Democratic Resilience in the AI Age

03/26/2018	George Tilesch, Chloe Morin	London

WORLD LEADERSHIP ALLIANCE

CLUB DE MADRID

Introducing Ipsos Global Affairs

Global Top 3 research companies

Research & Advisory for the public interest

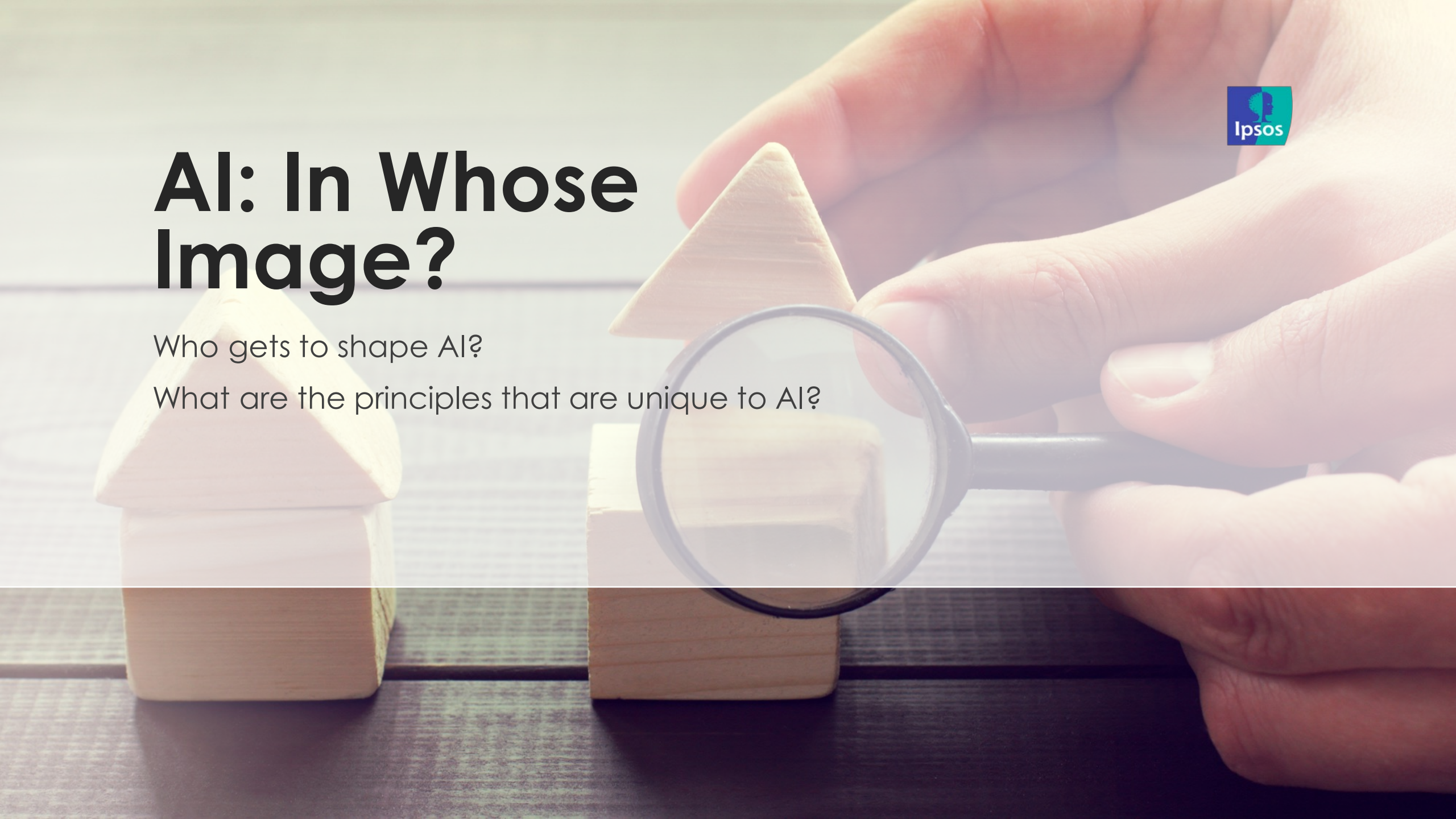
CEO: Najat Vallaud-Belkacem, former French Edu Minister

