THEME 1: ACCEPTABLE AI

- Bruno Jullien, TSE
- Jean-Francois Bonnefon, TSE
- Cesar A. Hidalgo, University of Toulouse, Manchester, & Harvard

Bruno Jullien
TSE

AI & Competition

Doh-Shin Jeon (Co Chair), Daniel Ershov (Co Chair), Alexandre de Cornière

UTC-TSE

Can Algorithms Learn to Collude?

High-frequency retail price data from Germany where AI is supposed to have been widely adopted in 2017

-Margins of adopters +10%
-Margins in duopolies +30% after both duopolists adopt
-No effect after only one duopolist adopts
-Margins only increase after adoption

Stephanie Assada, Robert Clarkb, Daniel Ershov, Lei Xud, "Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market"

How to understand and regulate data markets?

What are the incentives and risks faced by companies looking to sell personal data?

Alexandre de Cornière, and Greg Taylor, ìData and Competition: a General Framework with Applications to Mergers, Market Structure, and Privacy Policyî, TSE WP, n. 20-1076

Bruno Jullien, Yassine Lefouili, and Michael Riordan, ìPrivacy Protection, Security, and Consumer Retentionî, TSE WP, n. 18-947,

What are the tradeoffs and incentives that guide AI investment for the benefit of the consumer (e.g. recommender system) vis-à-vis AI investment for the benefit of the firm (e.g. price targeting).

How does using retention as an objective function biases the content of recommendation systems?

Jean-Francois Bonnefon
TSE

Moral AI

Moral AI

The Moral AI chair explores the the way **humans and machines treat each other** when they make **decisions with a moral component.**

Machines can **make decisions with that impact** human well-being. How do people want machines to behave in these cases?

Machines can **judge the morality of humans** and communicate this information to others. What are the social consequences?

Machines and humans can **work together toward a common goal.** Is this cooperation easy? Can it turn into unethical collusion?

European Commission

ETHICS OF Connected and Automated Vehicles

Independent Expert Report

Research and Innovation

• CHAPTER 1. ROAD SAFETY, RISK, DILEMMAS

Improvements in safety achieved by CAVs should be publicly demonstrable and monitored through solid and shared scientific methods and data; these improvements should be achieved in compliance with basic ethical and legal principles, such as a fair distribution of risk and the protection of basic rights, including those of vulnerable users; these same considerations should apply to dilemma scenarios.

• CHAPTER 2. DATA AND ALGORITHM ETHICS: PRIVACY, FAIRNESS, EXPLAINABILITY

The acquisition and processing of static and dynamic data by CAVs should safeguard basic privacy rights, should not create discrimination between users, and should happen via processes that are accessible and understandable to the subjects involved.

• CHAPTER 3. RESPONSIBILITY

Considering who should be liable for paying compensation following a collision is not sufficient; it is also important to make different stakeholders willing, able and motivated to take responsibility for preventing undesirable outcomes and promoting societally beneficial outcomes of CAVs, that is creating a culture of responsibility for CAVs.

20 RECOMMENDATIONS

1. Ensure that CAVs reduce physical harm to persons.

To prove that CAVs achieve the anticipated road safety improvements, it will be vital to establish an objective baseline and coherent metrics of road safety that enable a fair assessment of CAVs' performance relative to non-CAVs and thereby publicly demonstrate CAVs' societal benefit. This should be accompanied by new methods for continuously monitoring CAV safety and for improving their safety performance where possible.

2. Prevent unsafe use by inherently safe design.

In line with the idea of a human-centric AI, the user perspective should be put centre-stage in the design of CAVs. It is vital that the design of interfaces and user experiences in CAVs takes account of known patterns of use by CAV users, including deliberate or inadvertent misuse, as well as tendencies toward inattention, fatigue and cognitive over/under-load.

3. Define clear standards for responsible open road testing.

In line with the principles of non-maleficence, dignity and justice, the life of road users should not be put in danger in the process of experimenting with new technologies. New facilities and stepwise testing methods should be devised to promote innovation without putting road users' safety at risk.

4. Consider revision of traffic rules to promote safety of CAVs and investigate exceptions to non-compliance with existing rules by CAVs.

Traffic rules are a means to road safety, not an end in themselves. Accordingly, the introduction of CAVs requires a careful consideration of the circumstances under which: (a) traffic rules should be changed; (b) CAVs should be allowed to not comply with a traffic rule; or (c) CAVs should hand over control so that a human can make the decision to not comply with a traffic rule.

5. Redress inequalities in vulnerability among road users.

In line with the principle of justice, in order to address current and historic inequalities of road safety, CAVs may be required to behave differently around some categories of road users, e.g. pedestrians or cyclists, so as to grant them the same level of protection as other road users. CAVs should, among other things, adapt their behaviour around vulnerable road users instead of expecting these users to adapt to the (new) dangers of the road.

6. Manage dilemmas by principles of risk distribution and shared ethical principles.

While it may be impossible to regulate the exact behaviour of CAVs in unavoidable crash situations, CAV behaviour may be considered ethical in these situations provided it emerges organically from a continuous statistical distribution of risk by the CAV in the pursuit of improved road safety and equality between categories of road users.

7. Safeguard informational privacy and informed consent.

CAV operations presuppose the collection and processing of great volumes and varied combinations of static and dynamic data relating to the vehicle, its users, and the surrounding environments. New policies, research, and industry practices are needed to safeguard the moral and legal right to informational privacy in the context of CAVs.

8. Enable user choice, seek informed consent options and develop related best practice industry standards.

There should be more nuanced and alternative approaches to consent-based user agreements for CAV services. The formulation of such alternative approaches should: (a) go beyond "take-it-or-leave-it" models of consent, to include agile and continuous consent options; (b) leverage competition and consumer protection law to enable consumer choice; and (c) develop industry standards that offer high protection without relying solely on consent.

9. Develop measures to foster prot of individuals at group level.

CAVs can collect data about multiple individuals at the sa with assistance from researchers, should develop legal individuals' rights at group levels (e.g driver, pedestria drivers' rights) and should outline strategies to reso between data subjects that have claims over the sar data, computer vision data), or disputes between data su and other parties (e.g. insurance companies).

10. Develop transparency strategies inform users and pedestrians abo data collection and associated rig

CAVs move through and/or near public and private spaces monitoring and the collection of traffic-related data and it development or other measures can occur. Consequently, n strategies are needed to inform road users and pedestria a CAV operating area that may, directly or indirectly, pose

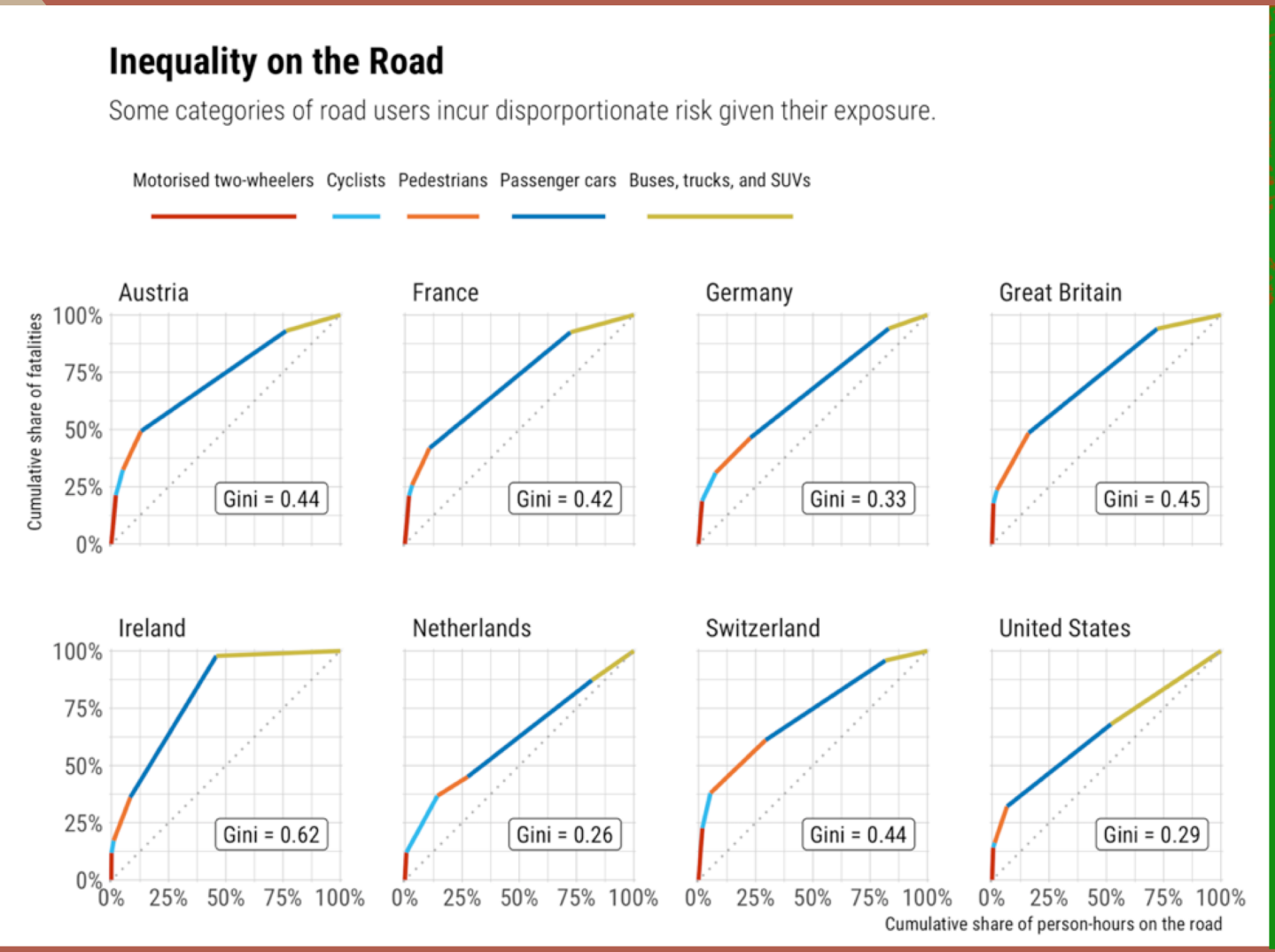
11. Prevent discriminatory differenti service provision.

