



INSTITUTE OF DIRECTORS
IN IRELAND



AI: Powering Change in Our Time

With Clare Dillon, Clare Dillon Consulting

Breakfast and Evening Briefings Series: Kindly sponsored by Mazars



@IoDIreland

@MazarsIreland

#IoDevents

AI Ethics

A path towards
trustworthy AI

Clare Dillon

www.linkedin.com/in/claredillon

clare.dillon@live.ie



The 4th Industrial Revolution

Physical

Autonomous vehicles, robotics,
3D printing, new materials



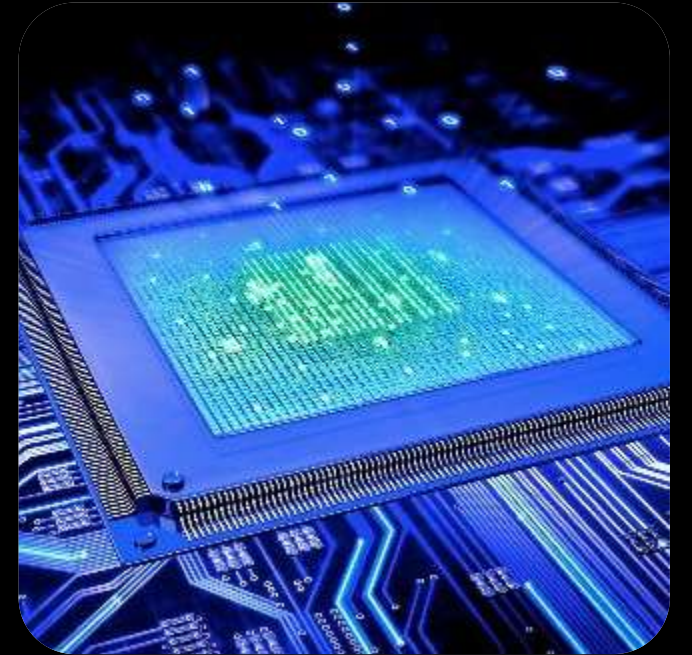
Biological

Genomic diagnostics,
treatment, engineering



Digital

IoT, Blockchain,
Artificial Intelligence (AI)



“Artificial Intelligence heralds **dramatic potential for growth** for both the economy and for humans.” - *Mark Purdy, Accenture*

“Huge value is going to come from the **new goods, services and innovations** AI will enable.” - *David Autor, Professor of Economics, MIT*

“By 2020, Artificial Intelligence Will **Create More Jobs Than It Eliminates**; AI Will Create 2.3 Million Jobs in 2020, While Eliminating 1.8 Million” - *Gartner*

AI Driving Economic Growth

- 1 | AI will potentially add **\$13 tr** (16%) to current global economic output by 2030.
- 2 | Spend estimated to grow to **\$77.6 bn** by 2022 (\$24 bn for 2018).
- 3 | By 2021, **75%** enterprise applications will use AI.
- 4 | US leads, (**60%** of spend) followed by Europe (retail/banking) and China (industry/gov).

What's Changing?



Automation

(Increased Efficiency)



Augmenting Capabilities

(Increased Productivity)



New Capabilities

(Disruptive Innovation)



New Industries

(Radical Innovation)



Automation
(Increased Efficiency)

What's
Changing?

Examples

Automation:

- Supply Chain Automation: Ocaco
- Robot Dispensing Drugs from CarePlus Pharmacies
- Robotic Process Automation: Blue Prism & NHS Case Study

Augmented Capabilities:

- Customer Service Chatbots: Julie, the Amtrak Chatbot
- Ellie: the AI Therapist “doing a better job than humans”
- Textio: AI improving inclusivity

New Capabilities:

- IDX-DR: FDA Approved Diagnostic Tool

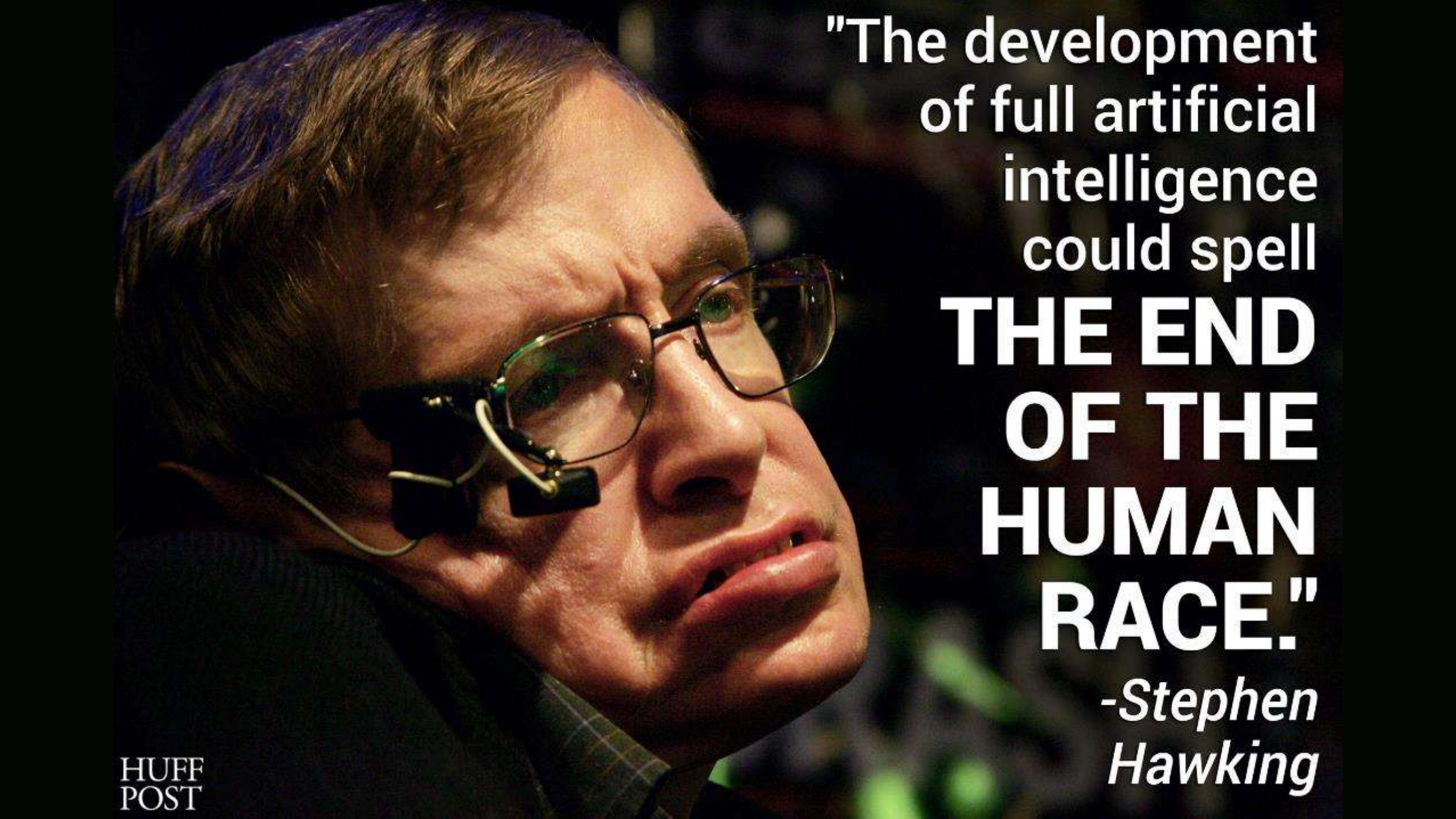
New Industries:

- Autonomous Vehicles
- Jewelry by Eva Nowak: shield against facial recognition



AI has huge
potential to
benefit the
human race.



A close-up, profile view of Stephen Hawking looking upwards and to the right. He is wearing his signature glasses and a small, black, rectangular facial tracking device attached to the side of his face with a thin wire. The background is dark and out of focus.

"The development
of full artificial
intelligence
could spell
**THE END
OF THE
HUMAN
RACE."**