Global Affairs: Total Social Understanding

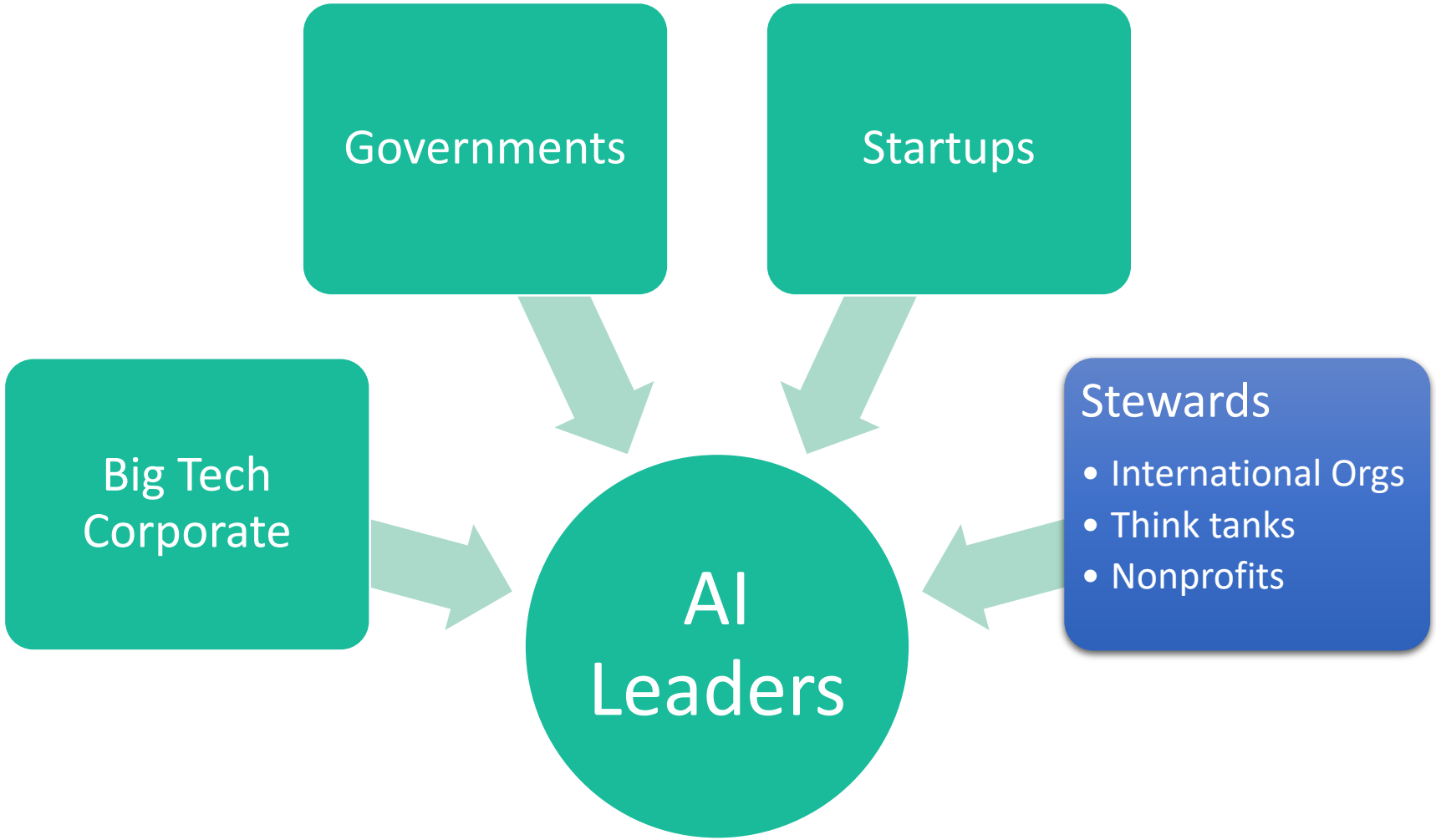
AI: In Whose Image?

Who gets to shape AI?

What are the principles that are unique to AI?



AI Forerunners: a Game of Unequals



Principles of AI Power: Why is AI Tech unprecedented?

Core

- AI is the Brain
- The core of Tech disruption in the 21st century & beyond

Imago

- Creators / Constructors call
- AI inherits moral priorities and omissions

Unfair Dominance

- Exponential Rewards
- Occupying niches pays off now

Endless Siege

- Penetration augmented 1000x
- Reaction → Optimization

Horror Vacui

- AGI is the Holy Grail
- Designed to be multi-purpose, omnipresent, self-augmenting

Inscrutable Infallibility

- Black Box (Intelligibility)
- Extended Black Box (Clueless Creator)

"AI is sufficiently informed, 'smart', autonomous and able to perform morally relevant actions independently of the humans who created them […]. (Floridi and Sanders 2004)"

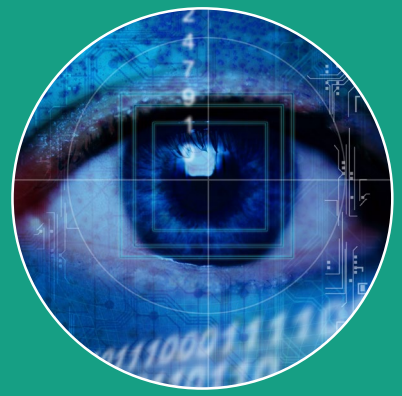
AI: Doubts & Threats

The threats of AI are a certainty. The boons of AI are a possibility that we need to work hard for.