CAVs should be designed and operated in ways tha against individuals or groups of users, nor create or social inequalities among users. They should also be d takes proactive measures for promoting inclusivity.

12. Audit CAV algorithms.

Investments in developing algorithmic auditing tools an adapted to and targeting the detection of unwanted conse system designs and operations of CAVS are recomme development of CAV specific means and methods of field evaluations, the results of which should be used for fo best practices and standards for CAV design, operation a counteracting any existing or emerging ethically and/or legall

Reduction of Risk Inequality



If mass surveillance and citizen scoring is part of our future, how do we introduce incentives for peer-to-peer cooperation, instead of blind obedience to the state?

Virtue Algorithms

The machine observes the online and offline behavior of the human
It aggregates these behaviors into one or several moral scores
These scores are communicated to other humans

- The Moral AI team is exploring several questions
- What are the possible inputs and outputs of these virtue algorithms?
- What are the privacy concerns of citizens, as objects and consumers of these algorithms?
- How do they use the moral scores of others when making decisions?
- Can these algorithms be used as tools of social control by the state?

César A. Hidalgo
U of Toulouse,
Manchester,
Harvard

Augmented Society

Augmented Society

Goal is to understand the mechanisms that affect collective intelligence and learning, and to design technologies to improve collective intelligence.

Economic Complexity

The use of machine learning tools and techniques applied to problems of economic devlopment

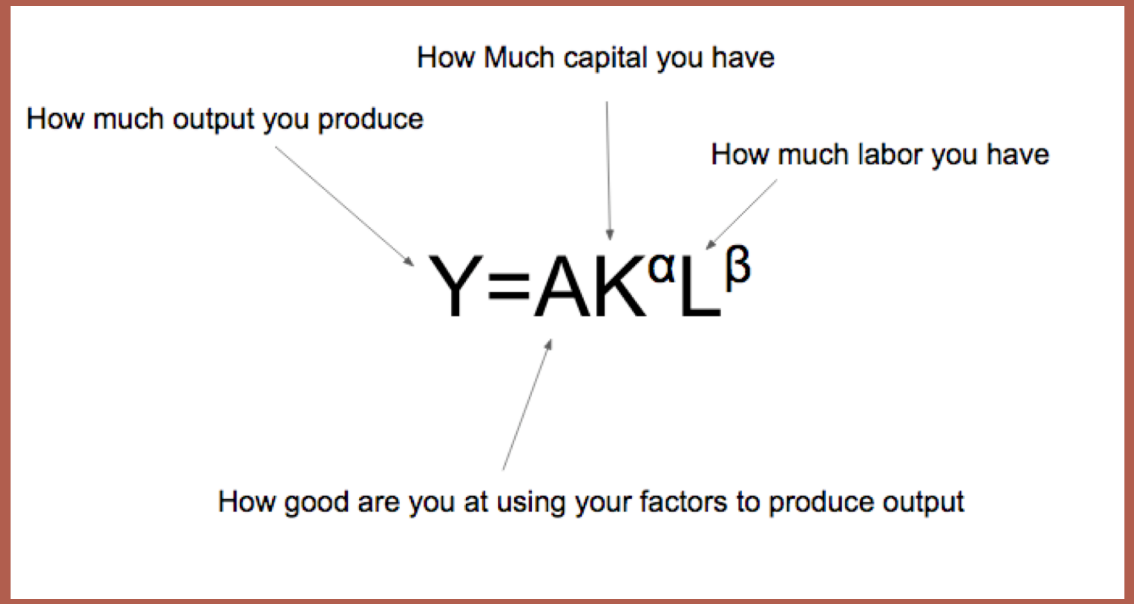
Human Perception of Technology

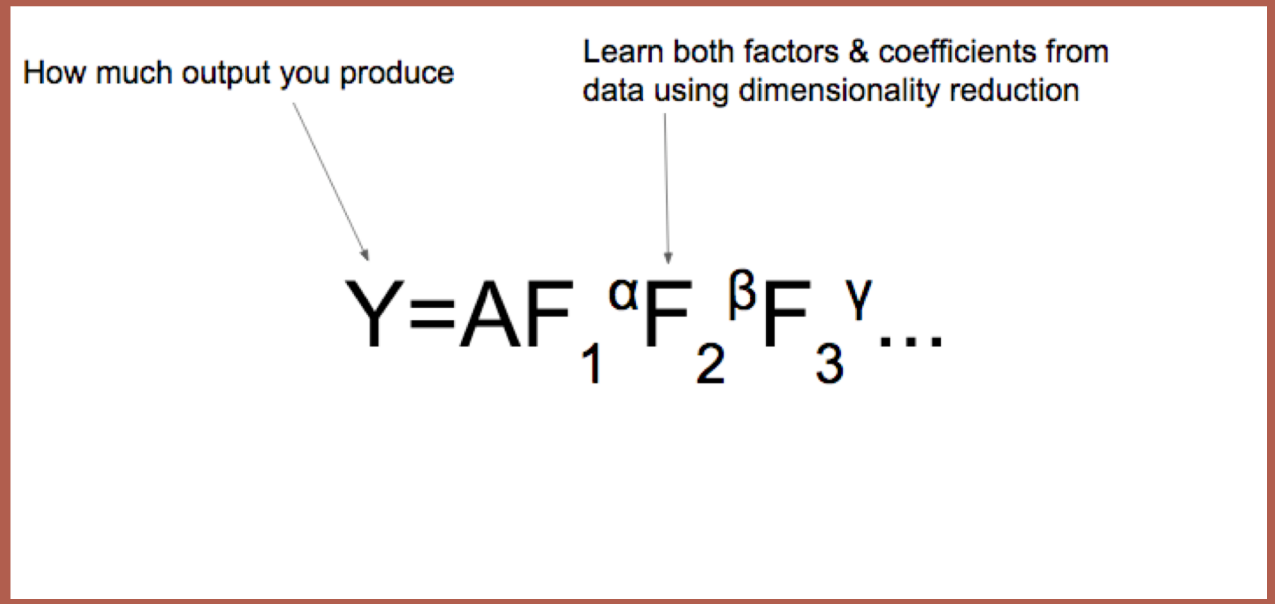
The experimental study of people's judgment of technology

Augmented Society

The creation of tools and processes to augment virtuous social participation

Economic Complexity





Traditional production functions, assume the existence of factors (e.g. capital, labor), and adjust coefficients (elasticities) to explain output as a function of these factors.

Economic complexity method attempt to learn both factors and coefficients directly from data using dimensionality reduction techniques (e.g. SVD).

2007 2009 2015 2020 First Field Review

Science Contents News Careers Journals

Read our COVID-19 research and news.

