How to translate universal principles to local realities: The Chilean experience in AI

Dr. Sofía P. Salas

Center for Bioethics, School of Medicine

Clínica Alemana Universidad del Desarrollo

Santiago, Chile
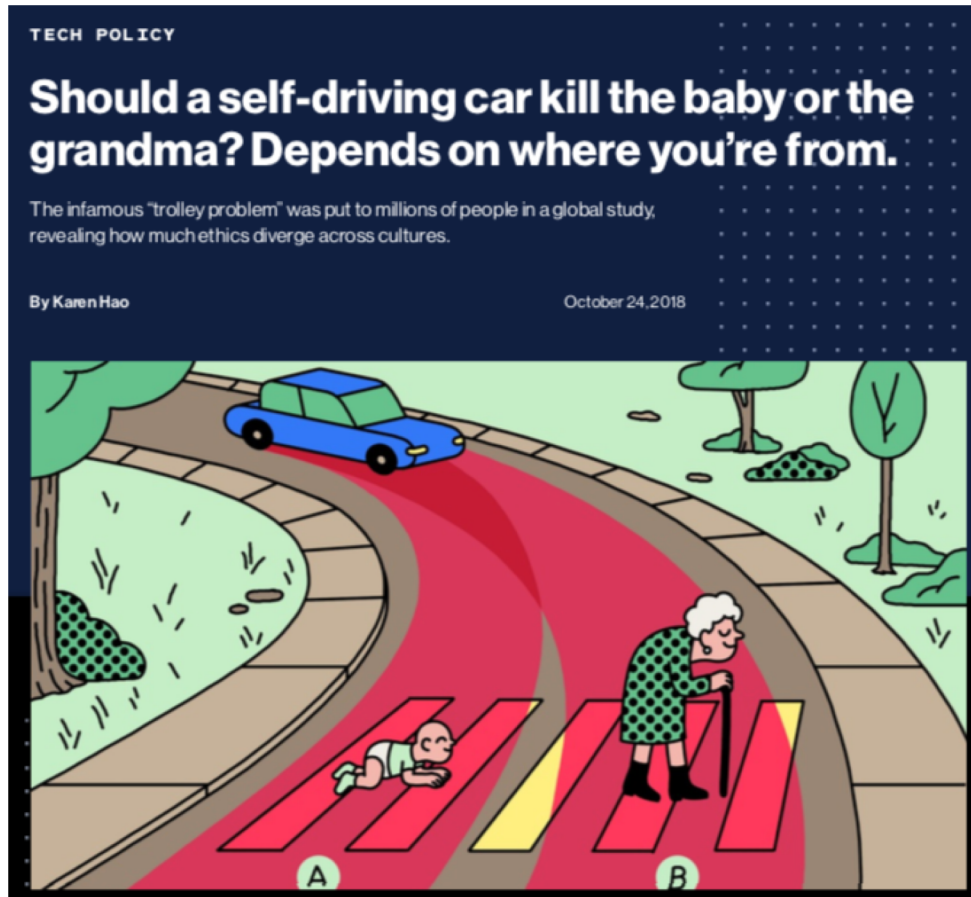
Cape Town. GFBR, November 2022

My agenda:

- Lessons from the moral machine experiment

- How to choose guidelines & frameworks

- The Chilean experience
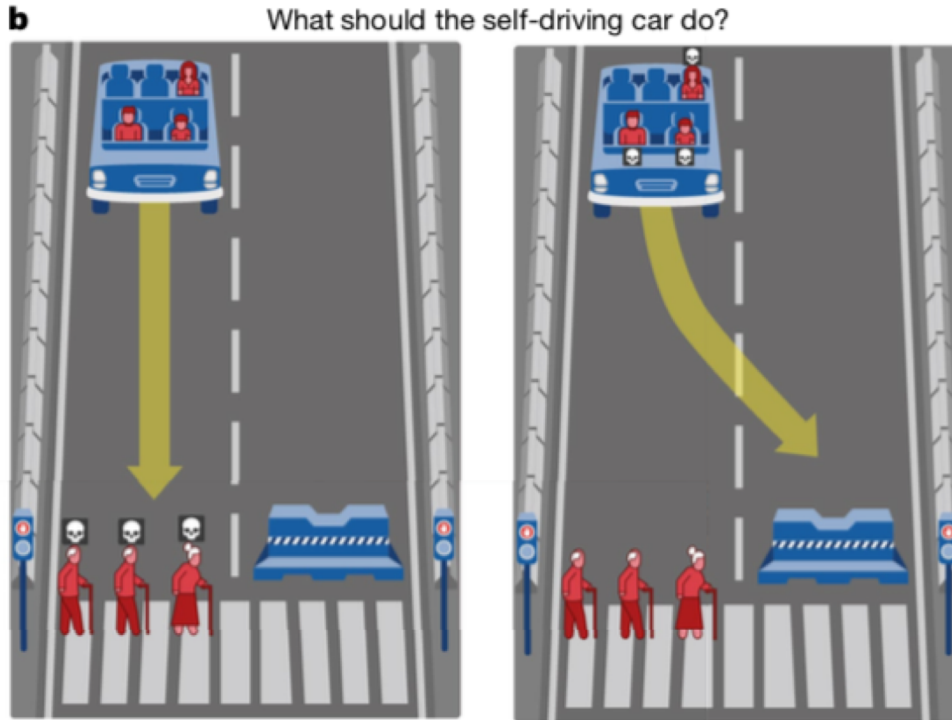
- Conclusions & recommendations

The AI version of the trolley dilemma

TECH POLICY

Should a self-driving car kill the baby or the grandma? Depends on where you're from.

