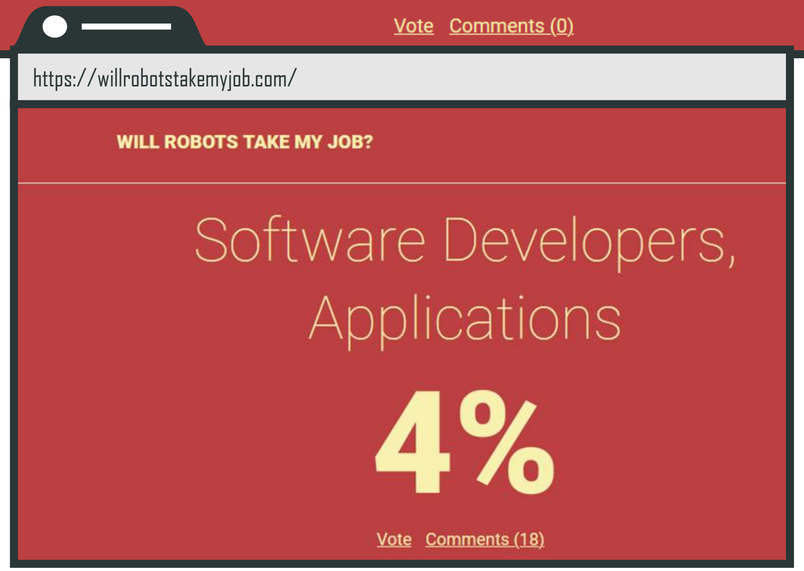
even you are taking part in this lecture and learn new, important aspects about digital technologies – how should a layperson understand everything that is important?

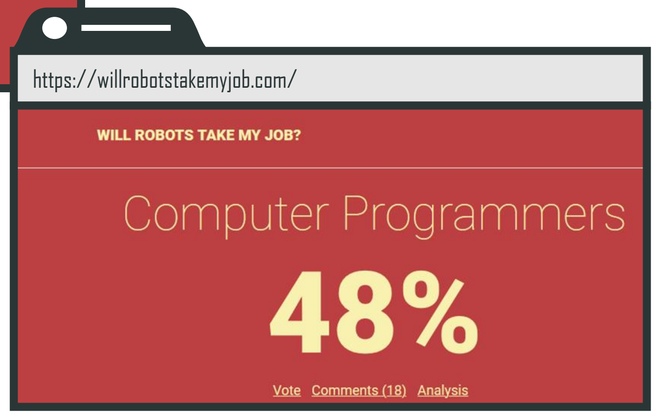
WILL A ROBOT TAKE MY JOB?

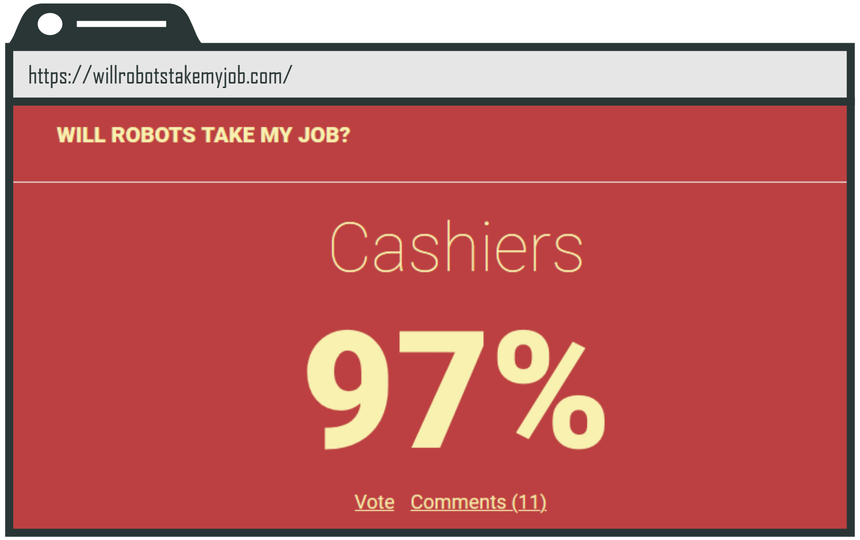
https://willrobotstakemyjob.com/











AUTOMATION AND SOCIETY

Automation-Dystopia

https://www.theguardian.com/sustainable-business/video/2016/feb/17/last-job-on-earth-automation-robots-unemployment-animation-video

Why could this be bad…

Automation-Utopia

https://www.youtube.com/watch?v=Vsy4mjSHtcE

…while this is good?

Automation-Dystopia

Why could this be bad…

- **Hedonism:**
  - while she tries to make the best of her situation, Alice does not seem to have a lot of pleasures
  - people have pain when losing their jobs or being unemployed
  - some do not have enough to eat, which is painful

Automation-Utopia

…while this is good?

- **Hedonism:**
  - people seem to have something to do, which they find pleasureable, e.g. developing themselves
  - hunger has been eliminated, so at least no pain through hunger, probably pleasure through enjoyable food

Automation-Dystopia

Why could this be bad…

- **Preference Theory:**
  - we see lots of frustrated preferences
    - preference to have work/to do something meaningful
    - preference to have enough to eat
    - preference to live an autonomous life
  - and only little satisfied ones

Automation-Utopia

…while this is good?

- **Preference Theory:**
  - people’s preference to do something meaningful seems fulfilled
  - their preference to develop themselves could be fulfilled or frustrated
  - no material preferences, thus no frustration or fulfillment
  - preference to have enough good food is satisfied

AUTOMATION AND SOCIETY

Automation-Dystopia

Why could this be bad…

Automation-Utopia

…while this is good?

Both hedonism and preference theory favor the automation-utopian scenario over the dystopian scenario.

How could both scenarios look in reality, outside of fiction?