SHARE

RESEARCH ARTICLE

The Product Space Conditions the Development of Nations

C. A. Hidalgo[1,2,3], B. Klinger[2,3], A.-L. Barabási[1], R. Hausmann[2]

[1]* These authors contributed equally to this work.

* See all authors and affiliations

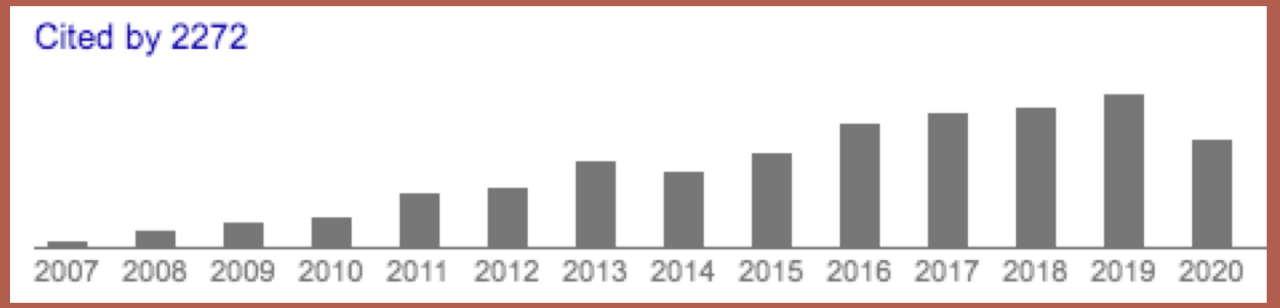
Science 27 Jul 2007: Vol. 317, Issue 5837, pp. 482-487 DOI: 10.1126/science.1144581

Article Figures & Data Info & Metrics eLetters PDF

You are currently viewing the abstract. View Full Text

Abstract

Economies grow by upgrading the products they produce and export. The technology, capital, institutions, and skills needed to make newer products are more easily adapted from some products than from others. Here, we study this network of relatedness between products, or "product space," finding that more-sophisticated products are located in a densely connected core whereas less-sophisticated products occupy a less-connected periphery. Empirically, countries move through the product space by developing goods close to those they currently produce. Most countries can reach the core only by traversing empirically infrequent distances, which may help explain why poor countries have trouble developing more competitive exports and fail to converge to the income levels of rich countries.

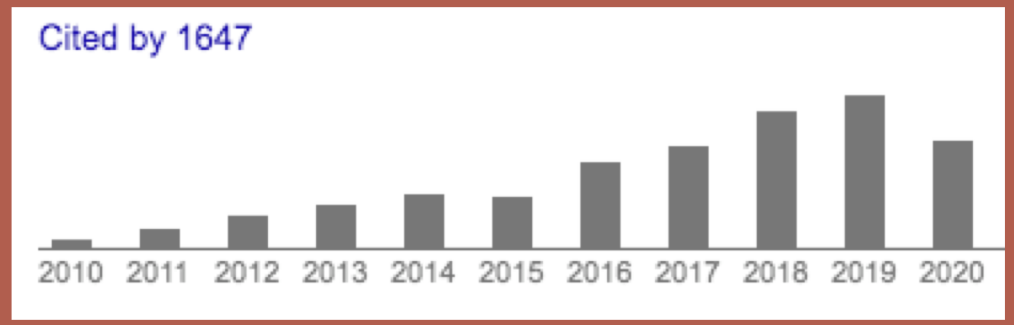


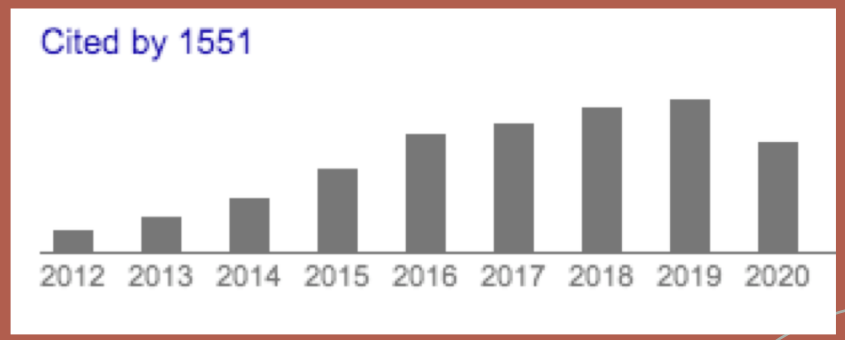
PNAS

The building blocks of economic complexity

RESEARCH ARTICLE

Abstract





nature reviews physics

Economic complexity theory and applications

César A. Hidalgo

Artificial and Natural Intelligence Toulouse Institute, University of Toulouse, Toulouse, France

Alliance Business School, University of Manchester, Manchester, UK

School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

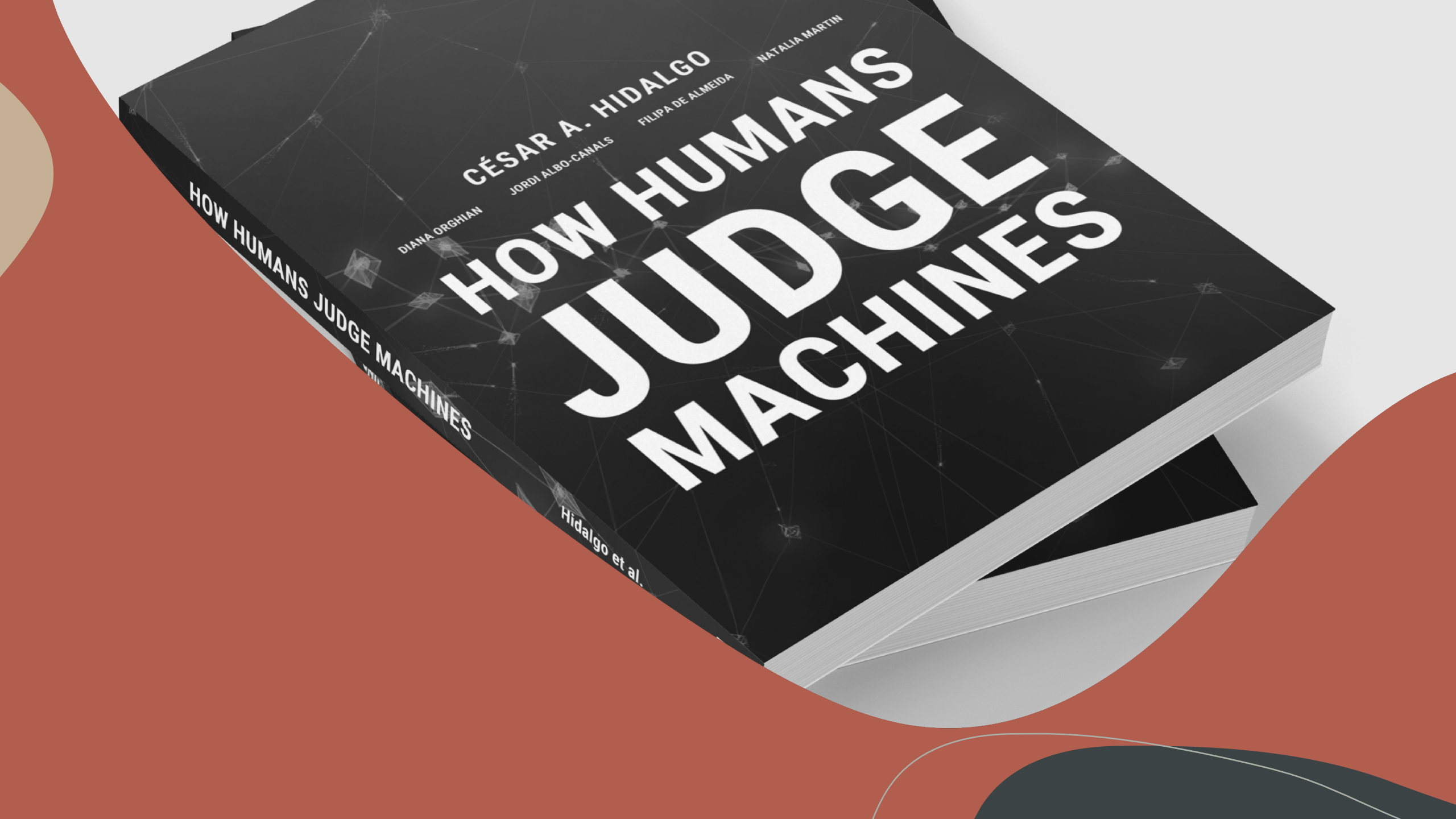
Email: cesifoti@gmail.com

Table-of-contents summary

Key points

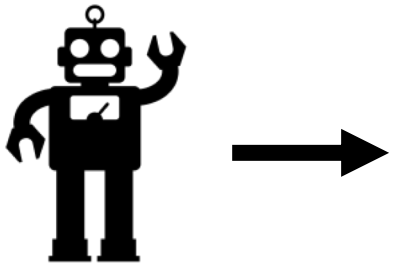
Abstract

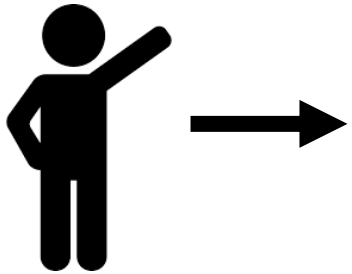
Economic complexity methods have become popular tools in economic geography, international development and innovation studies. Here I review economic complexity theory and applications with a particular focus on two streams of literature: the literature on relatedness, which focuses on the evolution of specialization patterns, and the literature on metrics of economic complexity, which is the application of dimensionality reduction techniques to specialization matrices.