The infamous "trolley problem" was put to millions of people in a global study, revealing how much ethics diverge across cultures.

By Karen Hao   October 24, 2018

- Should a self driving car avoid a pedestrian in the road or deliberately hit a single pedestrian if it means avoiding a larger group of people?

- The decision about killing the baby or the grandma is an updated version of the old trolley dilemma.

Image from https://www.inc.com/magazine/201811/tom-foster/artificial-intelligence-ethics.html

Facultad de Medicina
Clínica Alemana · Universidad del Desarrollo

Results of the moral machine experiment
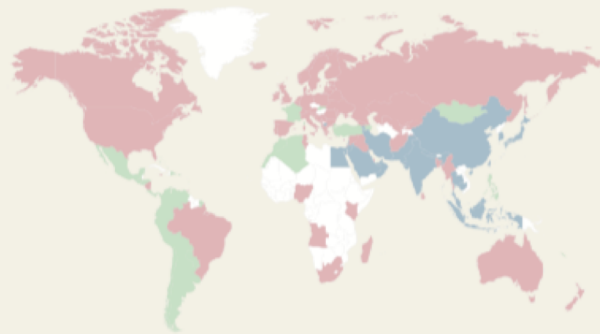

What should the self-driving car do?

- Online experimental platform.
- 40 million decisions from millions of people in 233 countries/territories.
- Different scenarios about what should the self-driving car do (ie. kill the baby or the grandma).
- The data generated provided insight into the collective ethical priorities of different cultures.
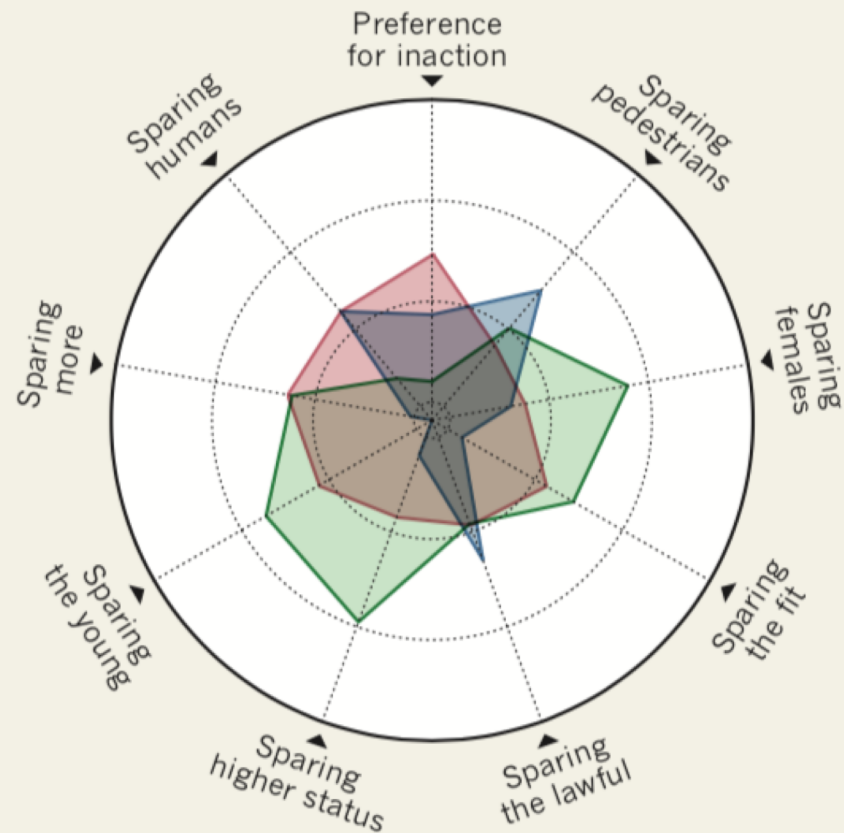
Facultad de Medicina
Clínica Alemana - Universidad del Desarrollo