AUTOMATION AND SOCIETY

Automation-Dystopia

- every company acts like Uber and gradually replaces human workers with autonomous systems

  → unemployment and poverty

- only high qualified, creative and probably communicative jobs remain

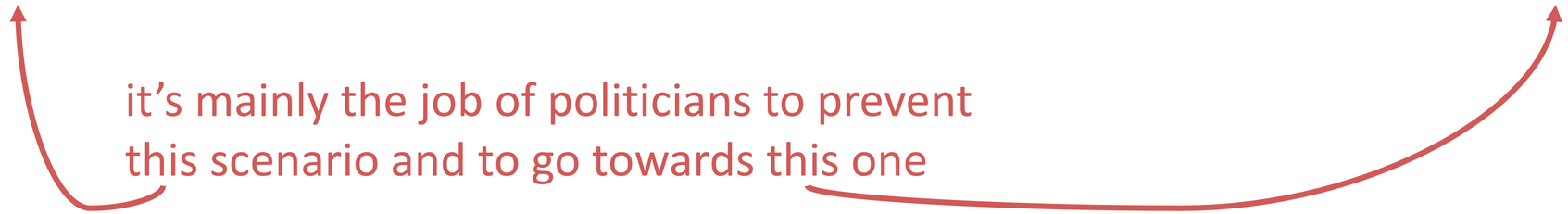
  → less and less people can keep up with the raising level of education

→ very much divided society, lots of poverty

Automation-Utopia

- Robots are doing work nobody wants to do
  - dangerous jobs
  - monotone jobs
  - dirty jobs
- a solution is implemented so that unemployment does not result in poverty anymore

→ society of lots of happy people

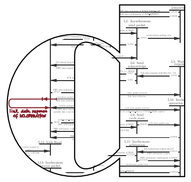


Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics C5.2
Societal and Environmental Effects

Environment and Sustainability

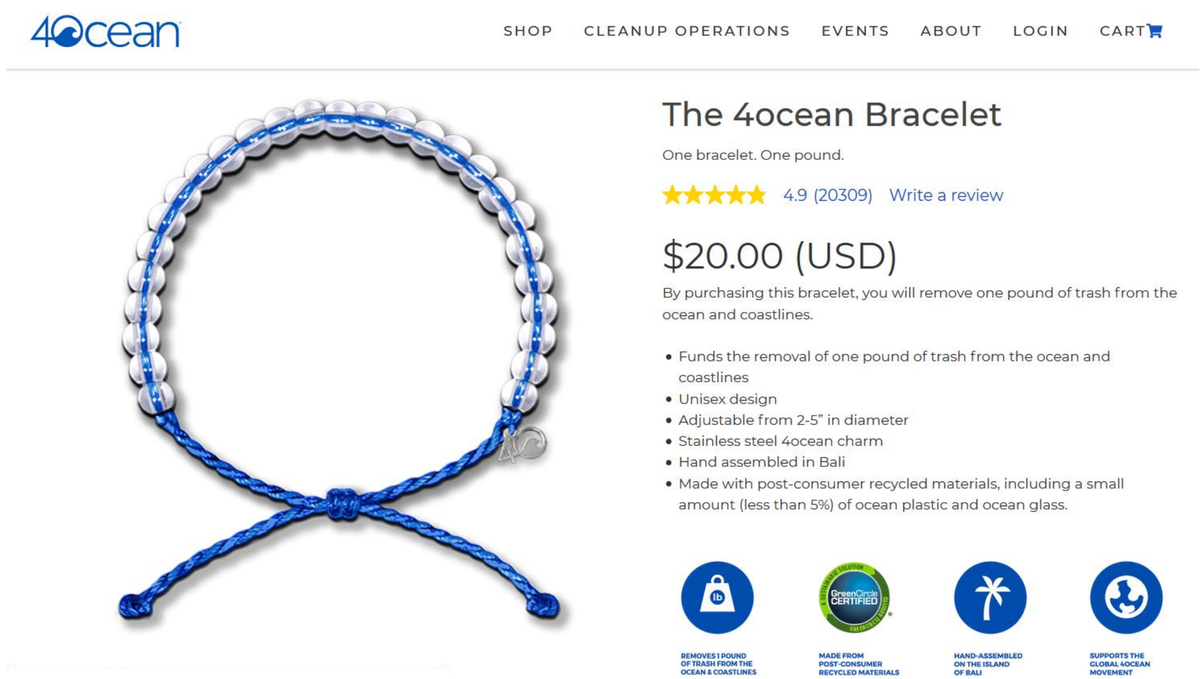


Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

ANOTHER RHETORICAL PLOY

Appeal to Alleged Sustainability



https://www.truthinadvertising.org/six-companies-accused-greenwashing/

A CHARACTERIZATION?