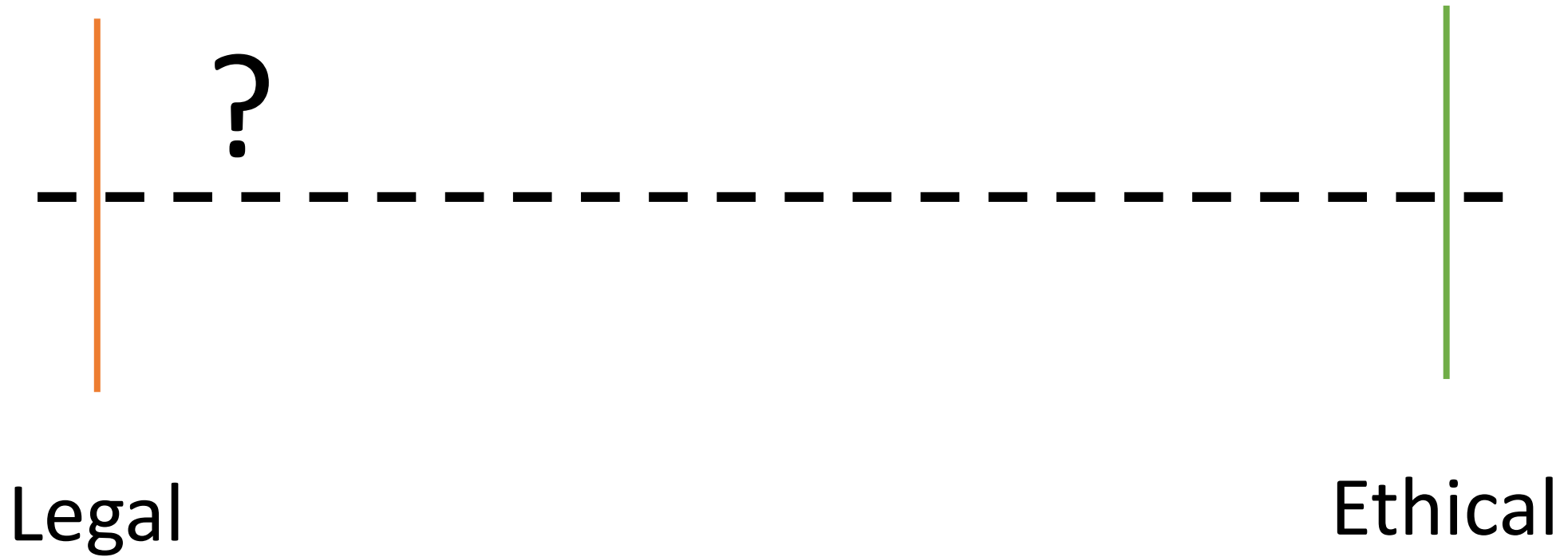
*-Stephen
Hawking*

Building AI ART

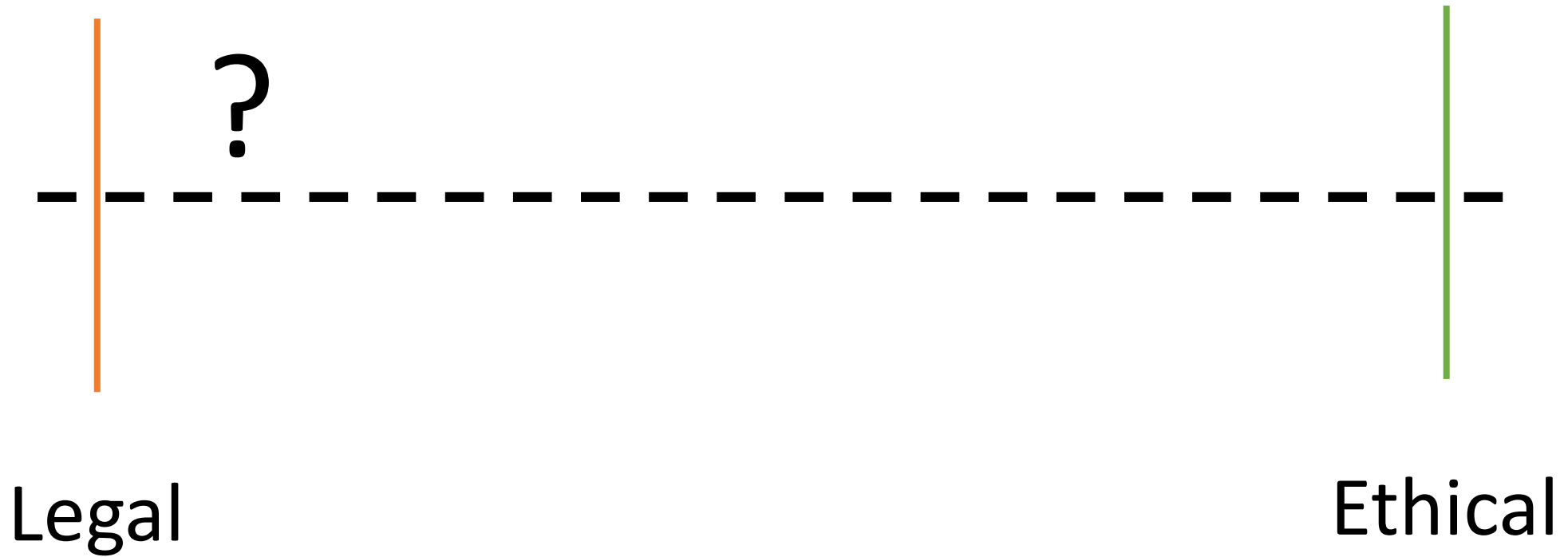
Accountability
Responsibility
Transparency