Results indicate that there is no universal moral code
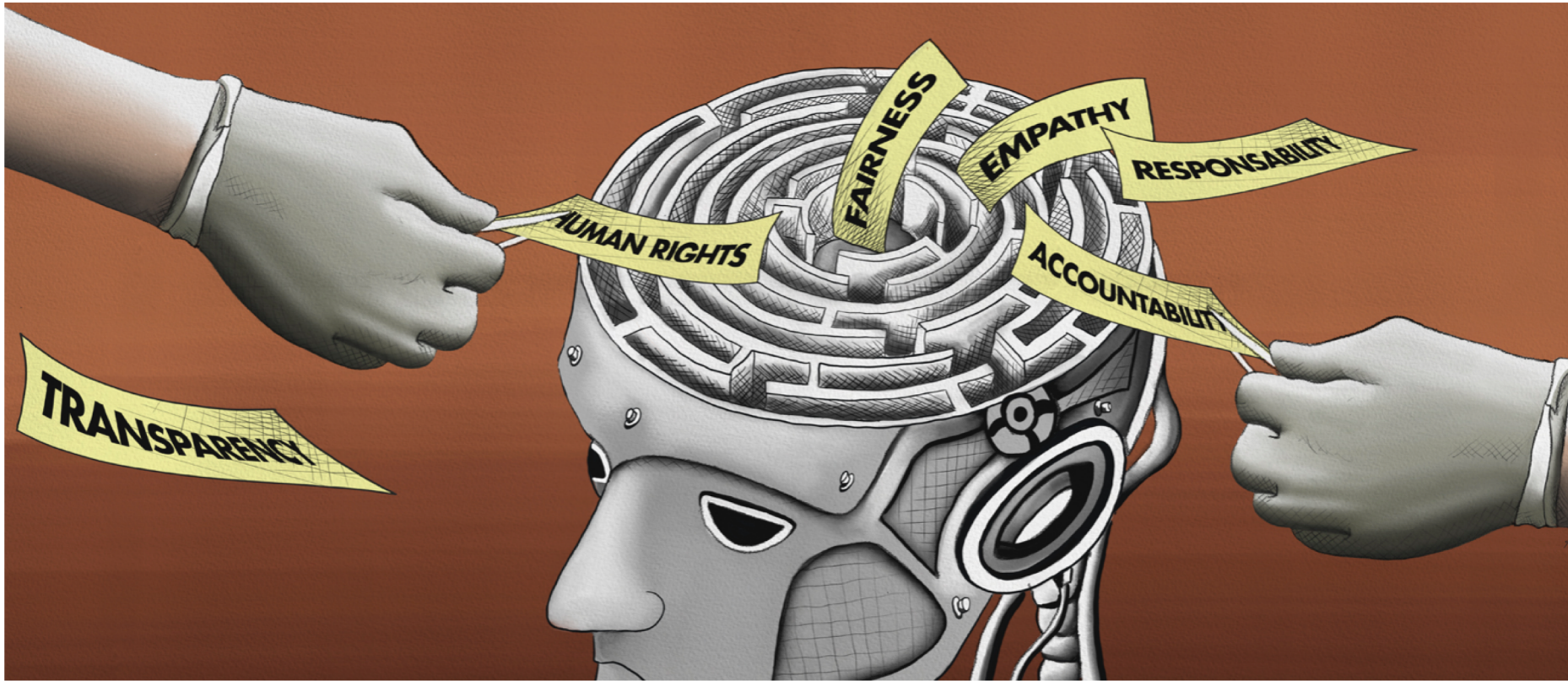

MORAL COMPASS

A survey of 2.3 million people worldwide reveals variations in the moral principles that guide drivers' decisions. Respondents were presented with 13 scenarios, in which a collision that killed some combination of passengers and pedestrians was unavoidable, and asked to decide who they would spare. Scientists used these data to group countries and territories into three groups based on their moral attitudes.

Preference for inaction · Sparing pedestrians · Sparing females · Sparing the fit · Sparing the lawful · Sparing higher status · Sparing the young · Sparing more · Sparing humans