The AI Threats & Challenges Framework

AI-supported Bias
Hardwiring Consciousness, Compassion
Cognitive Dependence & Resilience
Data: Ownership, Consent, Privacy
Delegation & Human supervision
Distributed Responsibility
Automation: Job/Tax loss
Misinformation/Democratic collapse
People-centeredness/ Values embedding
Trust & Identity

Exposing AI's Dark underbelly - Examples



Curtailing Freedom

- The lure of the Surveillance State
- Cognitive automation and distortion, loss of free will

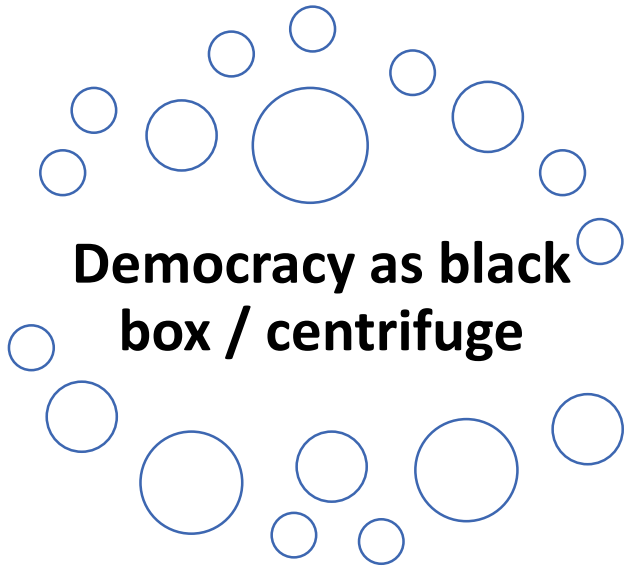
Magnifying Inequality

- Exponential concentration of data and resources
- Rise of the "useless class" etc.

Dismantling Trust

- Weaponized AI, "deepfakes"
- Mind hacks, proliferated AI cyberwarfare

A Framework for Democratic Resilience



Democracy as black box / centrifuge

- Trust has fallen, Empathy gap (tribalization)
- Fragmentation of Commons
- Fear of losing one's place in society, becoming useless
- Responsibility is diluted or nonexistent
- Citizens as a multitude, and not a mass
- Complexity of issues, difficult to grasp, but opinionated & emotional
- Cognitive bias / unconscious constraints on our choices

Principles & values to guide policymaking and action

- Simplicity, calculability, intelligibility (starts with AI Literacy)
- Accountability (participation, but also co-creation, proofs by results)
- Openness/ Cognitive Resilience (Make constraints on our choices and behaviours visible)
- Agility/Fluidity/Deep Listening

Pillars of democratic resilience

- Foster Trust and Empathy (Fraternity)
- Create a digital concept of citizenship, options & rights mindset (Liberty)
- Foster & reward debates that are more productive, inclusive and meaningful (Equality and reciprocity)

Towards AI Resilience

The AI promise to democratic guardians:

Best practices and how to make them work

A framework for Democratic Resilience

For guardians of democracy

Liberty

- Simplicity, transparency, building mental resilience (vs. cognitive biases/manipulation) are the principles that should dictate policymaking and actions, in order to empower citizens to participate in the public debate; make free and informed choices.