The Ethical Continuum



The Ethical Continuum

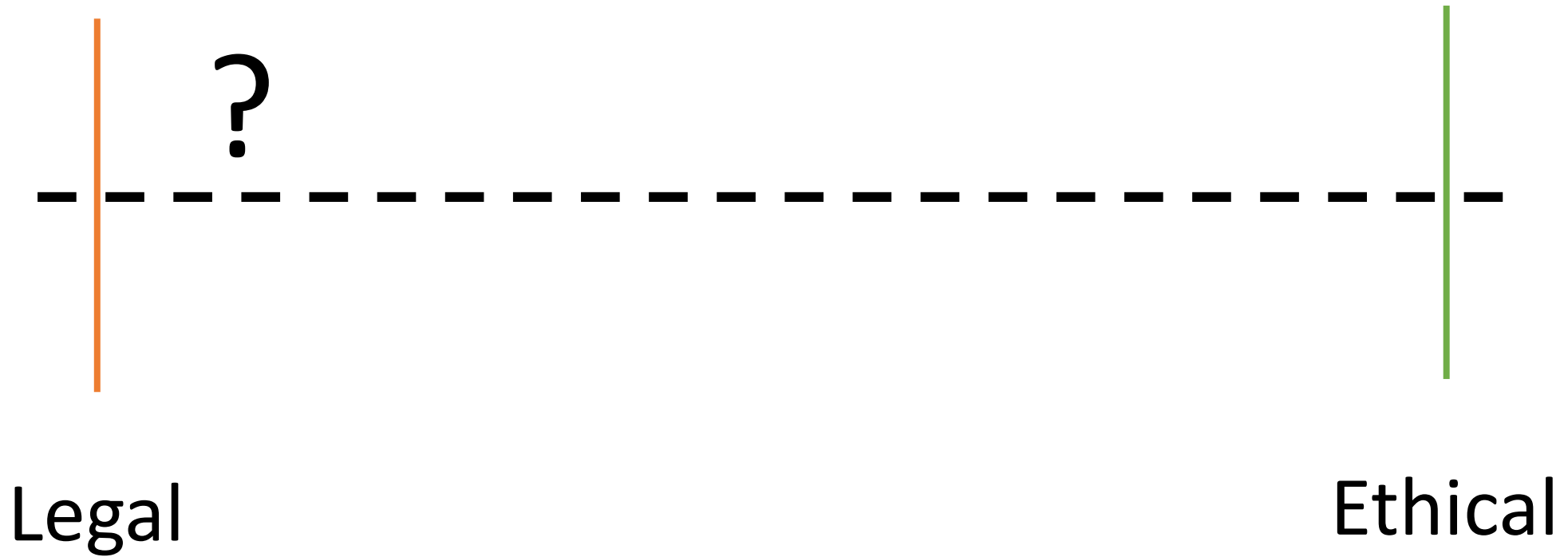


AI & GDPR

Individuals “shall have the right **not to be subject to a decision based solely on automated processing, including profiling**, which produces legal effects concerning him or her or **similarly significantly affects** him or her”.

Data controllers must notify consumers how their data will be used, including “the **existence of automated decision-making**, and, at least in those cases, **meaningful information about the logic involved**, as well as **the significance and the envisaged consequences of such processing** for the data subject.”

The Ethical Continuum



Trustworthy AI Requirements



Source: European Commission Ethics Guidelines for Trustworthy AI

When AI
Goes Bad...





A silver laptop is shown from a front-facing perspective, open. The screen is white and displays the text "When AI goes bad...." in a black, sans-serif font. The laptop has a black bezel around the screen and a silver body. The keyboard and trackpad are visible below the screen. The laptop is set against a plain white background.

When AI
goes bad....



... How can I help you? 😊

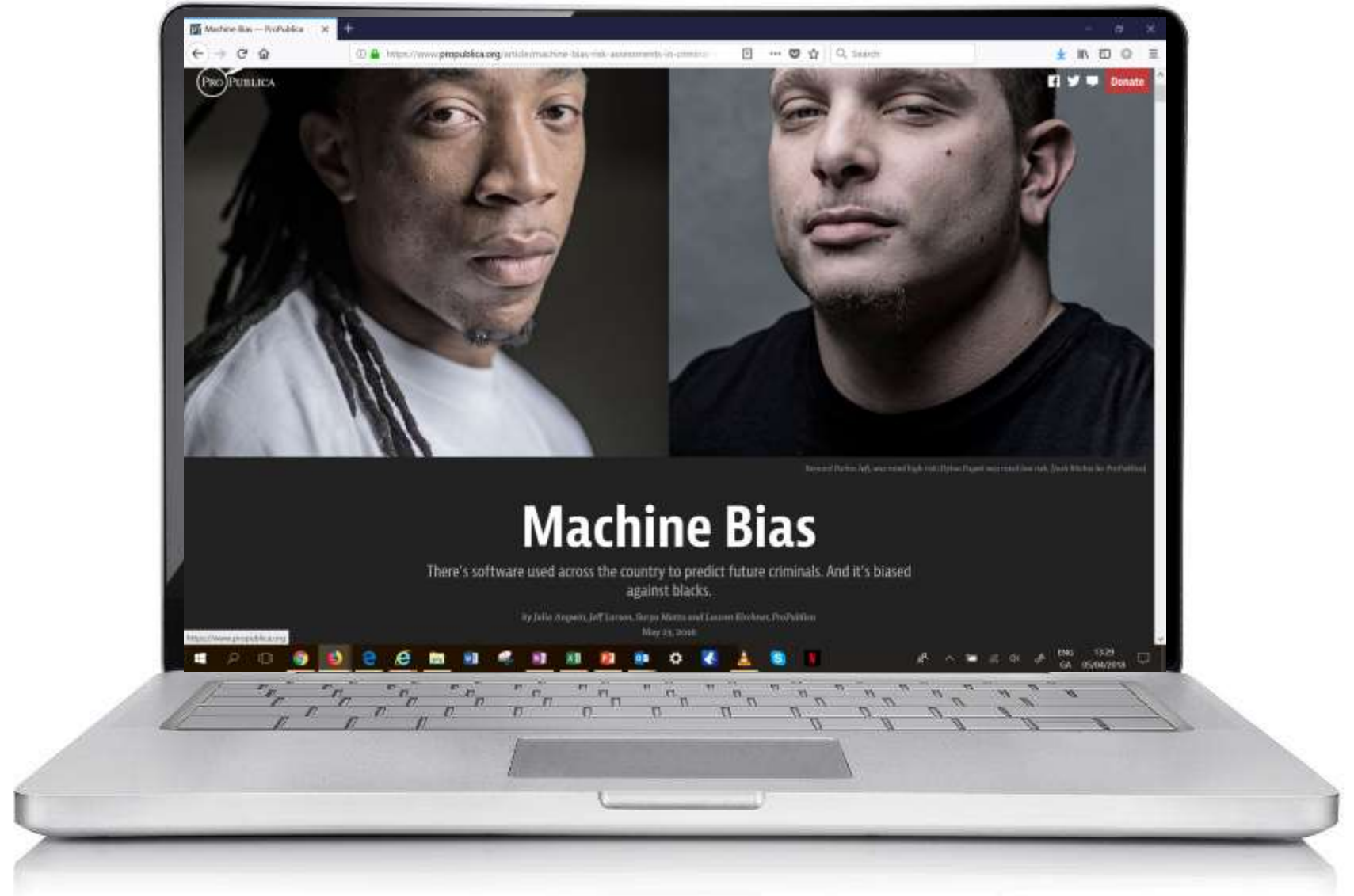
AI is not
right for
every
scenario!