Western · Eastern · Southern

Facultad de Medicina
Clínica Alemana · Universidad del Desarrollo
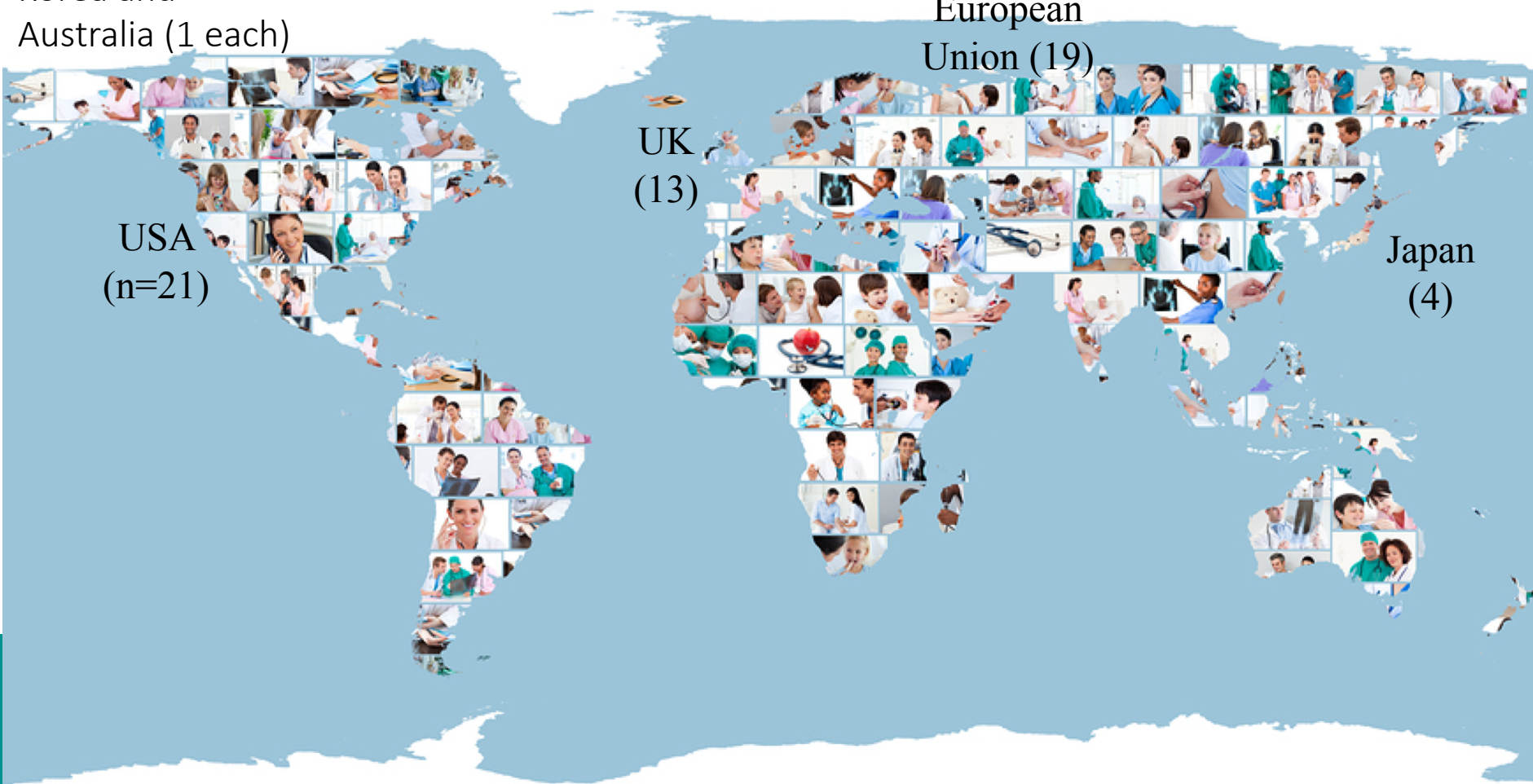
The challenges of machine ethics