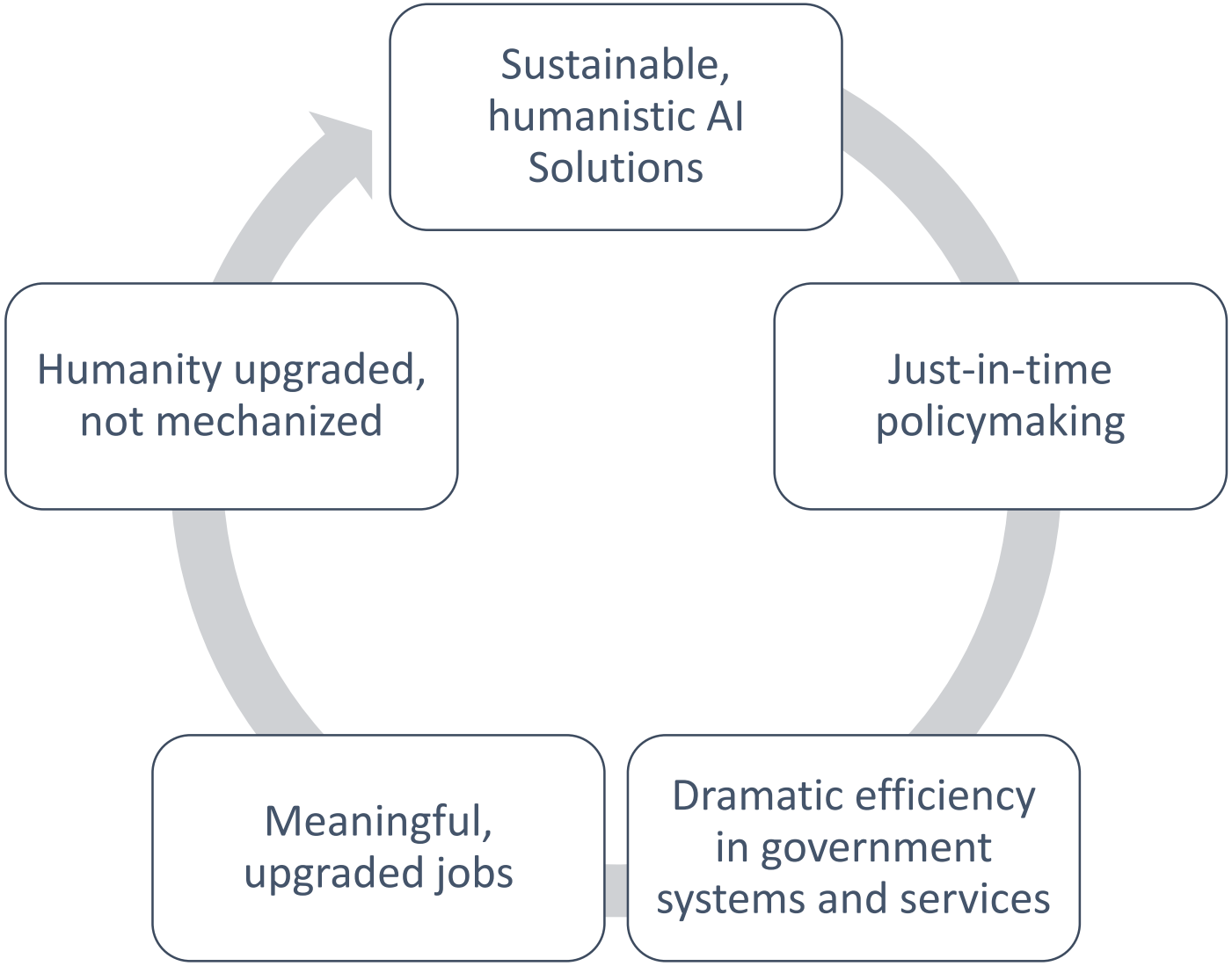
Fraternity

- Combatting the forces of social fragmentation and tribalization, by bridging the Elites/people distrust gap and fostering understanding and empathy.
- It is necessary in order to create the conditions for debate, and building consensus, which is the foundational trait of democracy.

Equality

- Inclusiveness of debates (accessible, simple, transparent
- Accountability and shared responsibility of all actors involved and fairness are the principles that should empower efficient debates.

The AI promise to democratic guardians



The AI promise to democratic guardians: Best practices

Promoting efficiency, openness and accessibility. New Zealand's Service innovation lab.

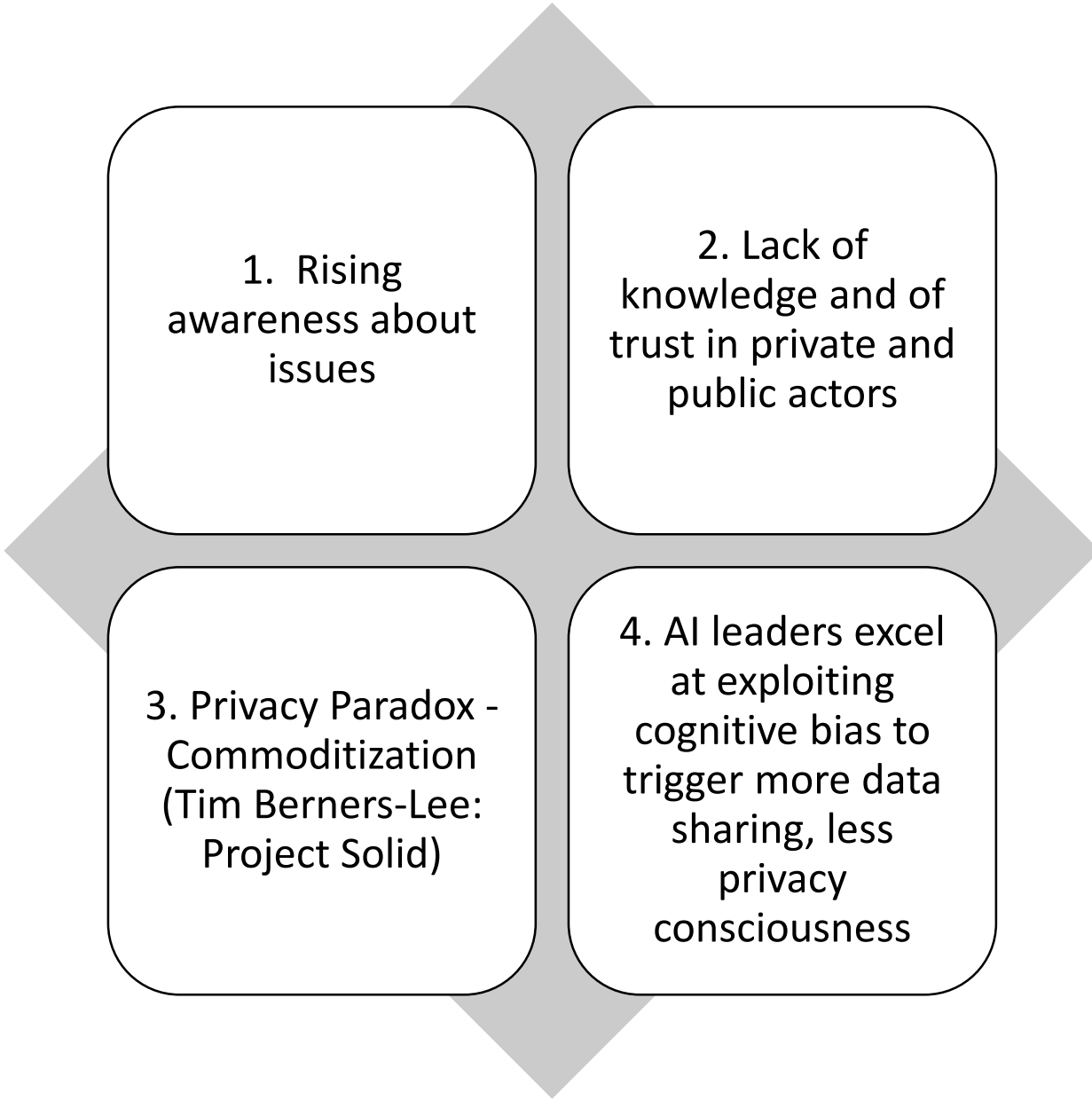
Improving access to public service, and services efficiency: Use of Big Data in Slovenia