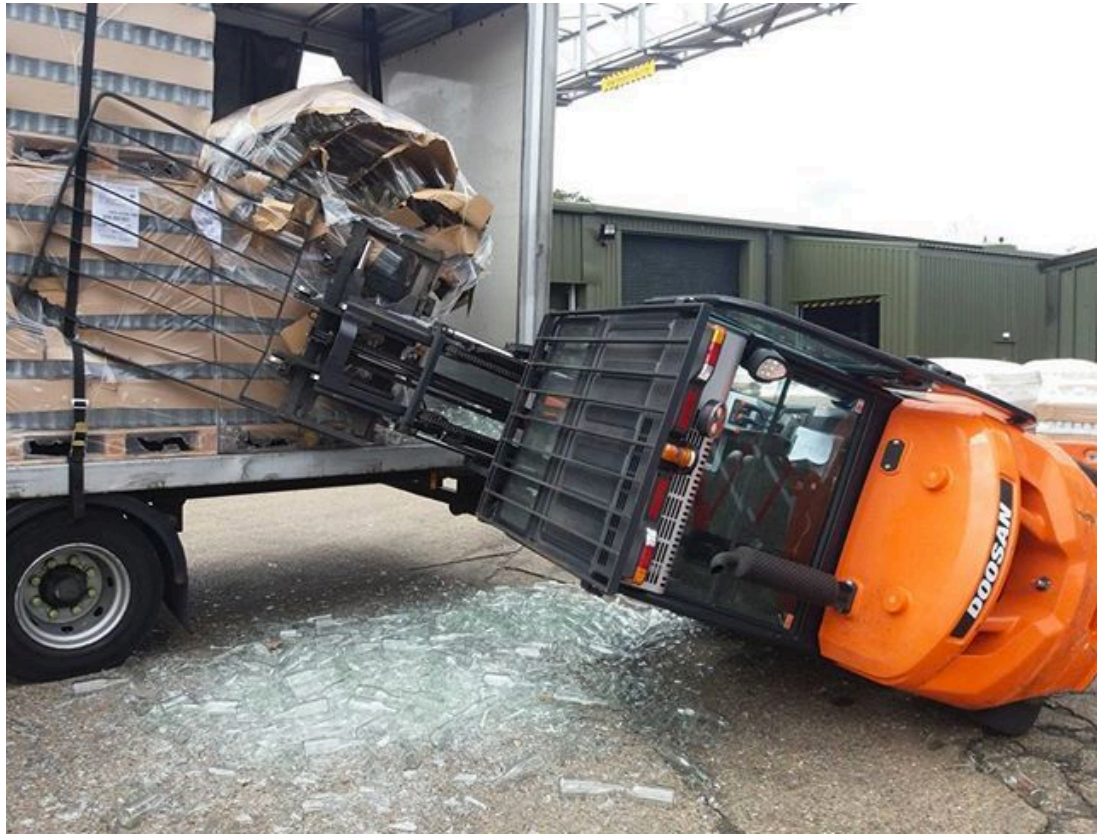

CÉSAR A. HIDALGO
JORDI ALBO-CANALS
FILIPA DE ALMEIDA
NATALIA MARTIN
DIANA ORGHIAN
HOW HUMANS JUDGE MACHINES
HOW HUMANS JUDGE MACHINES
Hidalgo et al.

Same Mistake

Same Reaction







=

Randomized Experiments

Assign people randomly to two groups of ~200 people each

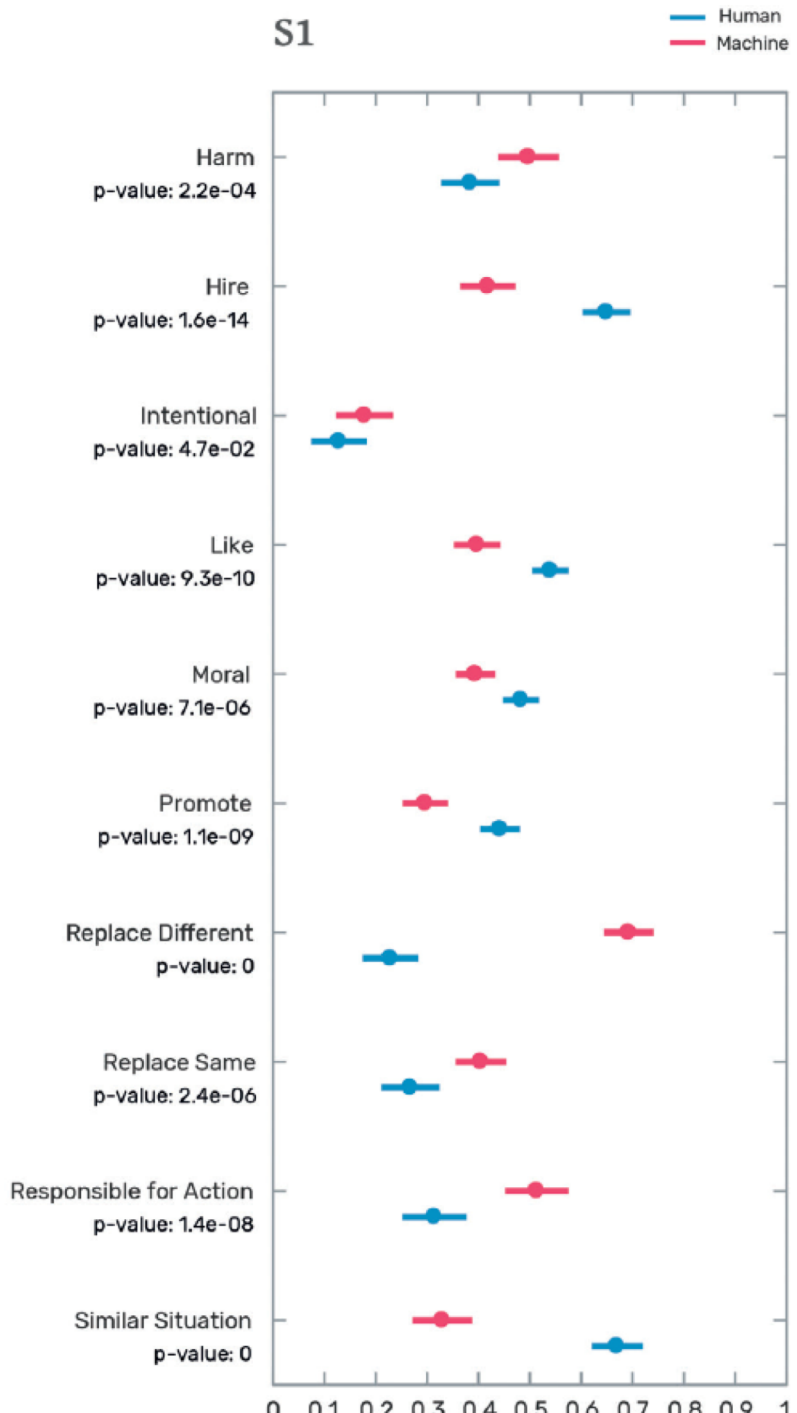
Each group sees the same exact scenario, except
that one group sees it as the action of a machine, while the
other, sees it as the action of a human

Differences in reactions tell us about biases that humans have for and against machines

Consider the following scenario

An excavator is digging up a site for a new building. Unbeknownst to the driver, the site contains a grave. The driver does not notice the grave and digs through it. Later, human remains are found.

Would you judge this differently if the driver was a **human** or a **machine**?