pre-theoretic notion will suffice

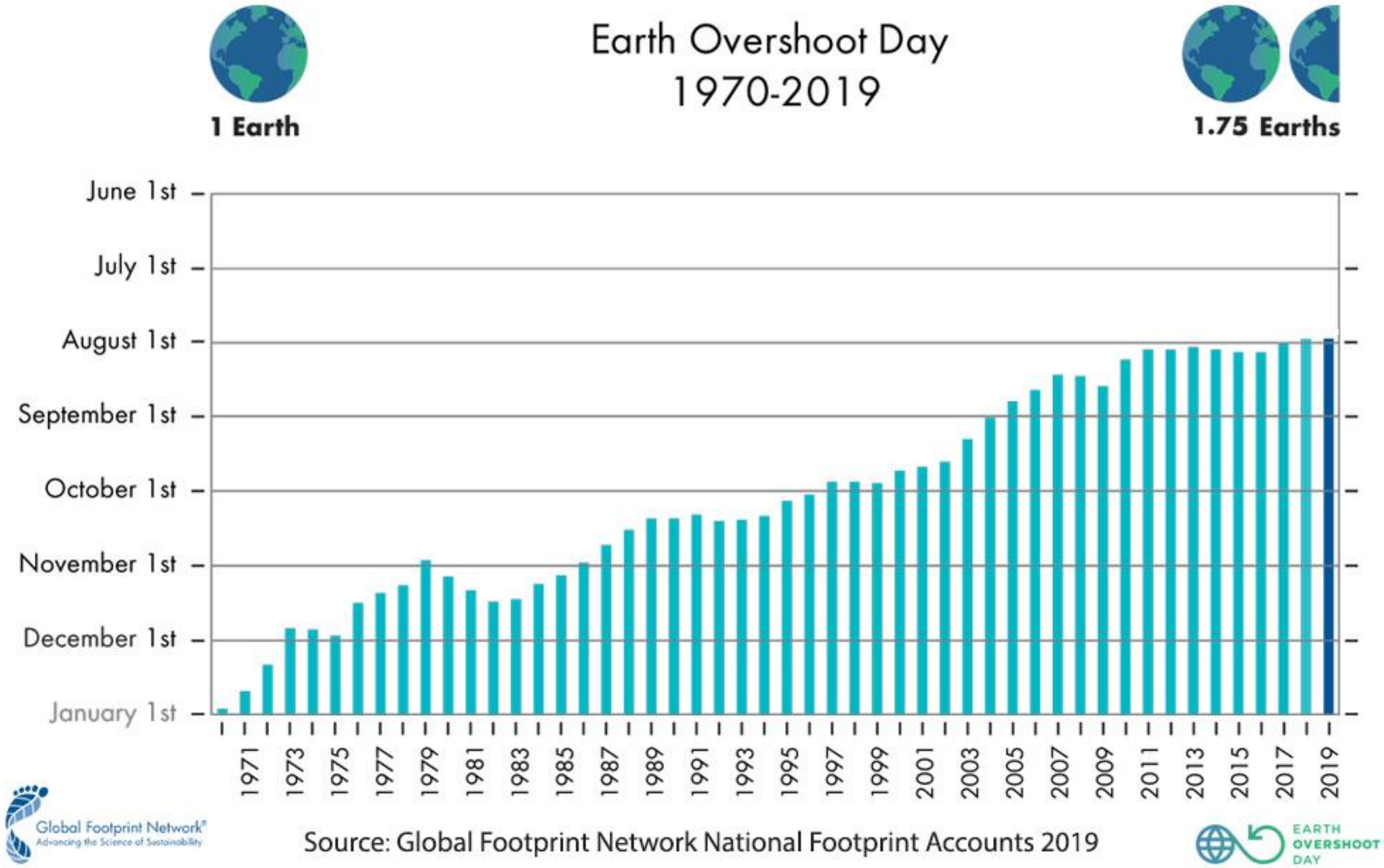
Different definitions of sustainability

- "**Development** that meets the needs of the present without compromising the ability of future generations to meet their own needs." (Bruntland Report for the World Commission on Environment and Development, 1992)
- "Sustainability is the **ability** to exist constantly. In the 21st century, it refers generally to the **capacity** for the biosphere and human civilization to coexist. It is also defined as the **process** of people maintaining change in a balanced environment, in which [various major decisions of society] are all in harmony and enhance both current and future potential to meet human needs and aspirations." (https://en.wikipedia.org/wiki/Sustainability)
- "The **ability** to be maintained at a certain rate or level." (Oxford Dictionary)
- "**Avoidance** of the depletion of natural resources in order to maintain an ecological balance." (Oxford Dictionary)
- "A **network of interactions** that achieves a consistent sum of resource components to operate and evolve indefinitely without collapse or additional influx of energy." (https://intercongreen.com/2010/04/29/what-does-sustainability-mean-anyway/)

UN GOALS

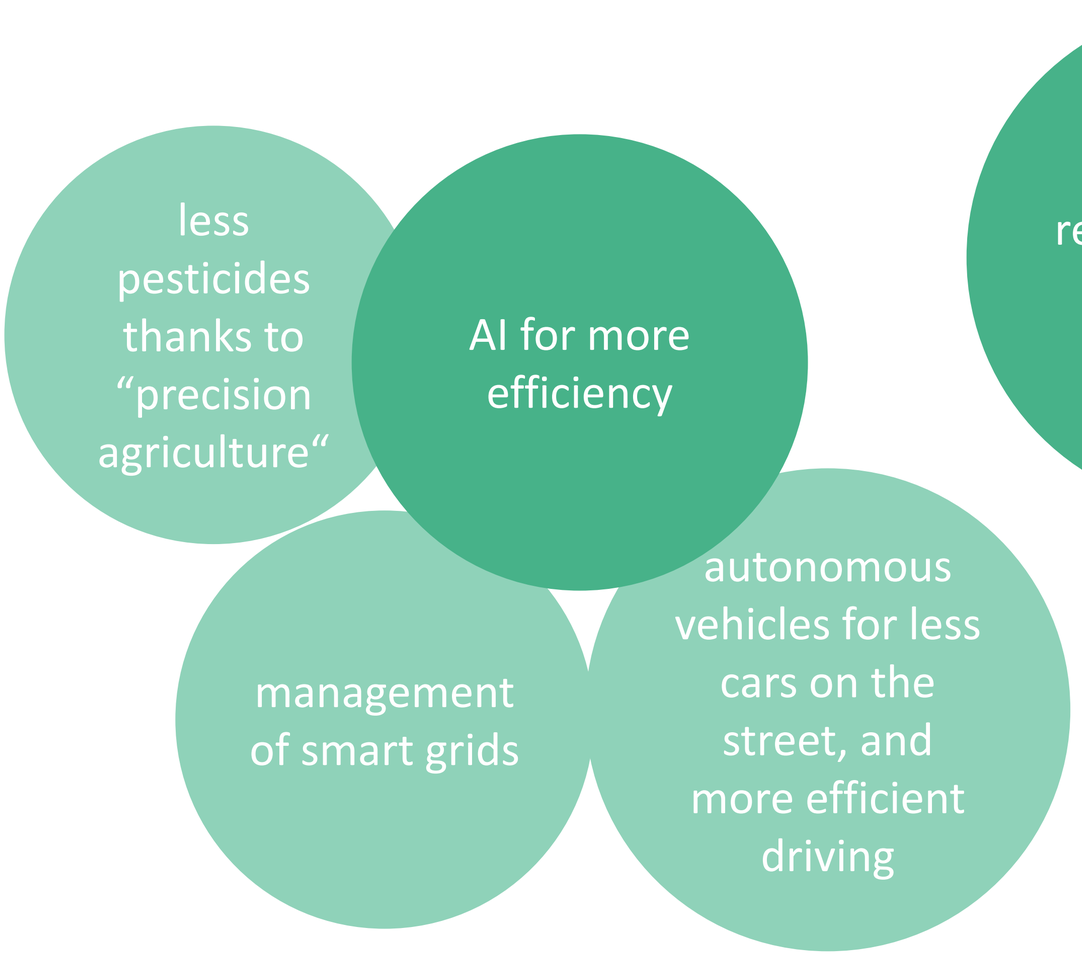
Check out https://sustainabledevelopment.un.org/

WORLD OVERSHOOT DAY



https://www.overshootday.org/newsroom/press-release-june-2019-german/

POSITIVE EFFECTS OF IT ON ENVIRONMENT



AI for reinforcement of ‘green’ behaviour

Dematerialization: digital businesses models and businesses work nearly without resources, besides hardware and energy

https://www.iof2020.eu/blog/2019/04/artificial-intelligence-for-digital-precision-agriculture

NEGATIVE EFFECTS OF IT ON ENVIRONMENT

blockchain mining

Bitcoin’s annual energy consumption is greater than that of Austria

One BC transaction takes about four times as much energy as 100.000 VISA transactions