World distribution of AI guidelines

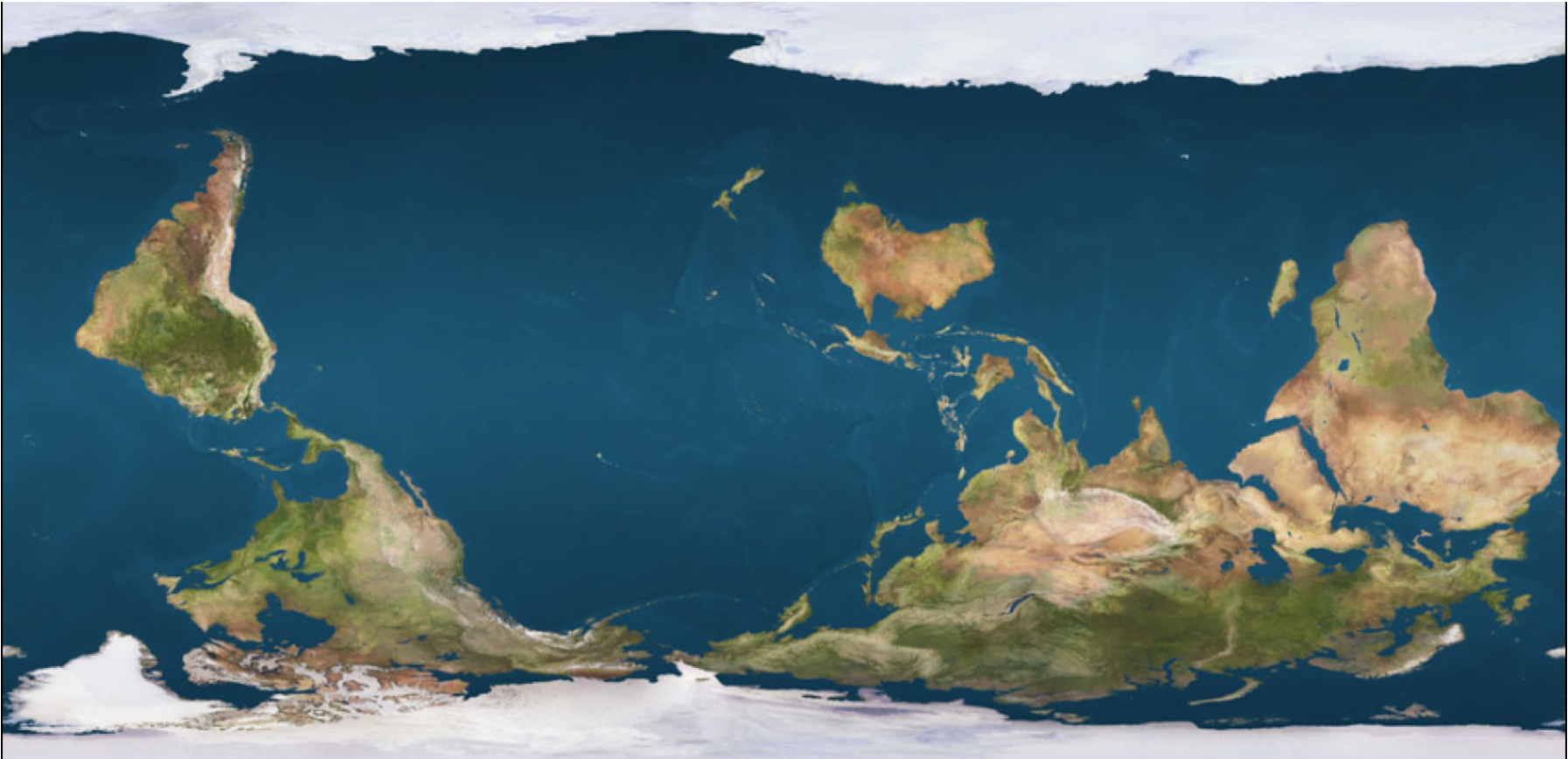Canada, Iceland, Norway, the United Arab Emirates, India, Singapore, South Korea and Australia (1 each)
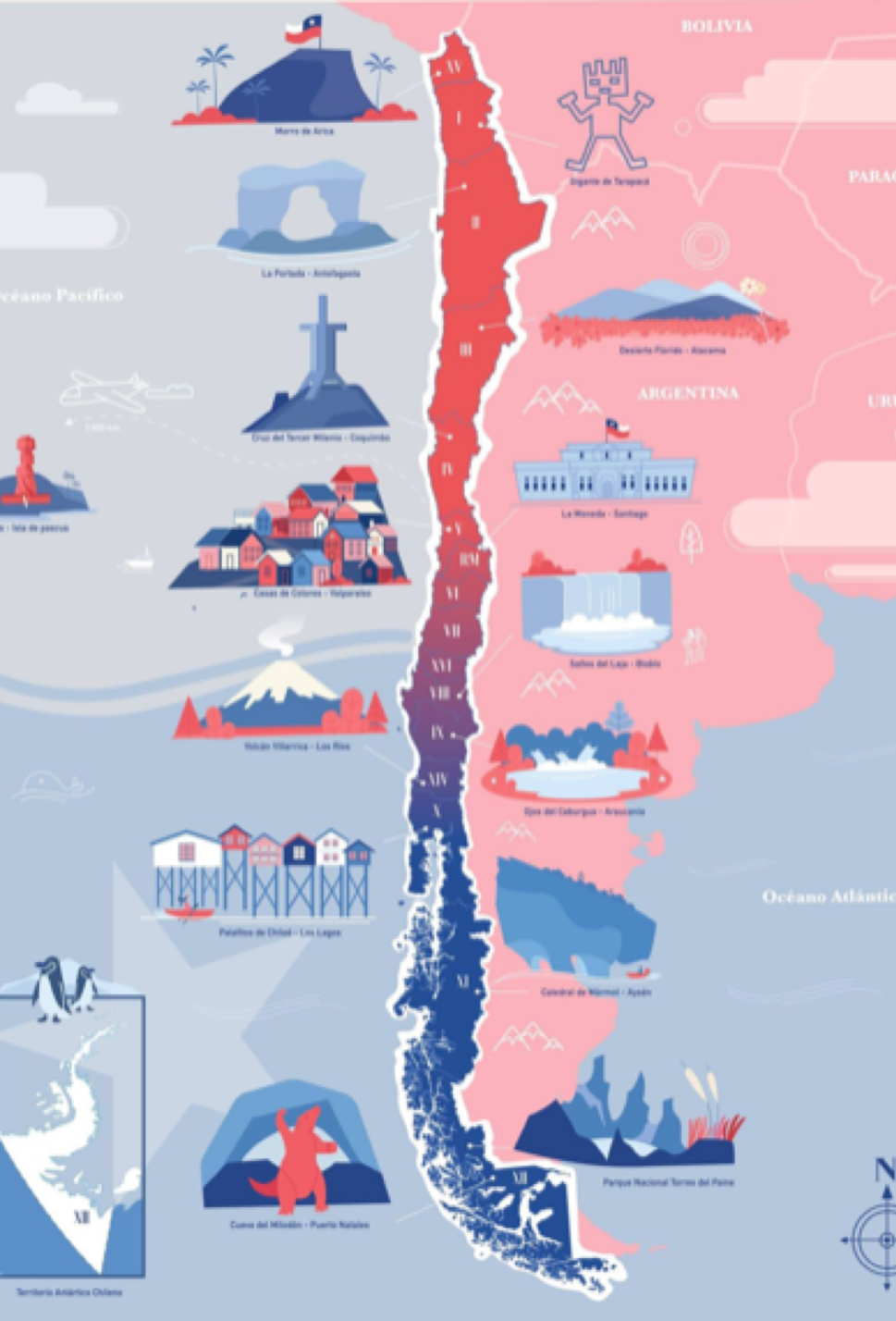
European Union (19)