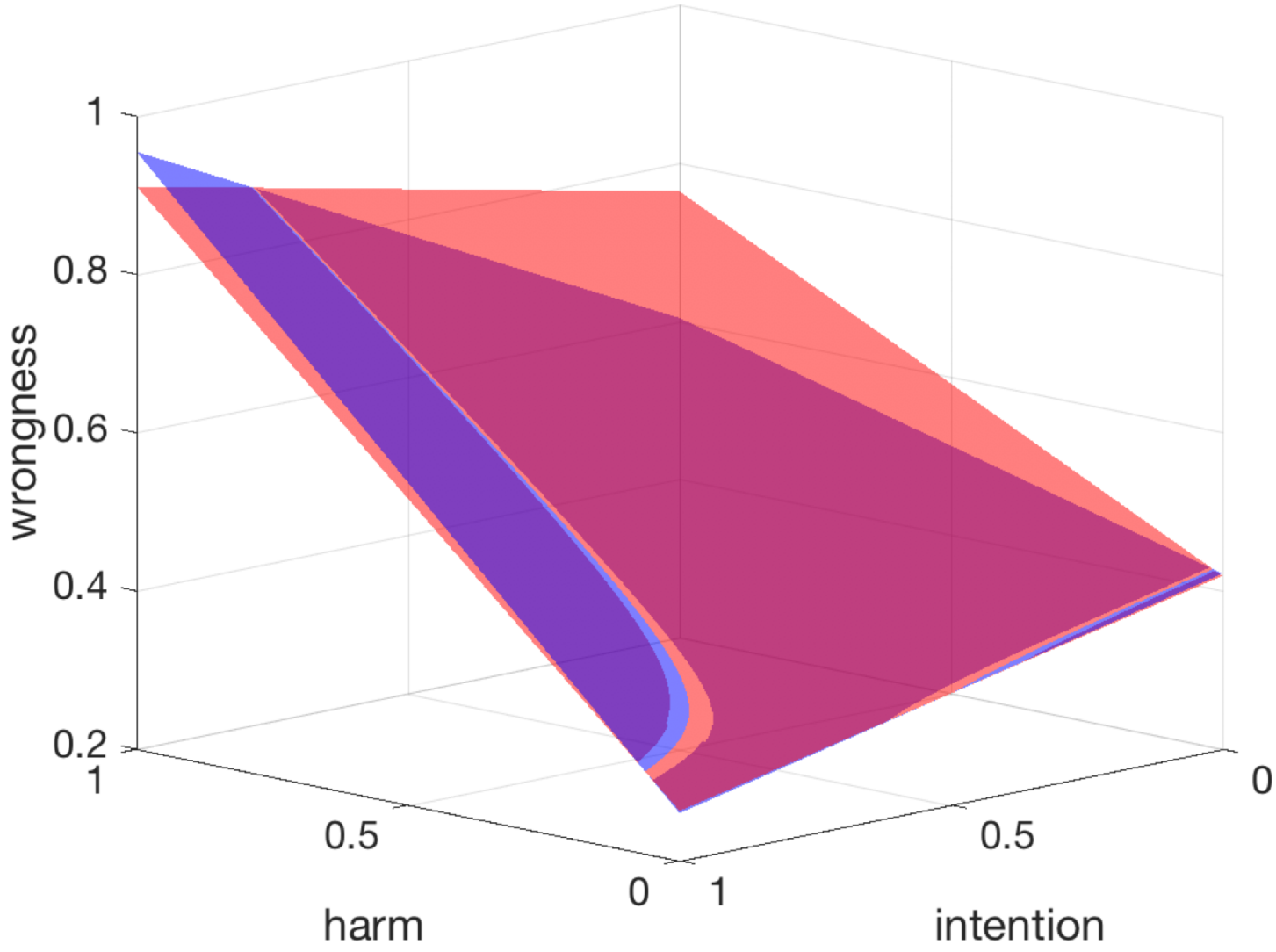


+80 scenarios

Dedicated chapters to algorithmic bias, privacy, & labor displacement.