Still you could do better by using https://www.ecosia.org/

search engines (though much less impact than you think)

hardware needs plenty of rare earth elements

smart meters can infer which TV show you are watching

privacy issues

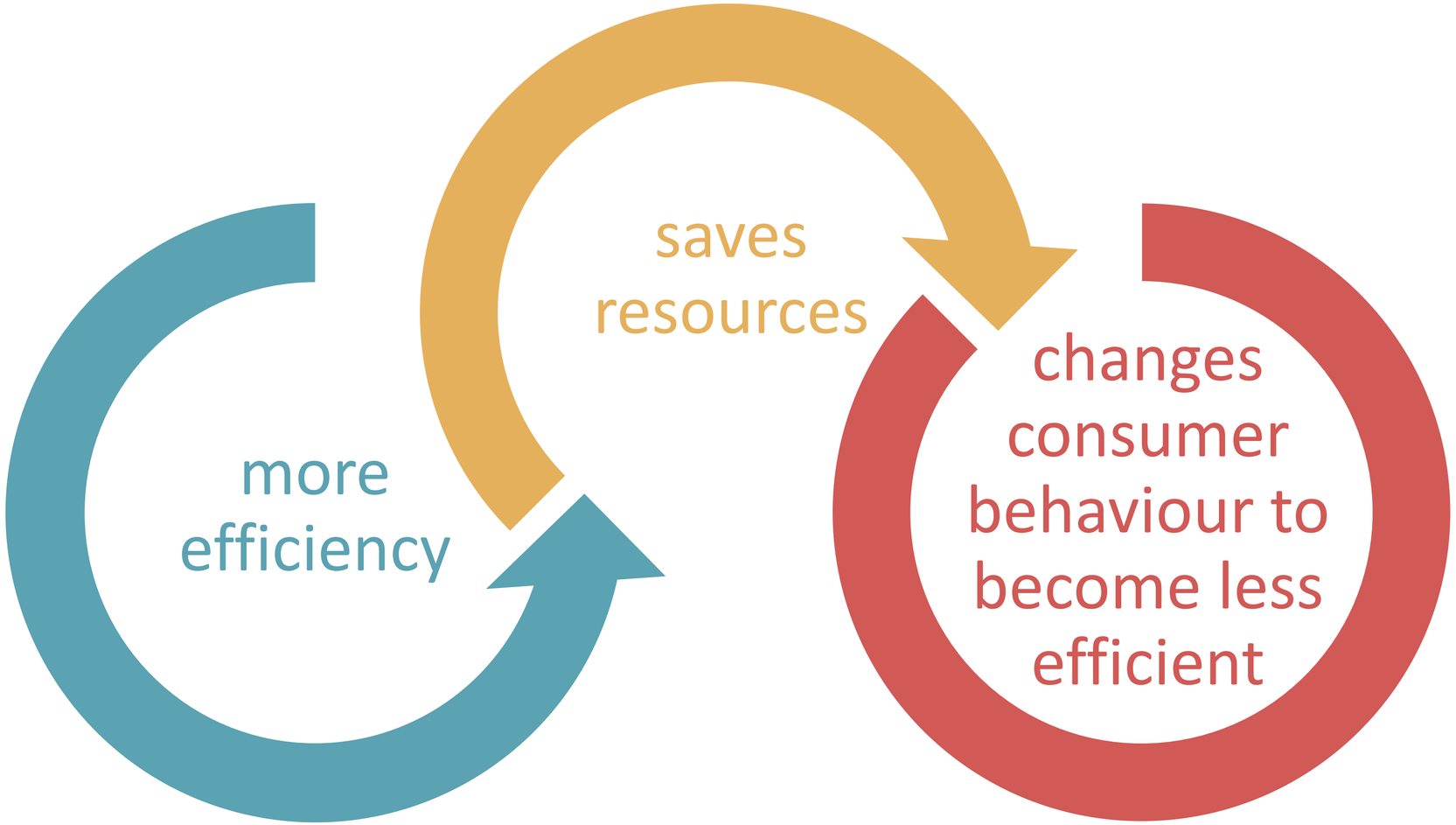
https://www.cnet.com/news/researchers-find-smart-meters-could-reveal-favorite-tv-shows/

costly streaming services

Streaming online pornography produces as much CO2 as Belgium

https://www.newscientist.com/article/2209569-streaming-online-pornography-produces-as-much-co2-as-belgium/