Unintended
uses of AI!



Historical
Datasets
can be
biased!



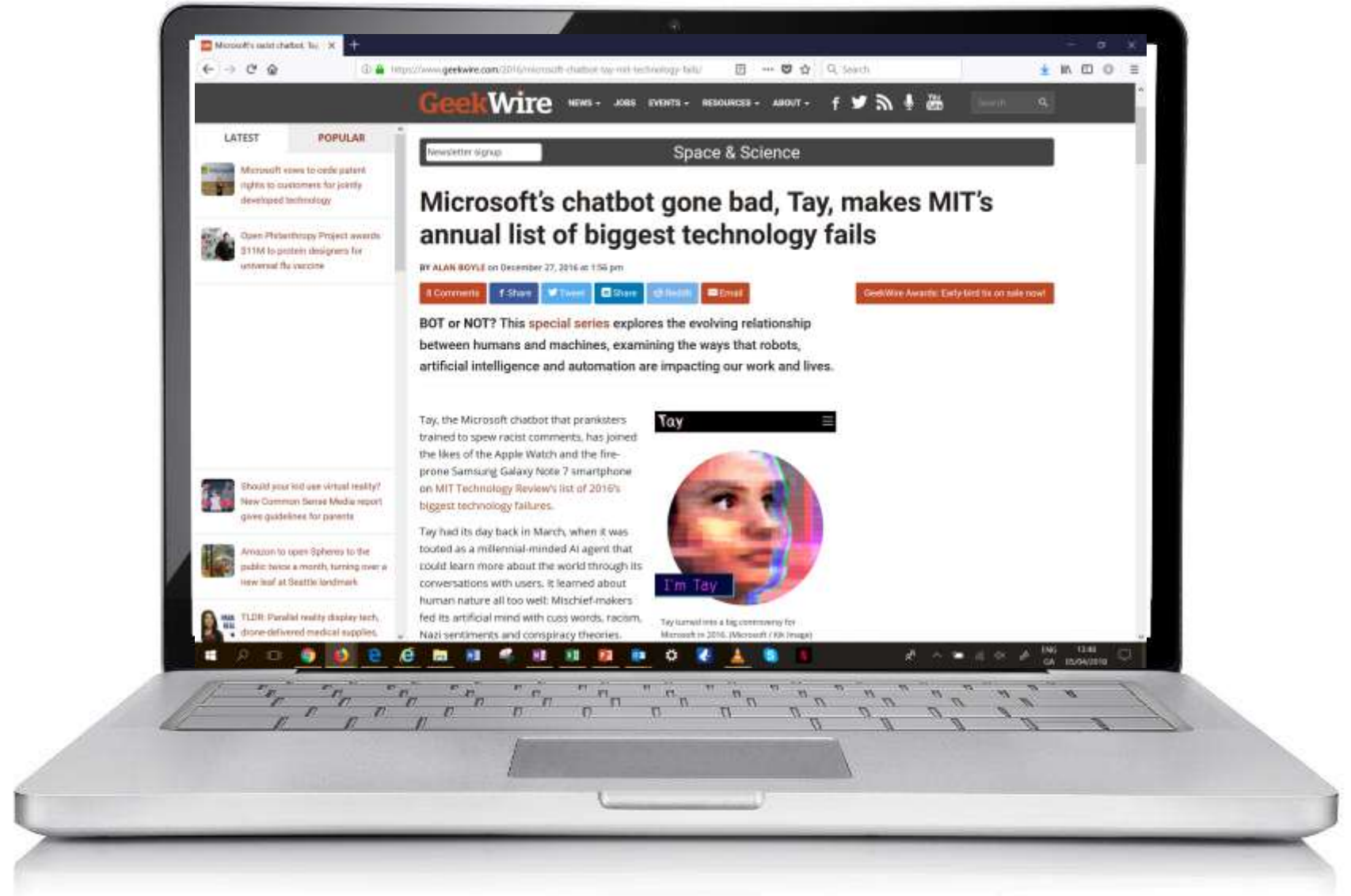
Training
datasets
may be
incomplete!