Moral Functions

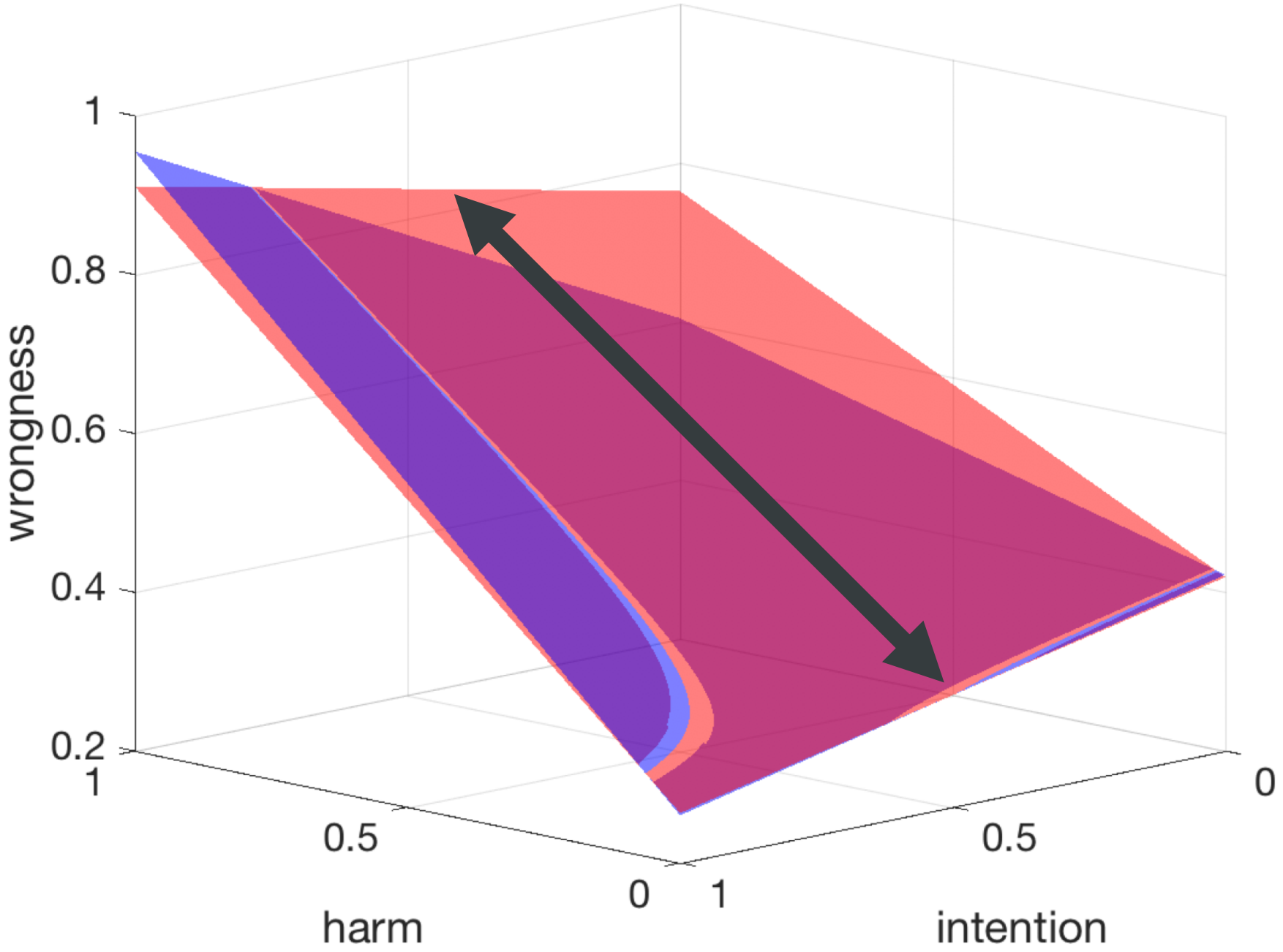
$$W=f_h(I,H)$$
$$W=f_m(I,H)$$



Moral Functions

$W=f_h(I,H)$

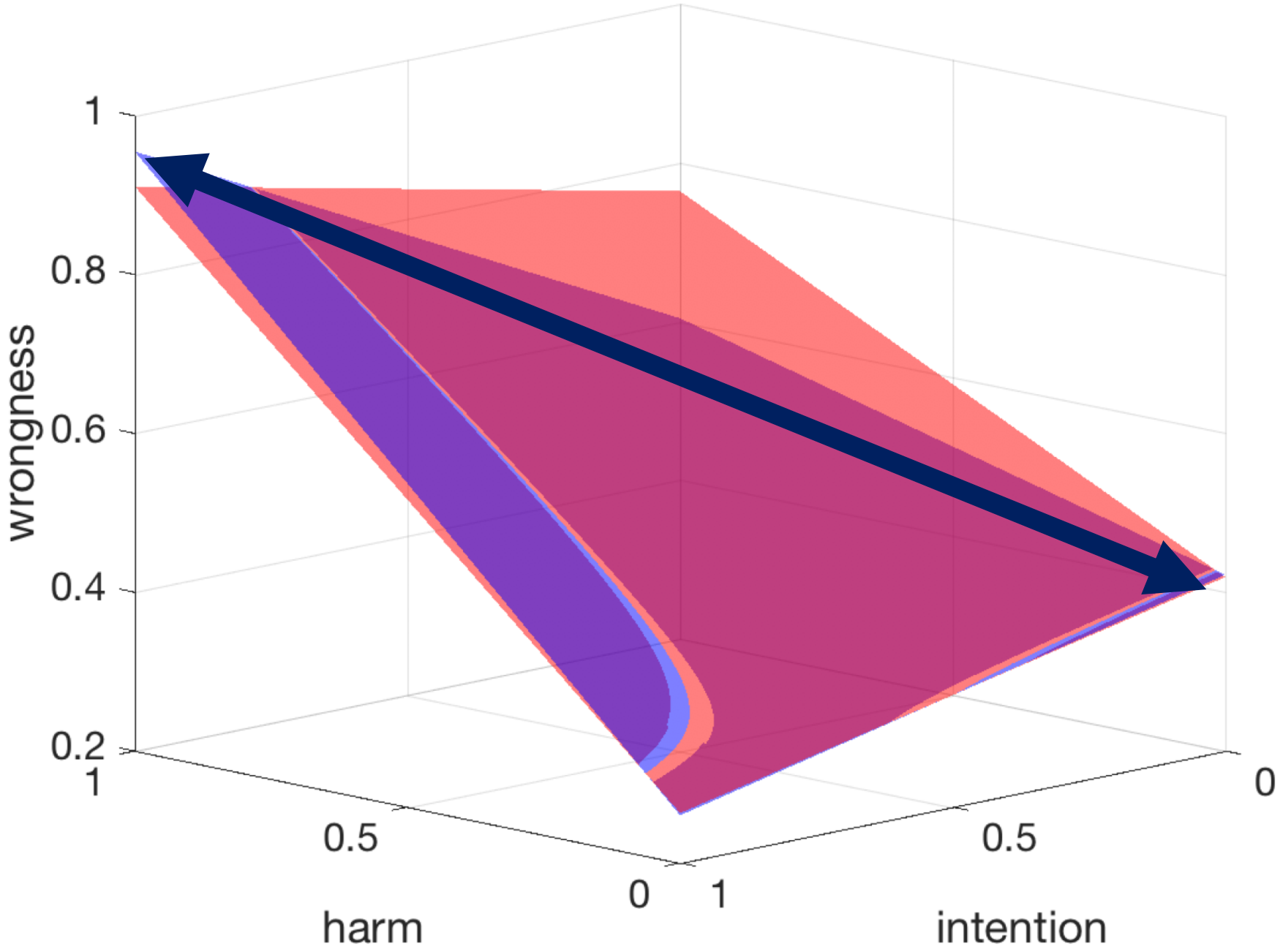
$W=f_m(I,H)$



Moral Functions

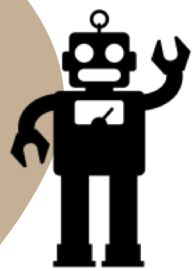
$W=f_h(I,H)$

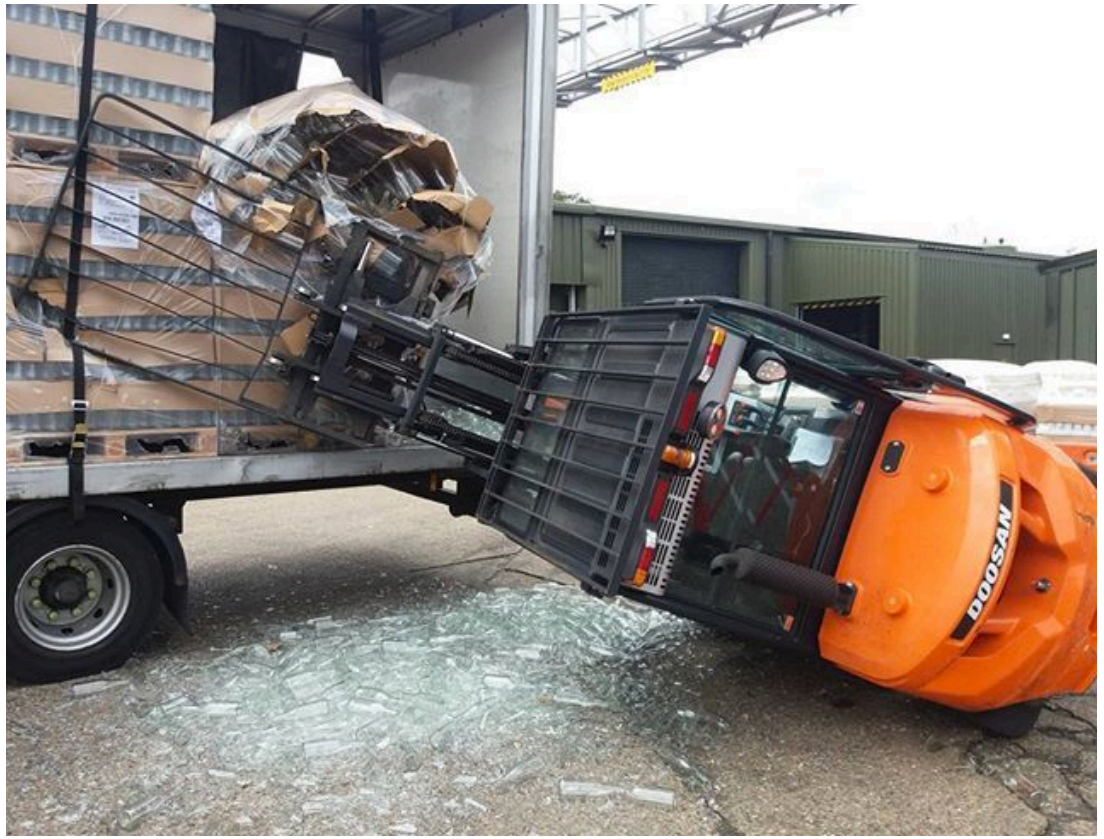
$W=f_m(I,H)$



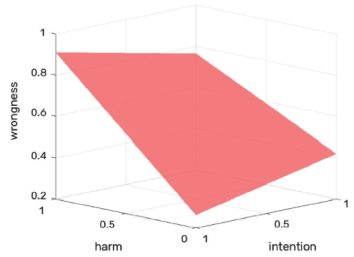
Same Mistake

Reaction



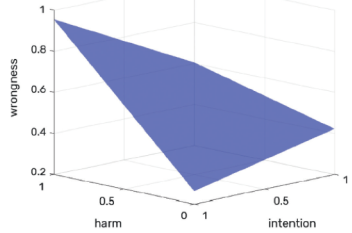


$f_h(\ldots)$

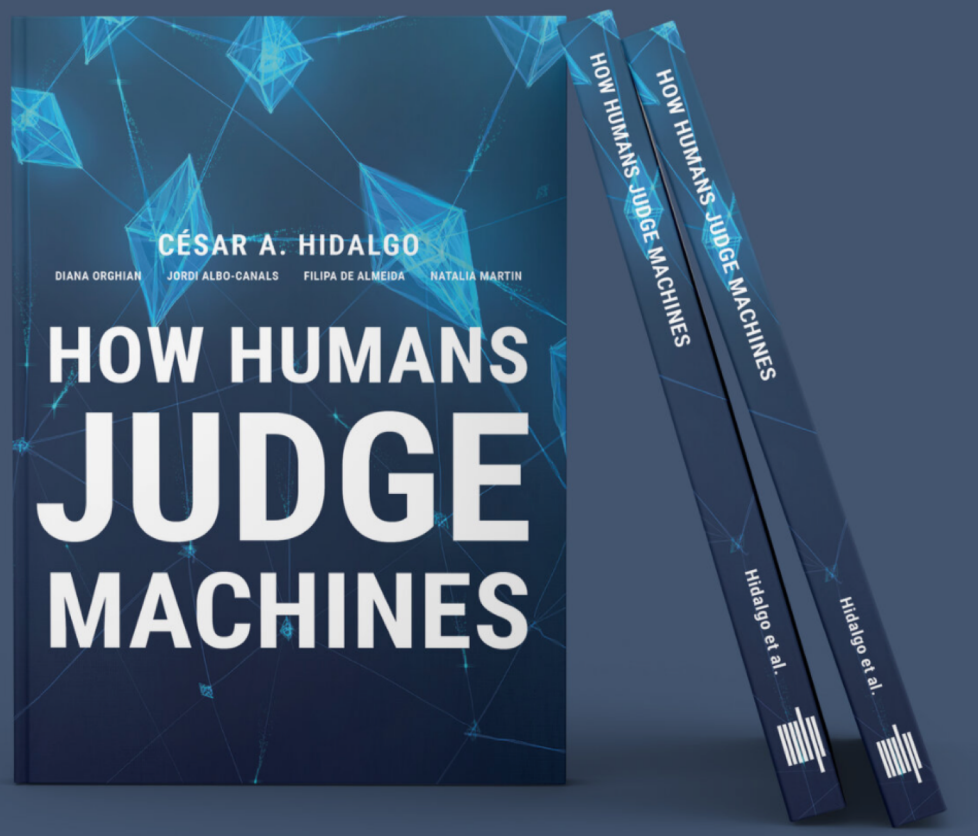


!=

$f_m(\ldots)$



People judge humans by their intentions, and machines by their outcomes



SEPT 29, 2020, 12pm EST

How would you feel about losing your job to a machine? How about an automated tsunami alert system that fails?

How Humans Judge Machines compares people's reactions to human and machine actions across dozens of experiments, revealing when and why humans are biased in favor or against machines.

Digital Edition
Free
Sept 29, 2020

Print edition
MIT press
Feb 2, 2021

2018 TED, Vancouver
A Bold Idea to Replace Politicians

OCTOBER, 2019, Santiago Chile

CHILECRACIA

más de 80 demandas ciudadanas

¿Qué priorizarías?