REBOUND EFFECTS

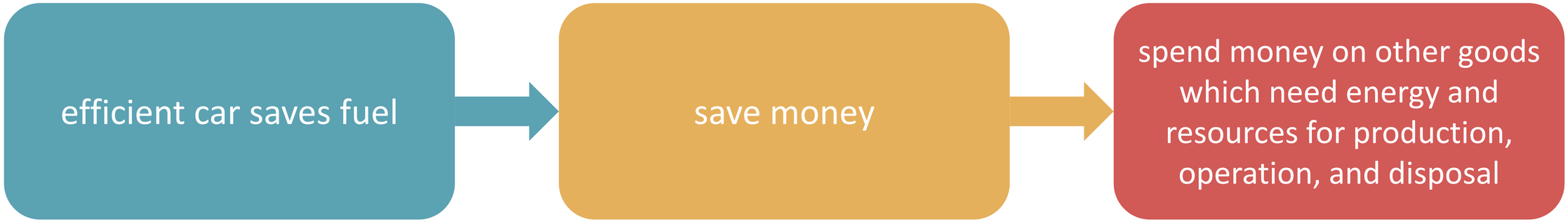


REBOUND EFFECTS

Direct rebound-effect
increased demand for the same good

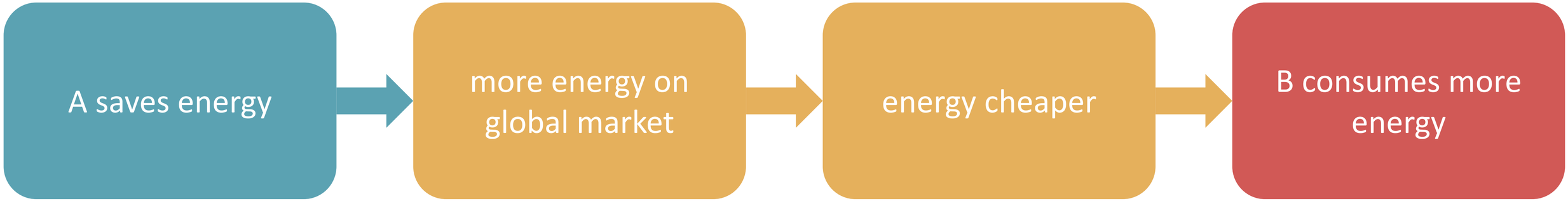
REBOUND EFFECTS

Indirect rebound-effect
increased demand for the other goods



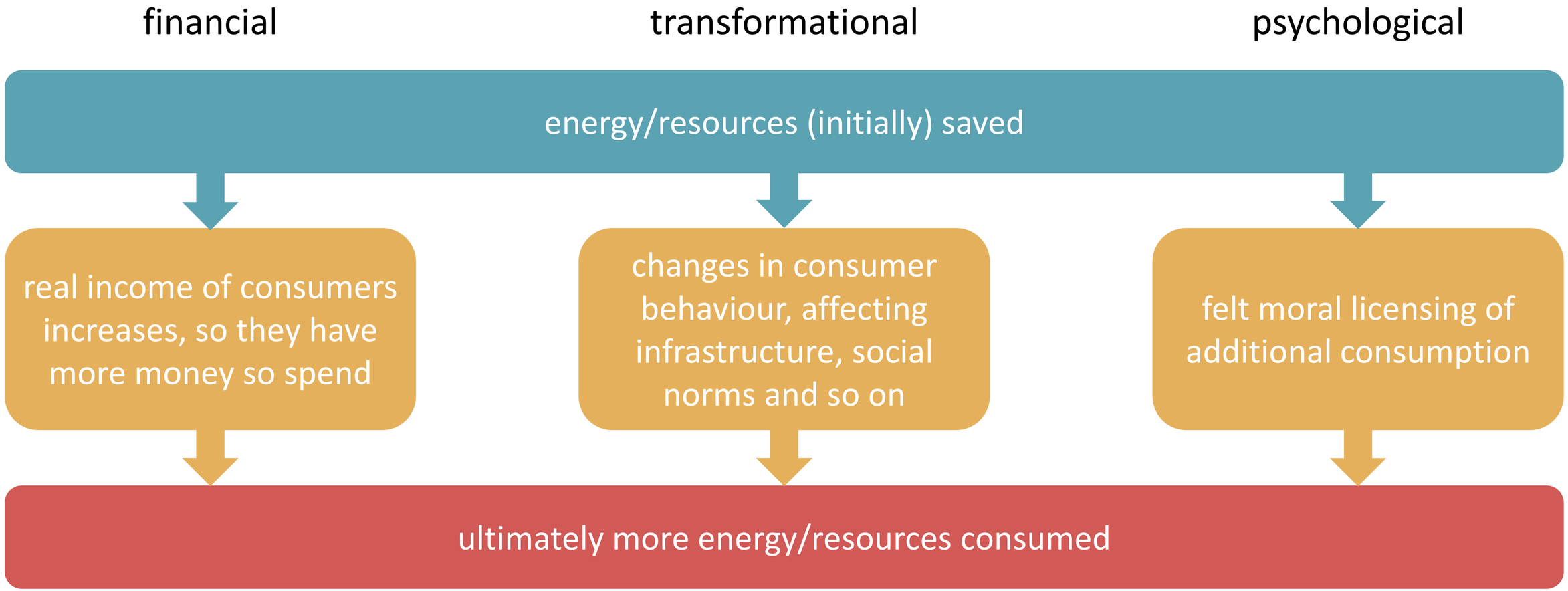
REBOUND EFFECTS

Macroeconomic rebound-effect



REBOUND EFFECTS

Reasons for rebound-effects



TWO LESSONS

Lesson 1:
It’s implausible that technological progress alone will suffice for making the world sustainable.

Lesson 2:
Lesson 1 does not imply that we can reasonably hope to achieve sustainability without technological progress – given that billions of people in the world strive for prosperity and wealth.

There are *good reasons* to aim at sustainability and live sustainably. Right?

EXCURSUS: PRECAUTIONARY PRINCIPLE

The Precautionary Principle

The precautionary principle shifts the burden of proof – it demands to disprove possible dangers:

> it says that action should be taken to prevent certain great dangers, even if there is scientific uncertainty about the likelihood of the danger. The ‘existence’ of the danger just must be ‘plausible enough’.

It sometimes will involve acting against what an expected utility calculation would recommend

Example (Coffee Maker) Suppose there is a researcher who invents the best and cheapest coffee maker ever. This, however, produces a byproduct of which he is not sure if it is dangerous. He fears that it may cause a severe threat to the environment and human health.

EXCURSUS: PRECAUTIONARY PRINCIPLE

The Precautionary Principle in Law

- in EU law
- applied to:
  - environmental protection
  - consumer policy
  - food
  - human, animal and plant health

"2. Union policy […] shall be based on the precautionary principle and on the principles that preventive action should be taken, that environmental damage should as a priority be rectified at source and that the polluter should pay. […]"

from Article 191 of the Treaty on the Functioning of the European Union

There are *good reasons* to aim at sustainability and live sustainably. Right?

MORAL REASONS FOR SUSTAINABILITY

Is nature valuable?

	Hedonism	Preference Theory	Objective List Theories
intrinsic?	pleasure ≠ nature pain ≠ nature nature is not intrinsically (dis)valuable	preference satisfaction ≠ nature preference frustration ≠ nature nature is not intrinsically (dis)valuable	nature could be an item on the list nature could be intrinsically valuable
extrinsic?	preserving nature can both generate pleasure and pain, actions that destroy the nature and also both generate pleasure and pain	preserving nature can both generate pref. satisfaction and pref. frustration, actions that destroy the nature and also both generate pref. frustration and pref. satisfaction	depends on the list, but most likely preserving nature can both contribute to generating values from the list and work against values on the list; same holds for actions that destroy nature

→ **we need to estimate the effects that are to be expected and make a trade-off**

Is nature valuable?

→ we need to estimate the effects that are to be expected and make a trade-off

It is very clear that nature and the environment in general have huge extrinsic value for us, as it is literally our entire basis of existence.

It is morally obligatory to protect the environment.

It also has to be rational to preserve it, right?