UK (13)

USA (n=21)

Japan (4)

South – up map orientation[1]

[1] https://en.wikipedia.org/wiki/South-up_map orientation

Facultad de Medicina
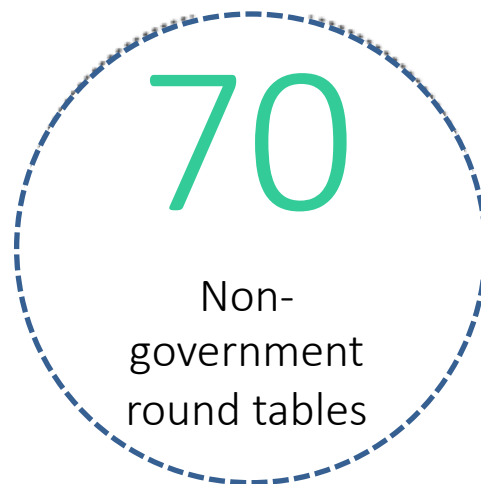Clínica Alemana · Universidad del Desarrollo

Chile needed to have local standards on the development and use of AI in order to position the country as a world hub for data science, entrepreneurship and innovation ecosystem.
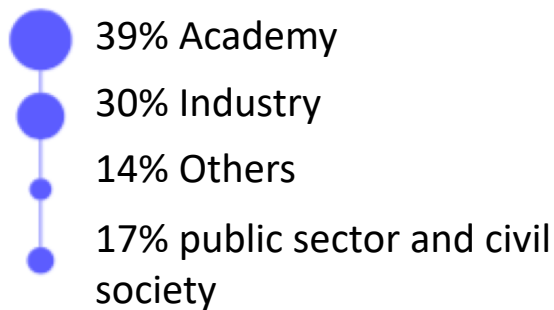
Initiative was organized by the Chilean National Council of Innovation for Development.

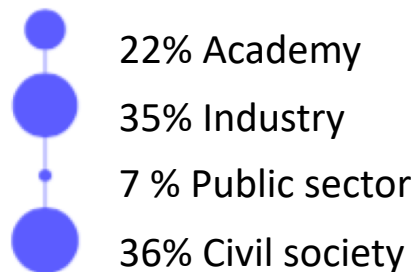AI was examined from multiple perspectives and disciplines.

From the beginning, there was an intention to be inclusive and participative.

69 Regional round tables

70 Non-government round tables

15 On-line thematic meetings

400 Participants
- 39% Academy
- 30% Industry
- 14% Others
- 17% public sector and civil society

1300 Participants
- 22% Academy
- 35% Industry
- 7 % Public sector
- 36% Civil society

Nearly **6600** participants

28 speakers

14   14

Gender equality

Figure 2: Synthesis of the participation process for the Chilean National AI policy

Facultad de Medicina
Clínica Alemana · Universidad del Desarrollo

Cross-cutting principles of Chilean AI policy

AI with focus on the well-being of people:
- respect to human rights
- security

AI for sustainable development:
- promote the use and development of technology
- Strengthen the ecosystem and the incorporation of AI as an axis of the sustainable development.
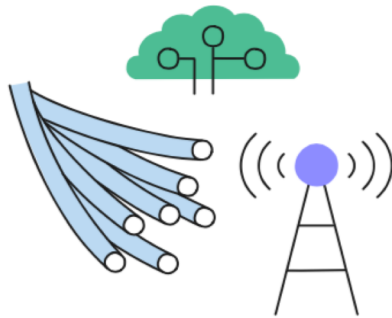
Inclusive AI:
- Manage possible biases
- With gender and cultural diversity

Globalized and evolving AI:
- Insertion of Chile in the global context
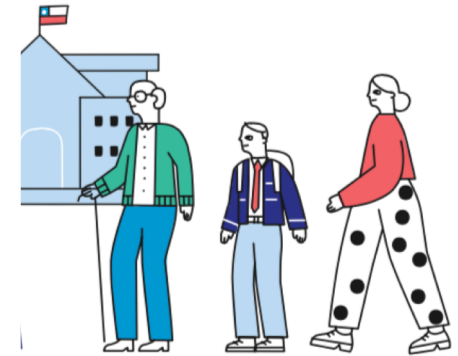- Aligned with agreements that Chile has signed (such as the OECD Principles on AI)