Tus votos: 102

1

Educación sobre uso de Redes Sociales

2

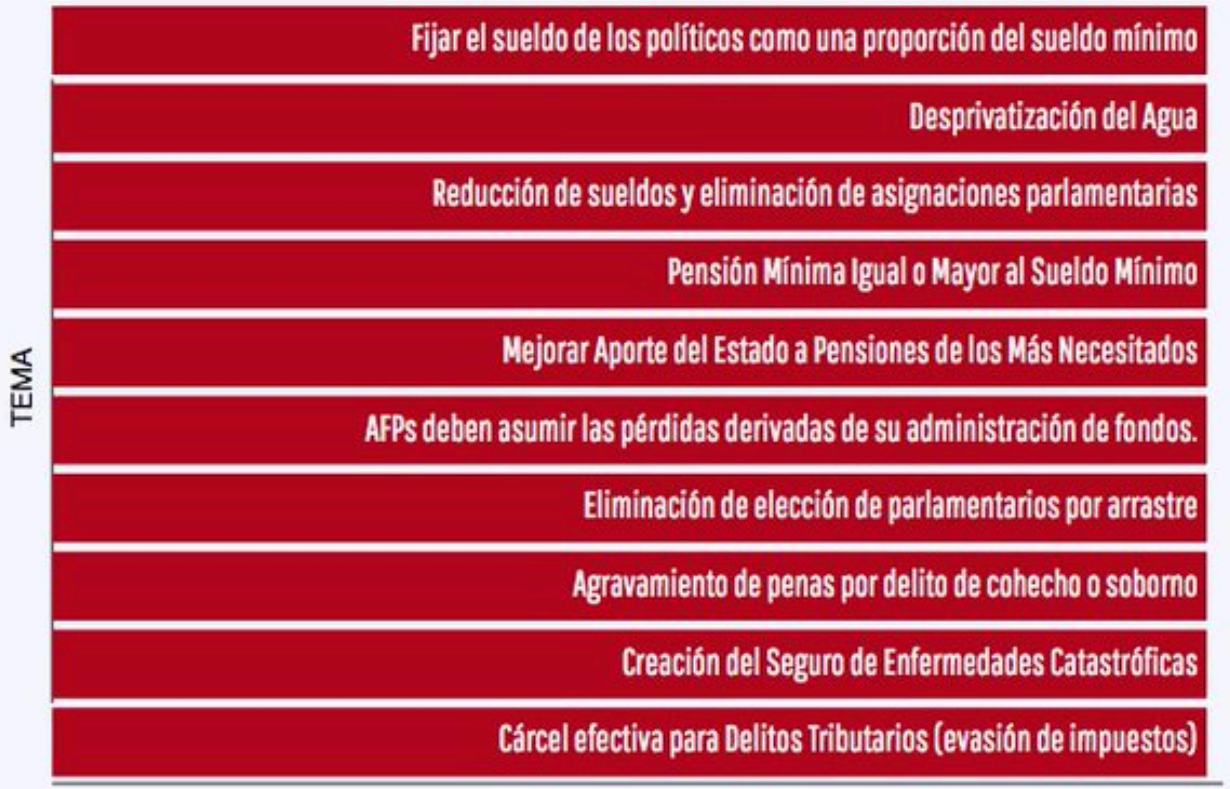
Desprivatización del Agua

NO TENGO PREFERENCIA

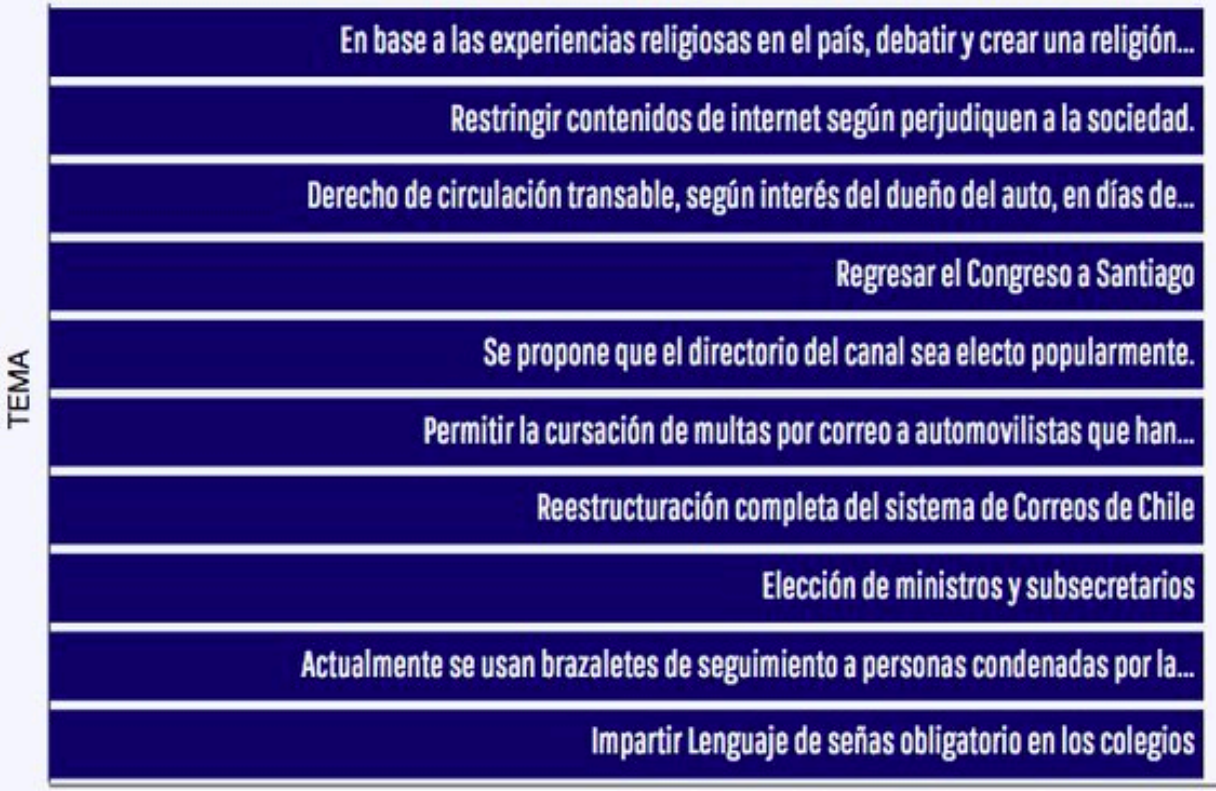
Algoritmo de ranking

TrueSkill | Eigenvector | % Victorias

Más relevantes

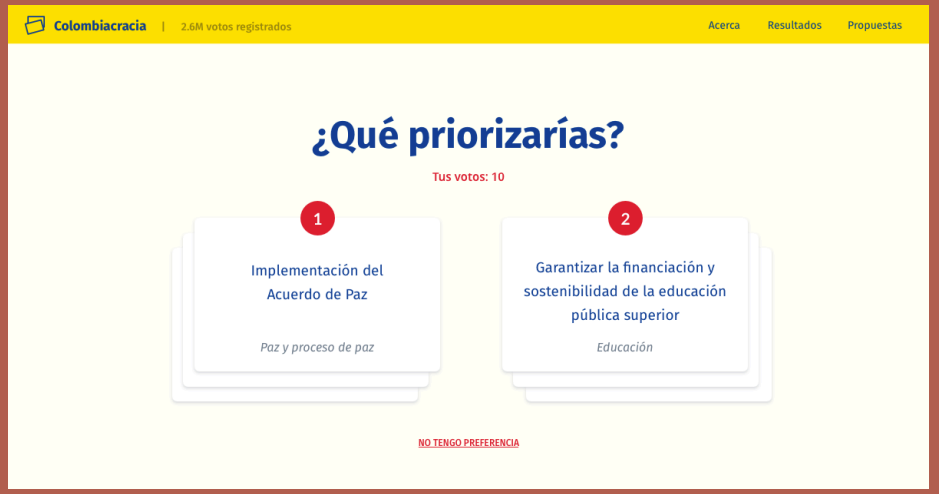


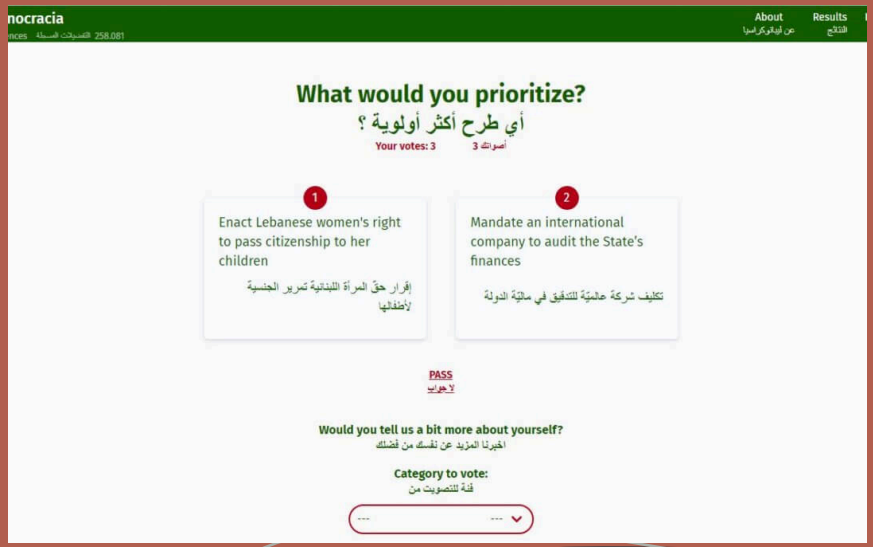
Menos relevantes



Chile

Colombia

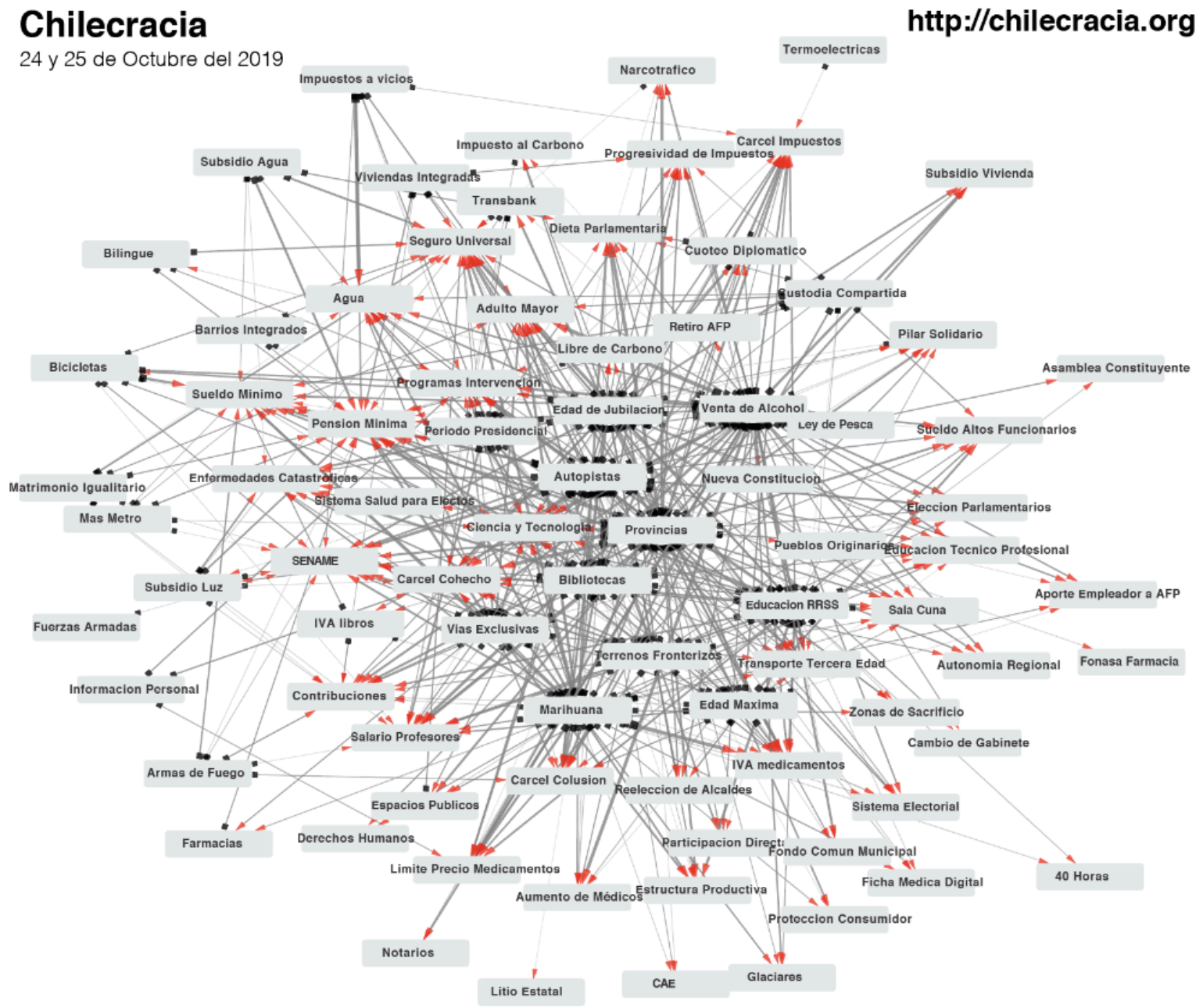




Georgia

Lebanon

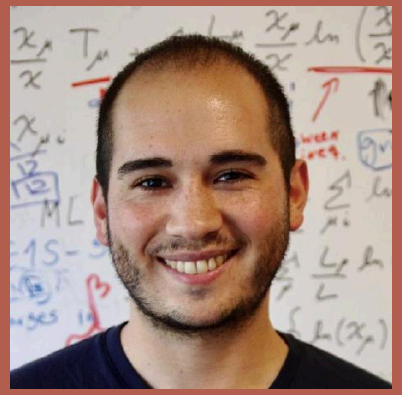
+8 million preferences
+100k users



Red de preferencias entre 90 demandas ciudadanas estimadas usando +1 millon de votos emitidos en chilecracia (chilecracia.org). Las flechas rojas apuntan a la propuesta ganadora del par cuando la probabilidad promedio de ganar es mayor a 75%.

Carlos Navarrete-Lizama
ANITI (IRIT PhD Student)

In collaboration with,
Cristian Jara-Figueroa (MIT)
Cristian Candia (Northwestern)



Working Together

The group is still young and we are getting to know each other. Nonetheless.

- Bonnefon and Hidalgo will explore the opportunity to write pieces on human perception and moral human machine interactions.
- The group will engage with other AI centers of excellence in France and the world in events at the intersection of AI, Psychology, and Economics.
- The group will explore other lines of collaboration, related to the use of AI in the economy, and applications of AI to understanding social and economic systems.