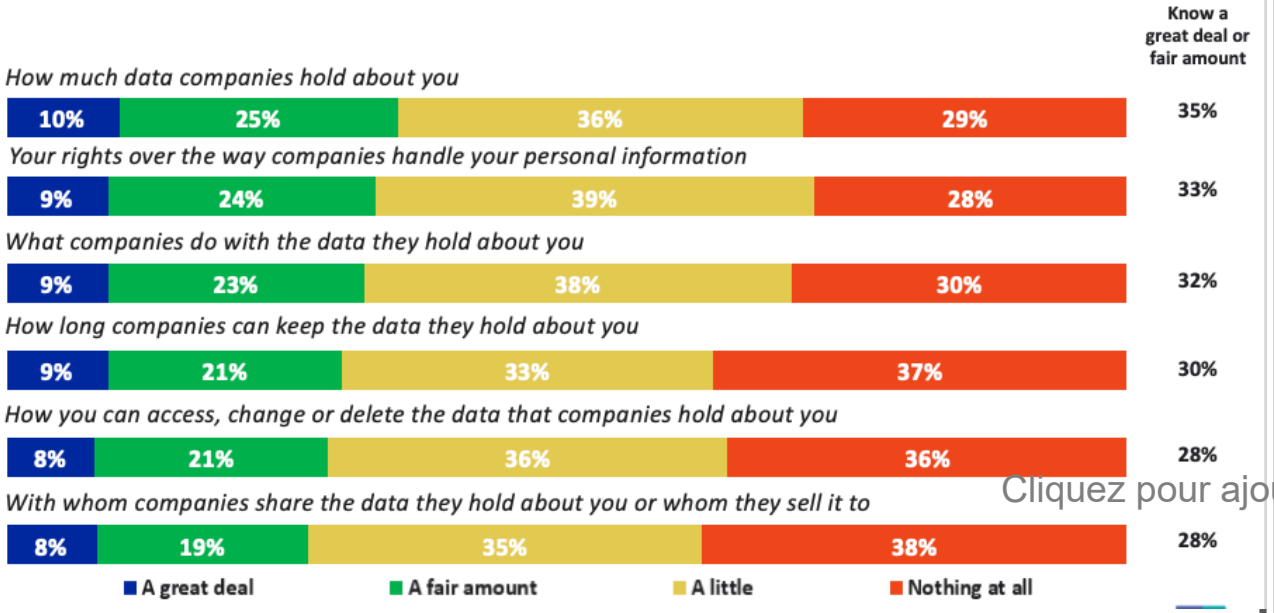
Promoting empathy and service to community : The Korea Power exchange.