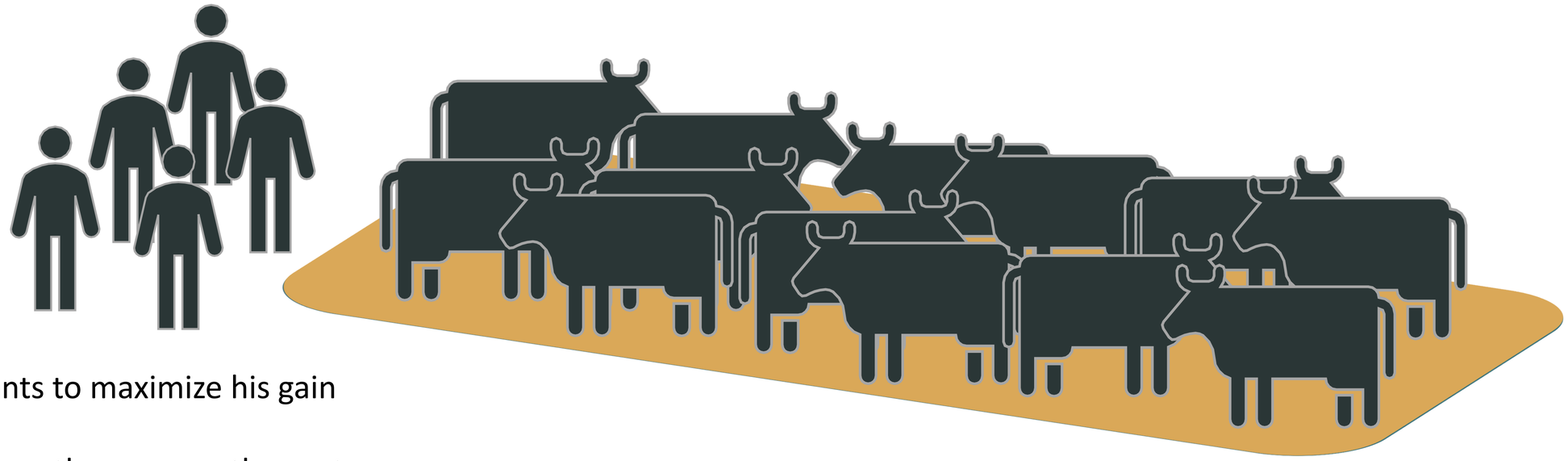
There are *good reasons* to aim at sustainability and live sustainably. Right?

RATIONAL REASONS FOR SUSTAINABILITY?

The Tragedy of the Commons by Garrett
in Science Vol. 162, Issue 3859, pp. 1243-1248, 1968:
https://science.sciencemag.org/content/162/3859/1243

Tragedy of the Commons

a situation in which the use of a resource by everyone spoils that resource



every herdsman wants to maximize his gain

utility from putting another cow on the pasture:

- $+u_1$ for him
- $-u_2$ for everybody (shared among all herdsmen)

as long as u_1 is greater than the herdsman's share of u_2, he gains by putting another cow on the pasture
→ everybody keeps putting their cows in the pasture until it becomes overgrazed and no cows can feed there anymore

RATIONAL REASONS FOR SUSTAINABILITY?

Tragedy of the Commons
a situation in which the use of a resource by everyone spoils that resource

	Everybody else puts their heard on pasture	**Everybody else does not put their heard on pasture**
I put my heard on pasture	3rd best for me	best for me
I do not put my heard on pasture	worst for me	2nd best for me

Putting my heard on the common is the dominant strategy in this game, no matter how many others put their heard on the common. But if everybody follows the dominant strategy, everybody will obtain a suboptimal result.

RATIONAL REASONS FOR SUSTAINABILITY?

Tragedy of the Commons
a situation in which the use of a resource by everyone spoils that resource

	Everybody else drives to work by car	**Everybody else does not drive to work by car**
I drive to work by car	3rd best for me	best for me
I do not drive to work by car	worst for me	2nd best for me

Driving to work by car is the dominant strategy in this game, no matter how many others go to work by car.
But if everybody follows the dominant strategy, everybody will obtain a suboptimal result.

RATIONAL REASONS FOR SUSTAINABILITY?

Tragedy of the Commons
a situation in which the use of a resource by everyone spoils that resource

	No other country lowers their emission	**All other countries lower their emissions**
Germany does not lower emissions	3rd best for Germany	best for Germany
Germany lowers emission	worst for Germany	2nd best for Germany

Not lowering your emission is the dominant strategy in this game, no matter how many others lower their emission. But if everybody follows the dominant strategy, everybody will obtain a suboptimal result.

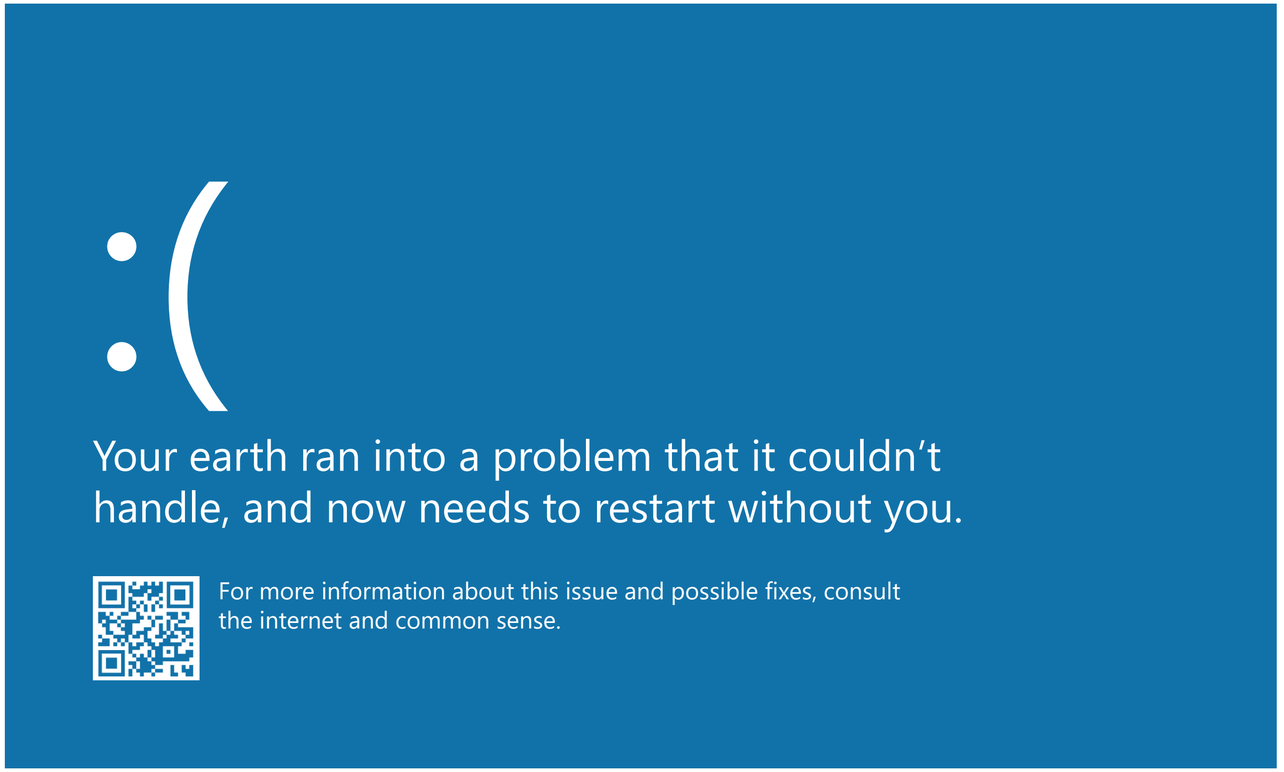
RATIONAL REASONS FOR SUSTAINABILITY? WELL…