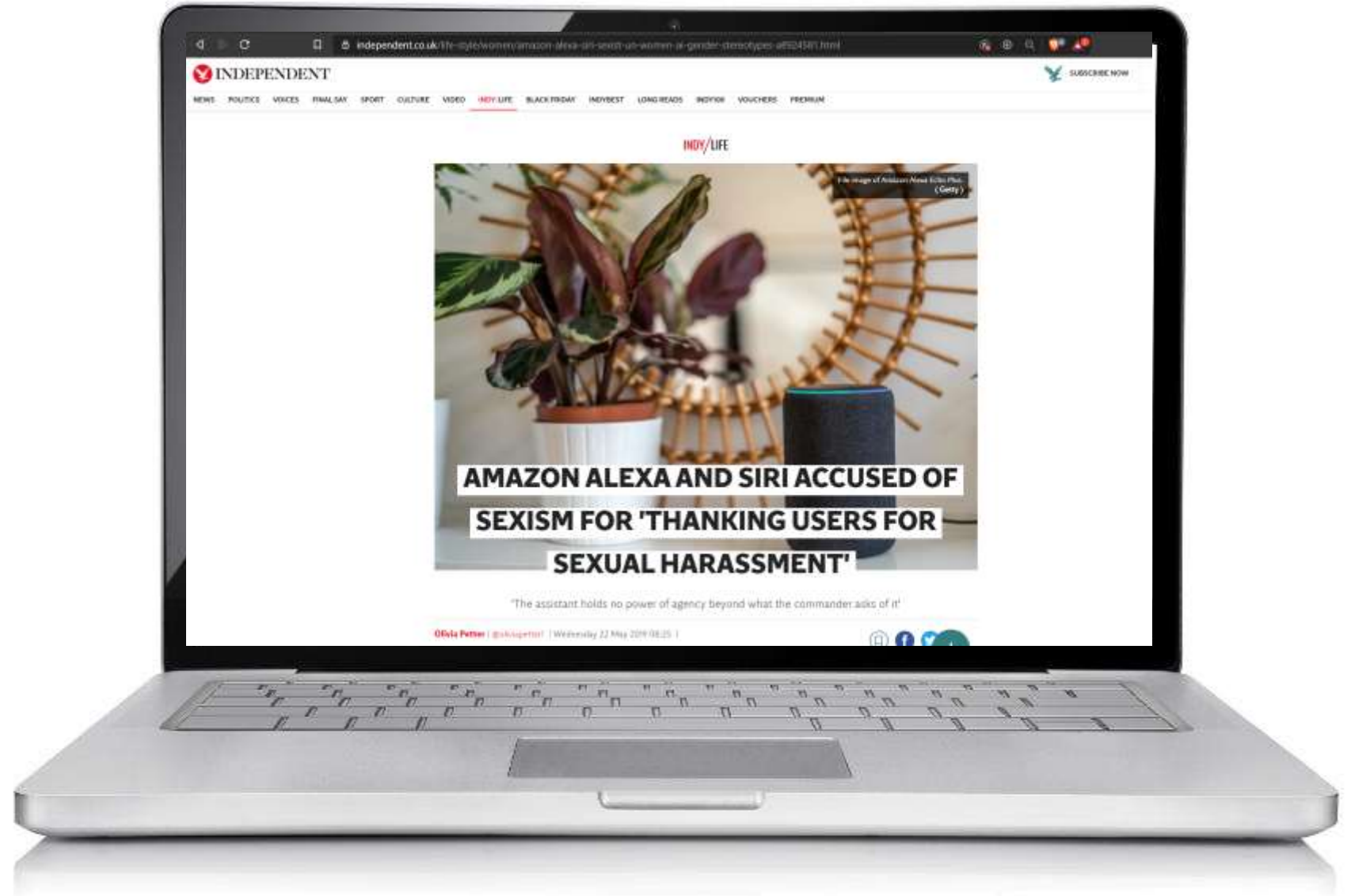
Harmful
even when
it works!



Failing to
prevent
malicious
use!