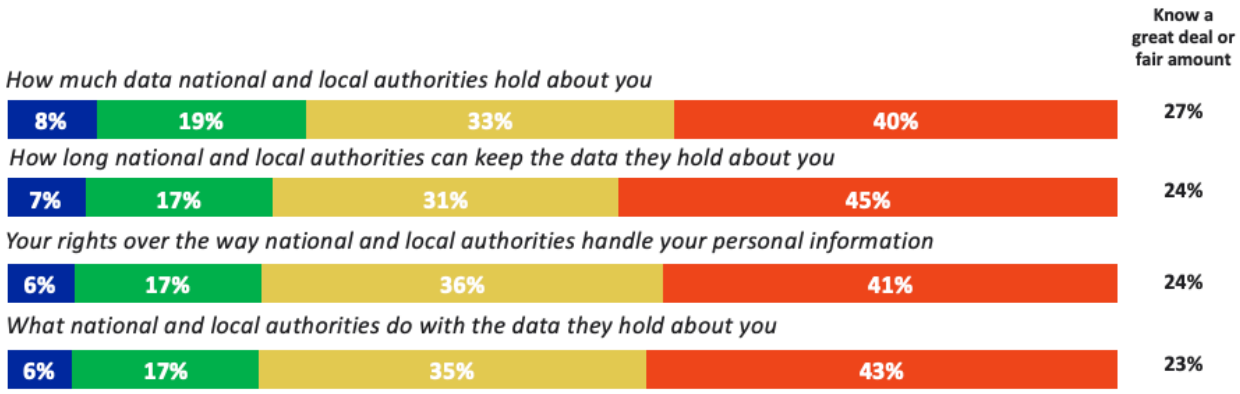
Keeping pace with tech change – Agile policymaking

=> Need for methodologies to **centralize best-practices, share information, and scale up.**

Deep dive: Fundamental rights, privacy and individual freedom in the digital era



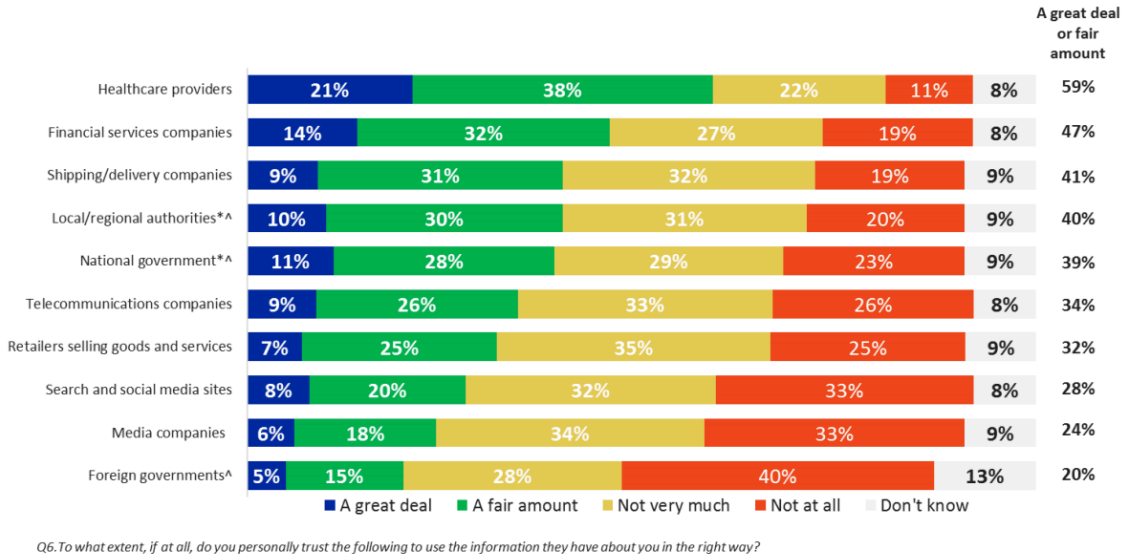




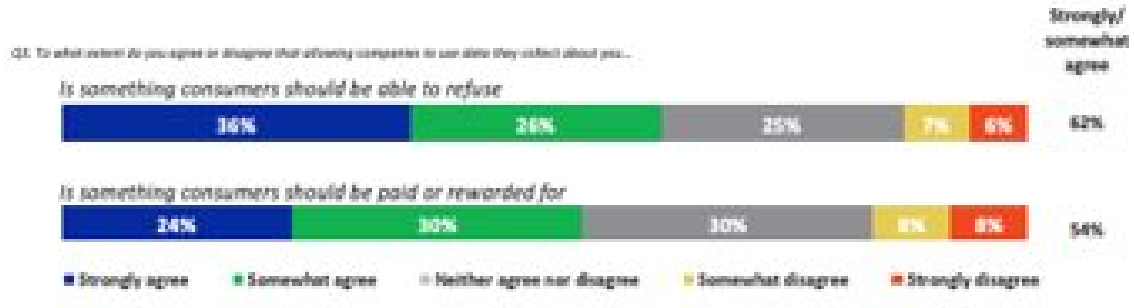
Cliquez pour ajouter du texte

Lack of knowledge...