So, it is rational to preserve nature and protect the environment?

For many problems in question, we prima facie would need a solution to the Tragedy of the Commons (which we do not have).

But that does not mean that we should not care:

- It is still **moral** to protect the environment.
- It still can be **rational** to behave in a way that gives others an incentive to protect the environment.
- You are still **responsible** for the damage you cause.

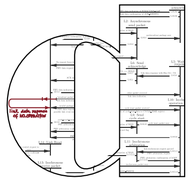


Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics C5.3
Societal and Environmental Effects

Responsibility



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

RESPONSIBILITY IN CS

Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/.

Prima facie, there are at least three things that are necessary for A being morally responsible for X:

causal connection

1 there is a causal connection between A's doing and X

A person should usually not be held responsible for something if she had nothing at all to do with it.

epistemic access

2 A had enough sufficiently well justified beliefs about the possible consequences of their doing that lead to X

We usually should not hold someone responsible if they could not have known that their actions would lead to a harmful event.

control

3 A had/was able to have some control over their doing that lead to X

It does not seem to make sense to hold someone responsible for an outcome if their doing was completely outside of their control.

All of them are potentially problematic in computer science…

EXCURSUS

Excursus: Causality (David Lewis, simplified)

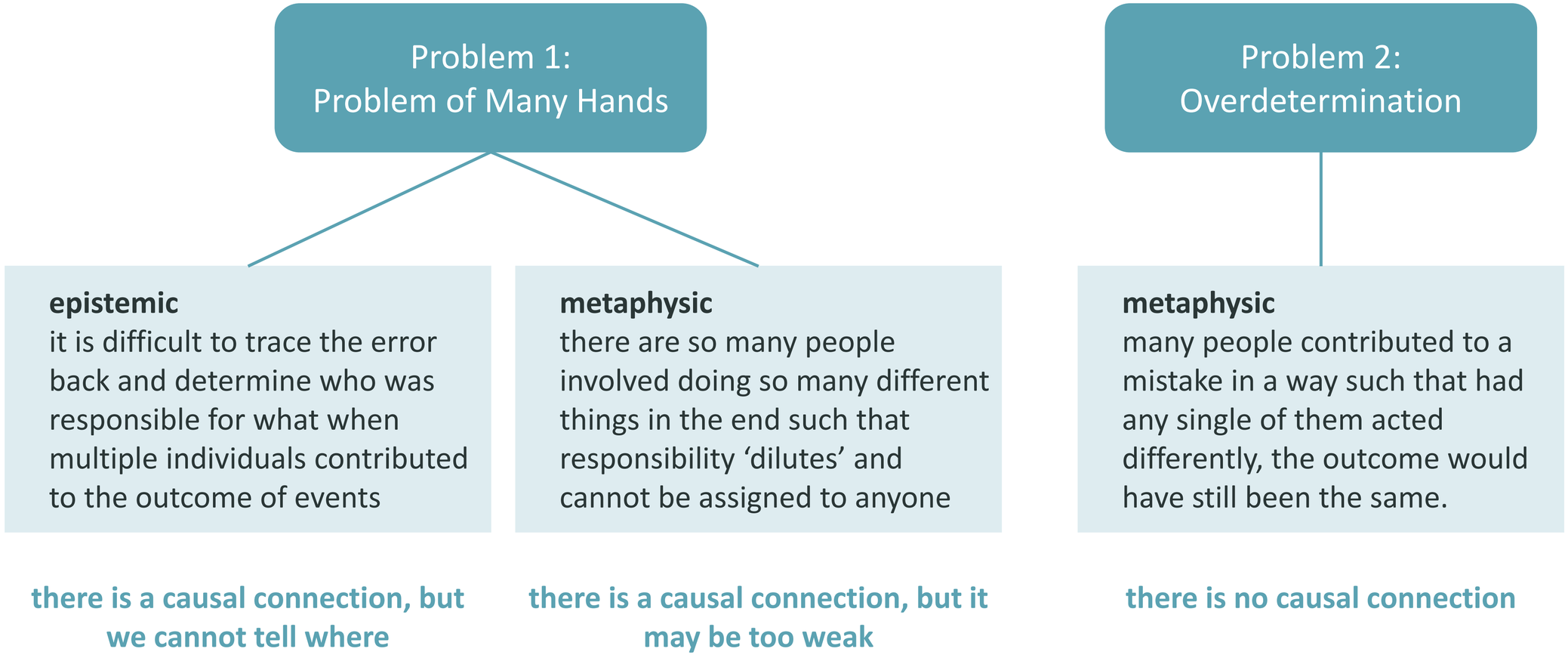
X is cause for Y iff the following counterfactual conditional holds:

> “If X did not occur, then Y would not have occured.”

RESPONSIBILITY IN CS

Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/.