Main Chilean policy axis

Enabling factors

Development & adoption

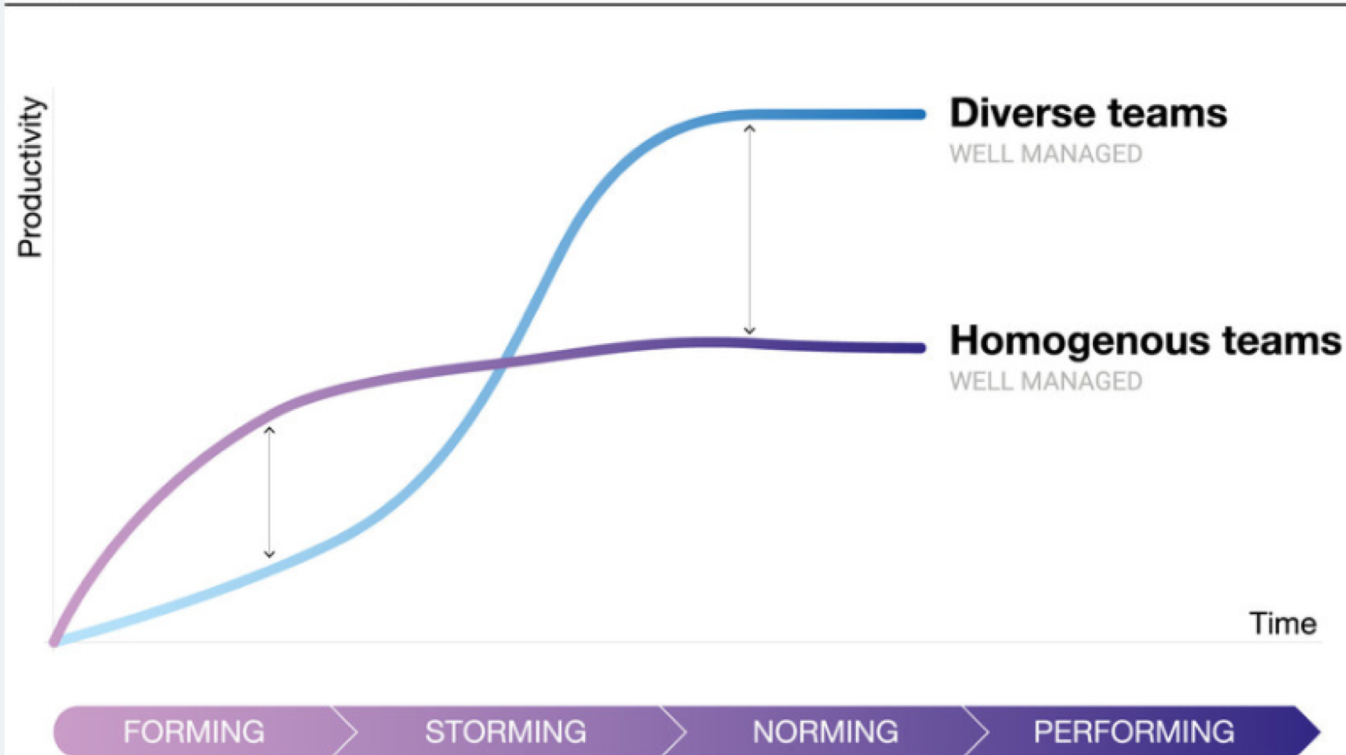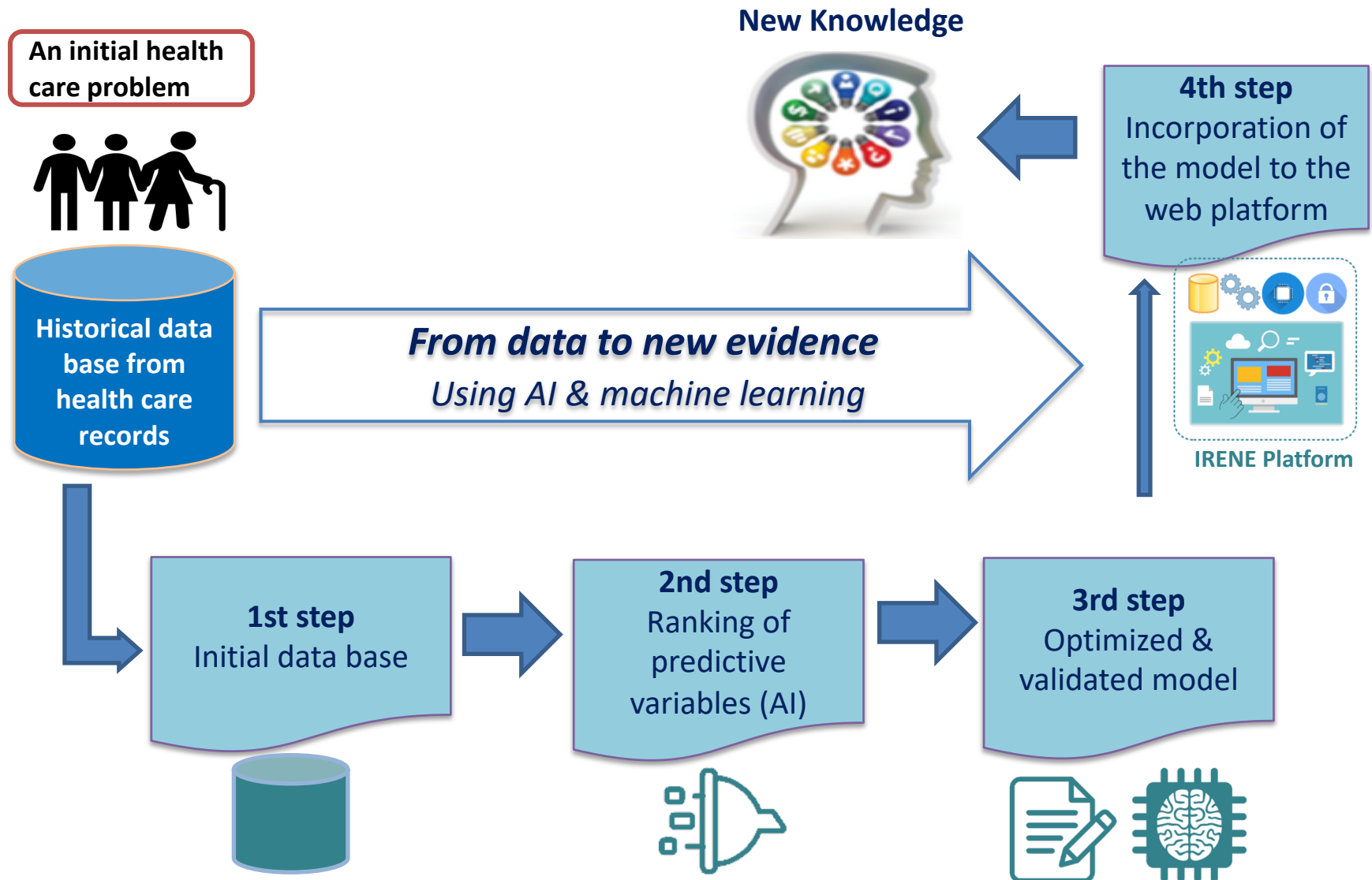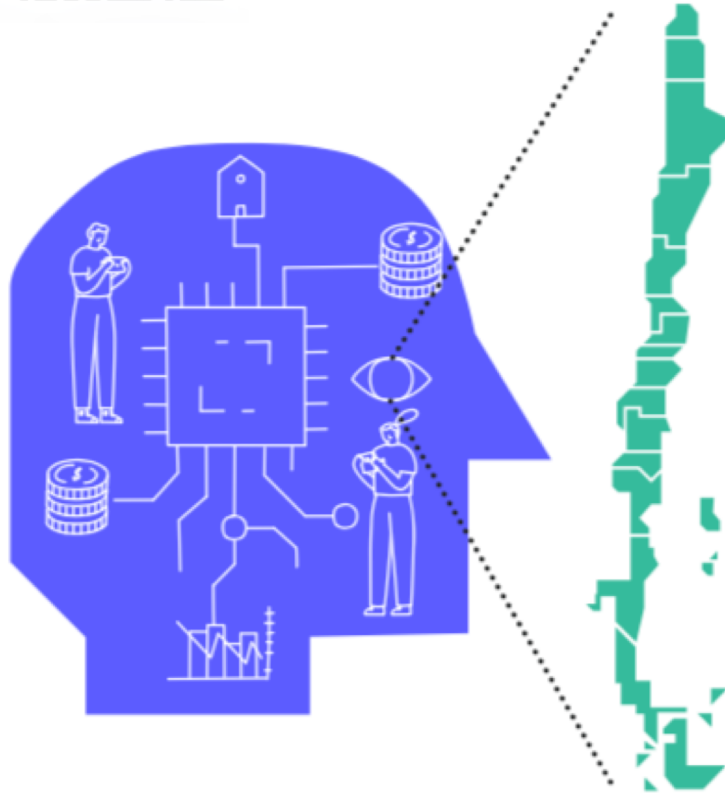Ethics, regulatory aspects, & socioeconomic effects

The importance of diversity

Figure 1: Impact of diversity on team performance
Source: Korn Ferry Institute, 2019

An example of a Chilean research based on AI

An initial health care problem

New Knowledge

Historical data base from health care records

From data to new evidence
Using AI & machine learning

4th step
Incorporation of the model to the web platform

IRENE Platform

1st step
Initial data base

2nd step
Ranking of predictive variables (AI)

3rd step
Optimized & validated model

Conclusions

- There is need to avoid "value imperialism".

- Each country should have/develop its own AI policy.

- This process should be participatory & inclusive.

Thanks, merci, gracias, gratzie, Vielen Dank, obrigada

Facultad de Medicina
Clínica Alemana · Universidad del Desarrollo