Deep dive: Fundamental rights, privacy and individual freedom in the digital era



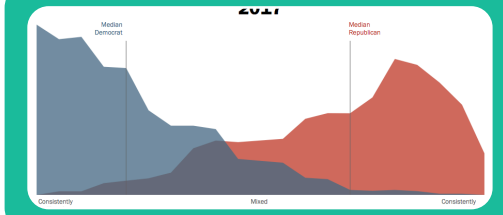
Lack of Trust



Control vs. Commoditization : privacy paradox

Deep dive: Fundamental rights, privacy and individual freedom in the digital era

Trust and public debate in the disinformation age



Overcoming Polarization

- Understand the "exhausted majority"
- Pol.is
- Media outlet AllSides

Empowering digital citizens

- Promote digital literacy & mental resilience :
  - Norway started early (2000s) with "The Knowledge Promotion" reform – digital literacy becoming the 5th basic competence along with writing, reading, arthmetic and oral skills

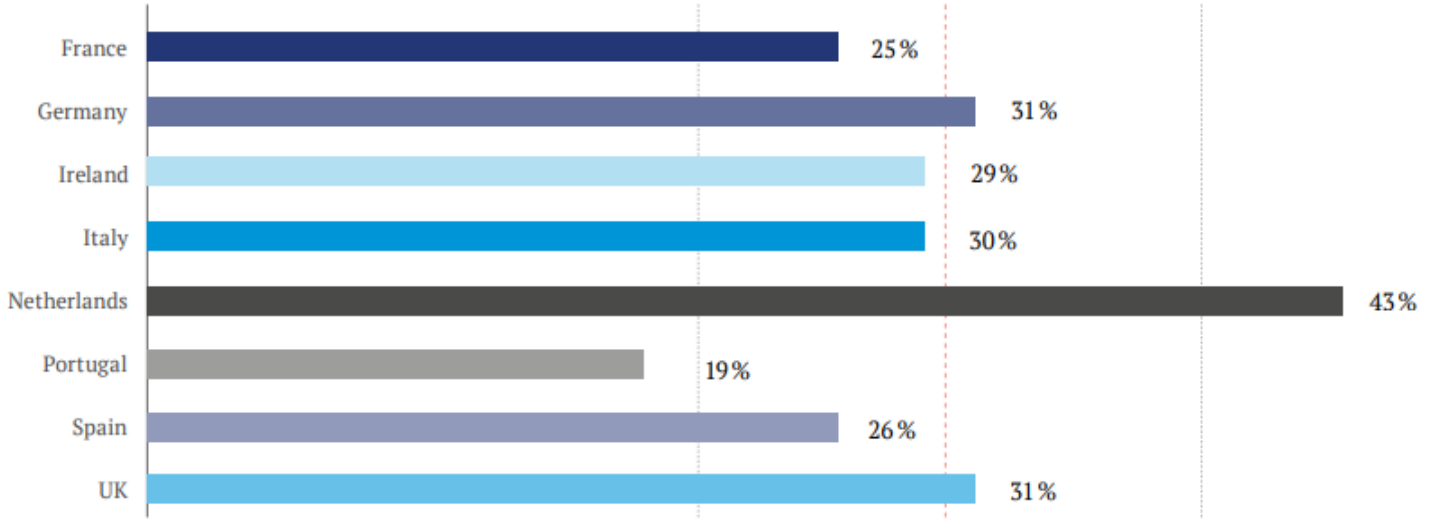
Promoting responsibility & accountability

- Of elites and citizens both
- USA Facts initiative

Hybrid Offline/online community movements

- a series of intimate community conversations Living Room Conversations (offline and online)
- and MisMatch, a political “dating app” that matches you with people on the opposite side for a polite video conversation.

AI Enthusiasm or establishment distrust?



One in every four Europeans would allow an artificial intelligence system to make important decisions about the running of their country. In nations such as the Netherlands, Germany, and the United Kingdom, the percentage is even higher – one in every three. This mindset, which probably relates to the growing mistrust citizens feel towards governments and politicians, underlies a questioning of the European model of representative democracy.

Source: IE Center for Governance of Change, 7 EU countries, 3000 participants

Technology, AI and the democratic social contract

Challenges

Overuse, underuse, misuse

Gap: Proliferating, diverging national strategies vs. binding global frameworks

Integrating media literacy with educational frameworks

Punishing anti-social behaviors on the net, preserving fundamental rights and liberties

“AI made me do it”- Humans shoving responsibility on AI

The Deepfake Race mimics gigahacks

Solutions

French and German hate speech and anti-fake news laws examples

OpenAI: Machine-to-machine interactions + human judges

New platforms for public debate should be built up that reward consensus algorithmically and are publicly funded.

New coalitions should start producing public benefit technology stacks.

What can we do together?