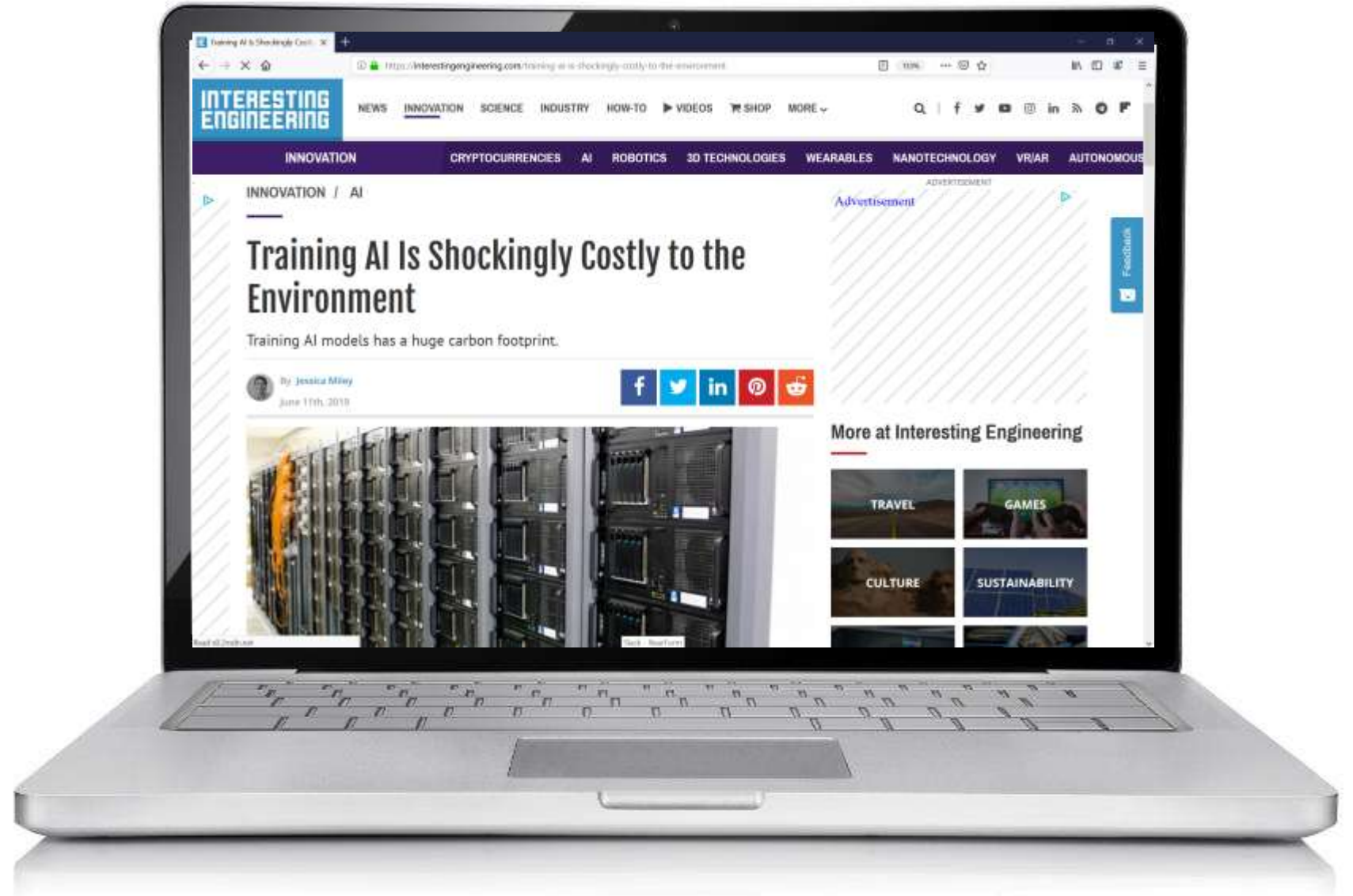
Encouraging Offensive Behaviours!



AI
deception!



Impact on
Environment!



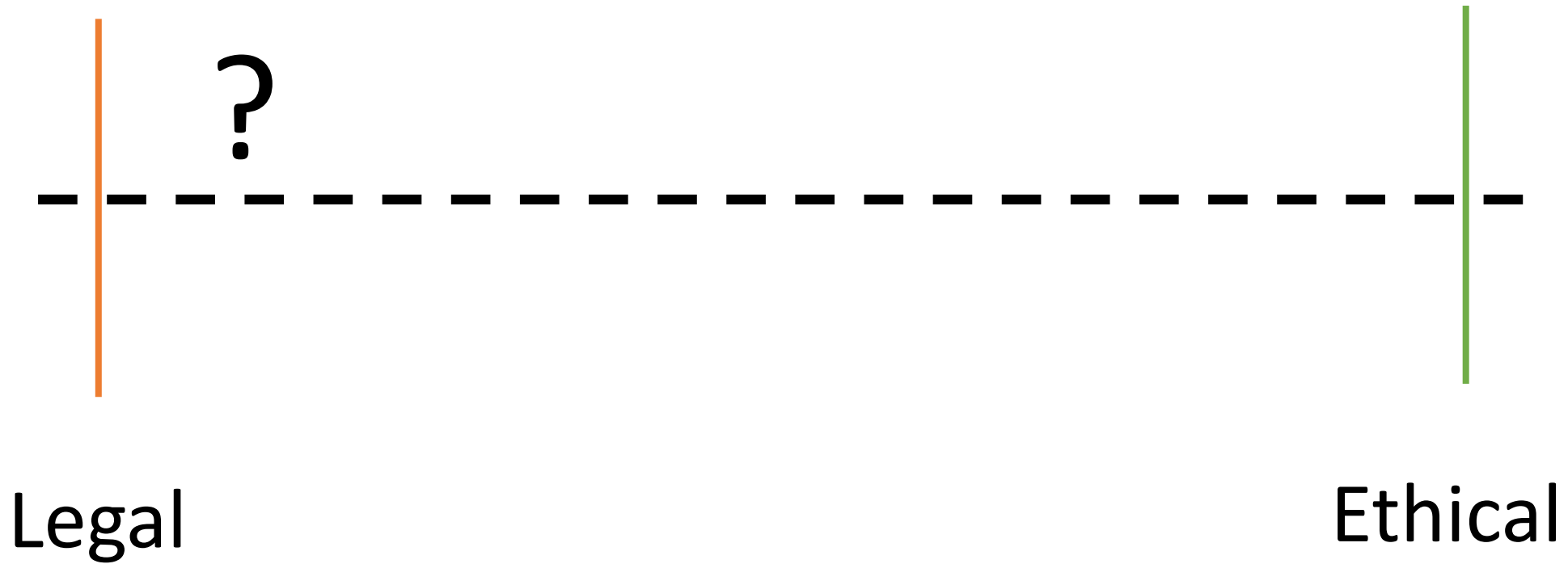
Tips for Ethical AI...