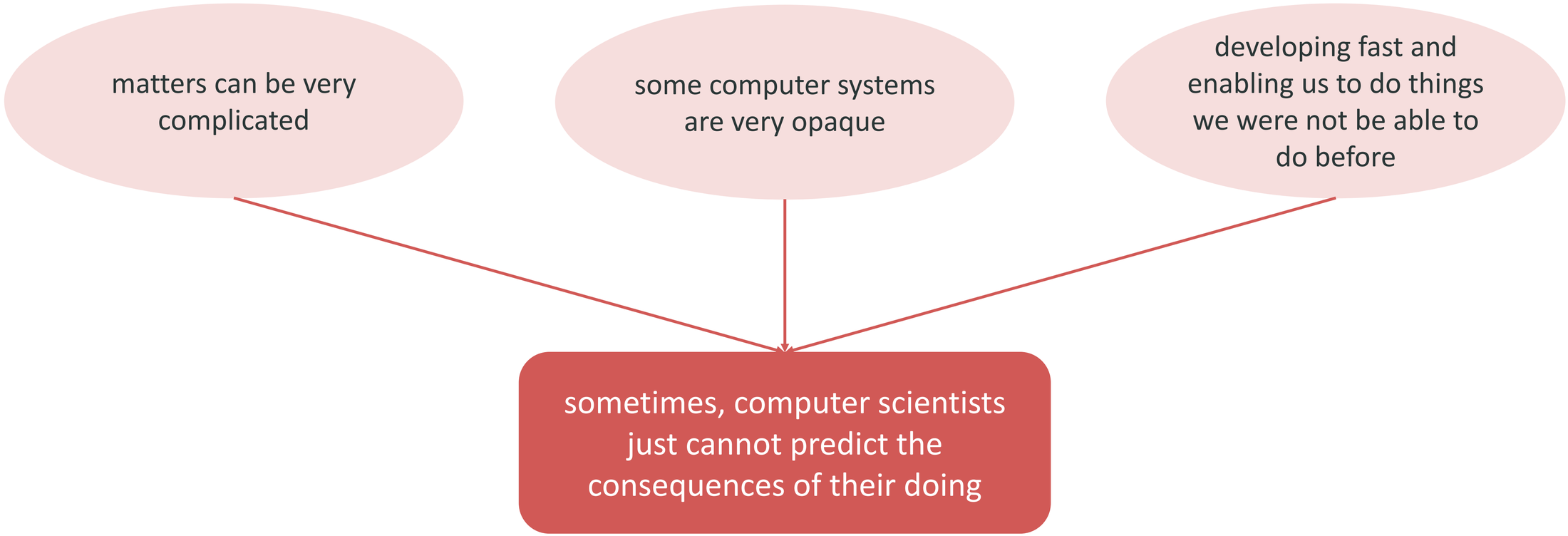
1. a causal connection between agent and outcome



RESPONSIBILITY IN CS

Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/.

2. agent has to have knowledge of and be able to consider the possible consequences of her actions



RESPONSIBILITY IN CS

Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/.

3. agent has to (be able to) have some control over their doing

some issues raise questions whether we really have sufficient control in all CS related matters

Example 1

we can choose to automate a decision making procedure or to take automation out of a decision making procedure, but once we implemented a certain system, we cannot control what the automated decision is

Example 2

sometimes technologies are specifically designed to nudge us in order to make us show a certain behaviour

RESPONSIBILITY IN CS

Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/.

For some contexts, they can also be taken as sufficient conditions. For this lecture, we will simplify and say that all three conditions together are sufficient for responsibility.

Prima facie, there are at least three things that are necessary for A being morally responsible for X:

causal connection

1 there is a causal connection between A's doing and X

A person should usually not be held responsible for something if she had nothing at all to do with it.

epistemic access

2 A had enough sufficiently well justified beliefs about the possible consequences of their doing that lead to X

We tend to excuse someone from blame if they could not have known that their actions would lead to a harmful event.

control

3 A had/was able to have some control over their doing that lead to X

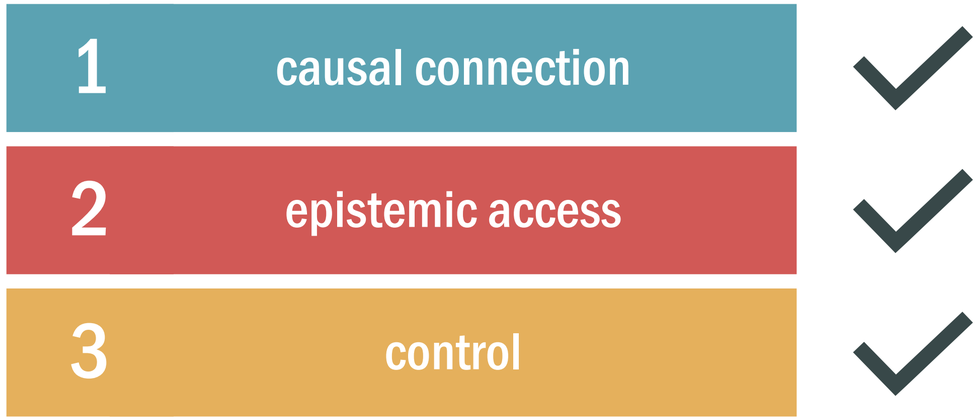
It does not seem to make sense to hold someone responsible for an outcome if their doing was completely outside of their control.

All of them are potentially problematic in computer science, **but still there are clear cases in which computer scientists and programmers are responsible for their work.**

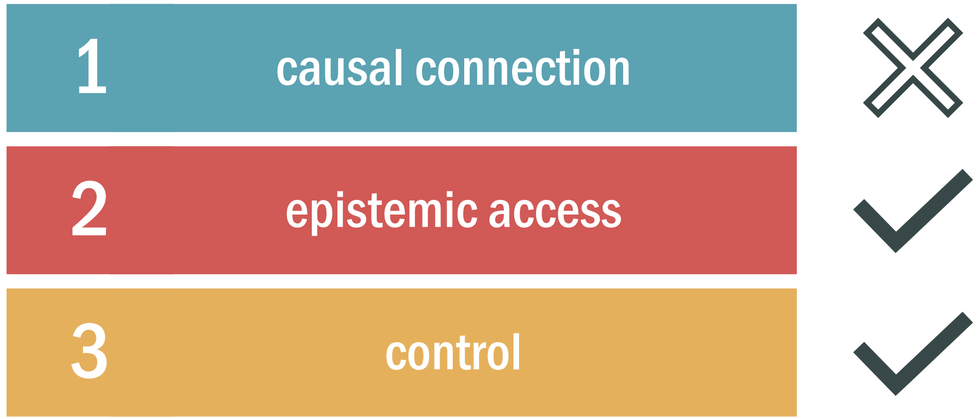
RESPONSIBILITY FOR THE CLIMATE CHANGE

Who is responsible for climate change under our definition?

If anyone, it's *politicians and other regulators…*



…but *you as an individual* are not responsible for climate change according to our definition.



RESPONSIBILITY FOR THE CLIMATE CHANGE

Who is responsible for climate change under our definition?

But you are responsible for individual damage you cause!