Political Vision-setting towards Humanistic AI Solutions

- Adding political weight & wisdom to the work of AI Stewards

Setting a framework for Citizen communications

- From mass awareness to mass literacy: e.g. Cross-society AI Literacy, Finland 55k citizens pilot example

Path-building from AI Ethical principles to policymaking

- Teaching AI human values through social surveying inputs (e.g. OpenAI initative → social surveying)

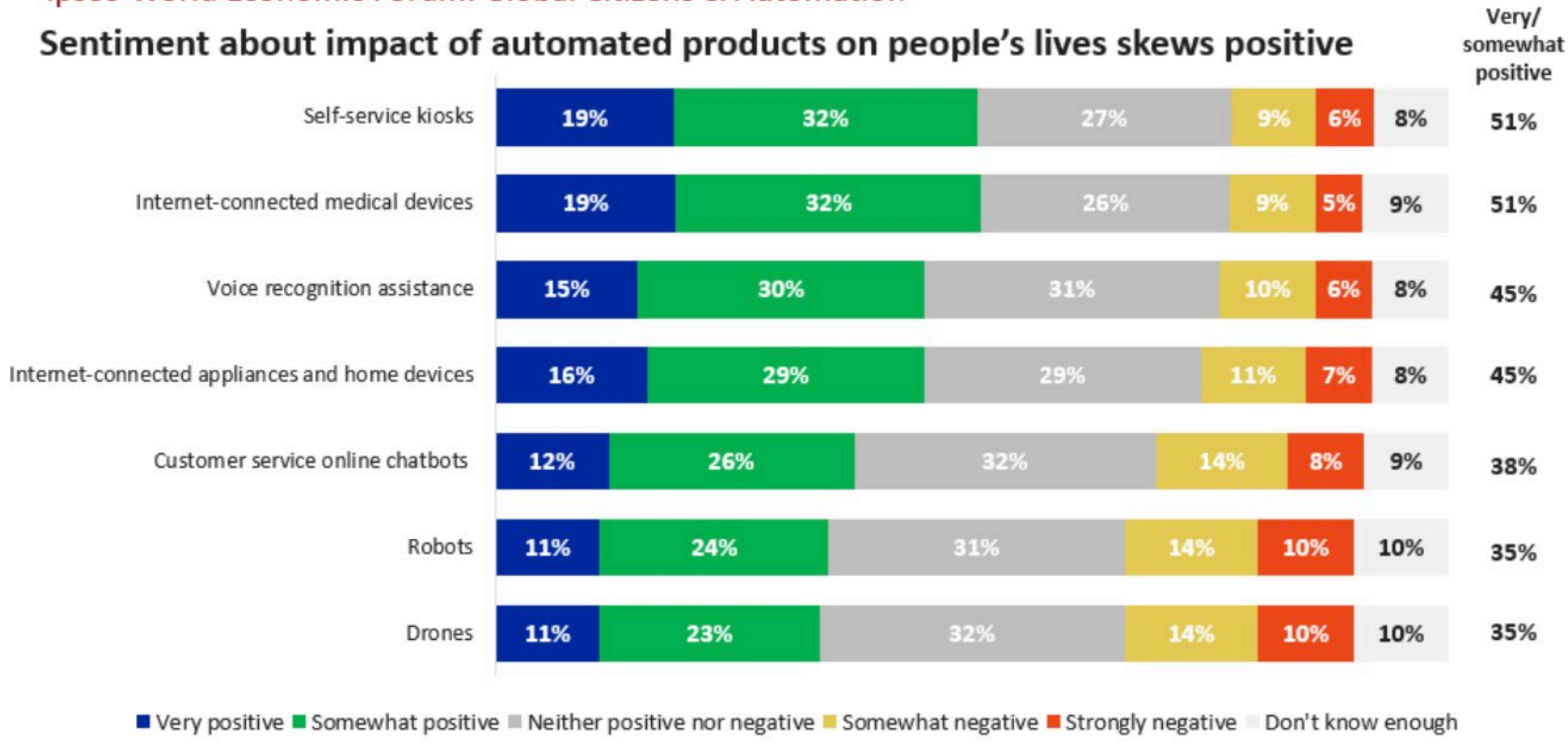
Joint AI Barometer

- Measuring societal AI perceptions globally and annually; identify gaps, ways to change behaviors and empower citizens

“Data & Wisdom” joint video content series

- Ipsos Insight & Wisdom by World Leaders: the best of both worlds

Sentiment about impact of automated products on people's lives skews positive



Q7.To what extent do you think that each of the following types of automated products and services has or will have a positive or a negative impact on people's lives?

Base: 18,813 adults aged 16-64 across Argentina, Australia, Belgium, Brazil, Canada, Chile, China, France, Germany, Great Britain, Hungary, India, Italy, Japan, Malaysia, Mexico, Peru, Poland, Russia, Saudi Arabia, South Africa, South Korea, Spain, Sweden, Turkey, United States. October-November 2018

Globalaffairs@Ipsos.com

Thank you!