Tips for Planning!

- 1 | Decide where you are on the Ethical Continuum.

The Ethical Continuum












Tips for Planning!

- 1 | Decide where you are on the Ethical Continuum.
- 2 | Connect AI implementation to valid business case.

Ethics Canvas

www.ethicscanvas.org

Ethics Canvas v1.8 - ethicscanvas.org © ADAPT Centre & Trinity College Dublin & Dublin City University, 2017.

<p>Individuals affected</p> <p>Identify the types or categories of individuals affected by the product or service, such as men/women, user/non-user, age-category, etc.</p> <p> 1</p>	<p>Behaviour</p> <p>Discuss problematic changes to individual behaviour that may be prompted by the application e.g. differences in habits, time-schedules, choice of activities, people behaving more individualistic or collectivist, people behaving more or less materialistic.</p> <p> 3</p> <p>Relations</p> <p>Discuss problematic differences in individual behaviour such as differences in habits, time-schedules, choice of activities, etc.</p> <p> 4</p>	<p>What can we do?</p> <p>Select the four most important Ethical impacts you discussed. Identify ways of solving these impacts by changing your project's product/service design, organisation. Or by providing recommendations for its use or spelling out more clearly to users the values driving the design.</p> <p> 9</p>	<p>Worldviews</p> <p>Discuss how the general perception of somebody's role in society can be affected by the project.</p> <p> 5</p> <p>Group Conflicts</p> <p>Discuss the impact on the relationships between the groups identified, e.g. employers and unions.</p> <p> 6</p>	<p>Groups affected</p> <p>Identify the collectives or communities, e.g. groups or organisations, that can be affected by your product or service, such as environmental and religious groups, unions, professional bodies, competing companies and government agencies, considering any interest they might have in the effects of the product or service.</p> <p> 2</p>
<p>Product or Service Failure</p> <p>Discuss the potential negative impact of your product or service failing to operate as intended, e.g. technical or human error, financial failure/ receivership/acquisition, security breach, data loss, etc.</p> <p> 7</p>		<p>Problematic Use of Resources</p> <p>Discuss possible negative impacts of the consumption of resources of your project, e.g. climate impacts, privacy impacts, employment impacts etc.</p> <p> 8</p>		

Assessment List for Trustworthy AI

TRUSTWORTHY AI ASSESSMENT LIST (PILOT VERSION)

1. Human agency and oversight

Fundamental rights:

- ✓ Did you carry out a fundamental rights impact assessment where there could be a negative impact on fundamental rights? Did you identify and document potential trade-offs made between the different principles and rights?
- ✓ Does the AI system interact with decisions by human (end) users (e.g. recommended actions or decisions to take, presenting of options)?
 - Could the AI system affect human autonomy by interfering with the (end) user's decision-making process in an unintended way?
 - Did you consider whether the AI system should communicate to (end) users that a decision, content, advice or outcome is the result of an algorithmic decision?
 - In case of a chat bot or other conversational system, are the human end users made aware that they are interacting with a non-human agent?

Human agency:

- ✓ Is the AI system implemented in work and labour process? If so, did you consider the task allocation between the AI system and humans for meaningful interactions and appropriate human oversight and control?
 - Does the AI system enhance or augment human capabilities?
 - Did you take safeguards to prevent overconfidence in or overreliance on the AI system for work processes?

Human oversight:

- ✓ Did you consider the appropriate level of human control for the particular AI system and use case?
 - Can you describe the level of human control or involvement?
 - Who is the "human in control" and what are the moments or tools for human intervention?
 - Did you put in place mechanisms and measures to ensure human control or oversight?
 - Did you take any measures to enable audit and to remedy issues related to governing AI autonomy?

Tips for Planning!

- 1 | Decide where you are on the Ethical Continuum.
- 2 | Connect AI implementation to valid business case.
- 3 | Determine measures of success and failure.
- 4 | Determine the need for Open or Explainable AI (XAI)

Tips for Getting Started!

- 5 | Hire a diverse team.
- 6 | Educate, educate, educate.
- 7 | Build a risk mitigation plan.

Tips for Building!

8 | Track datasets.

9 | Test, test, test.

Tips for After Launch!

10 | Keep testing.

11 | Monitor Usage Scenarios.

12 | Be transparent.

Let's Build AI ART!

Clare Dillon

www.linkedin.com/in/claredillon

clare.dillon@live.ie

