

# Info Sec, AI, and Ethics

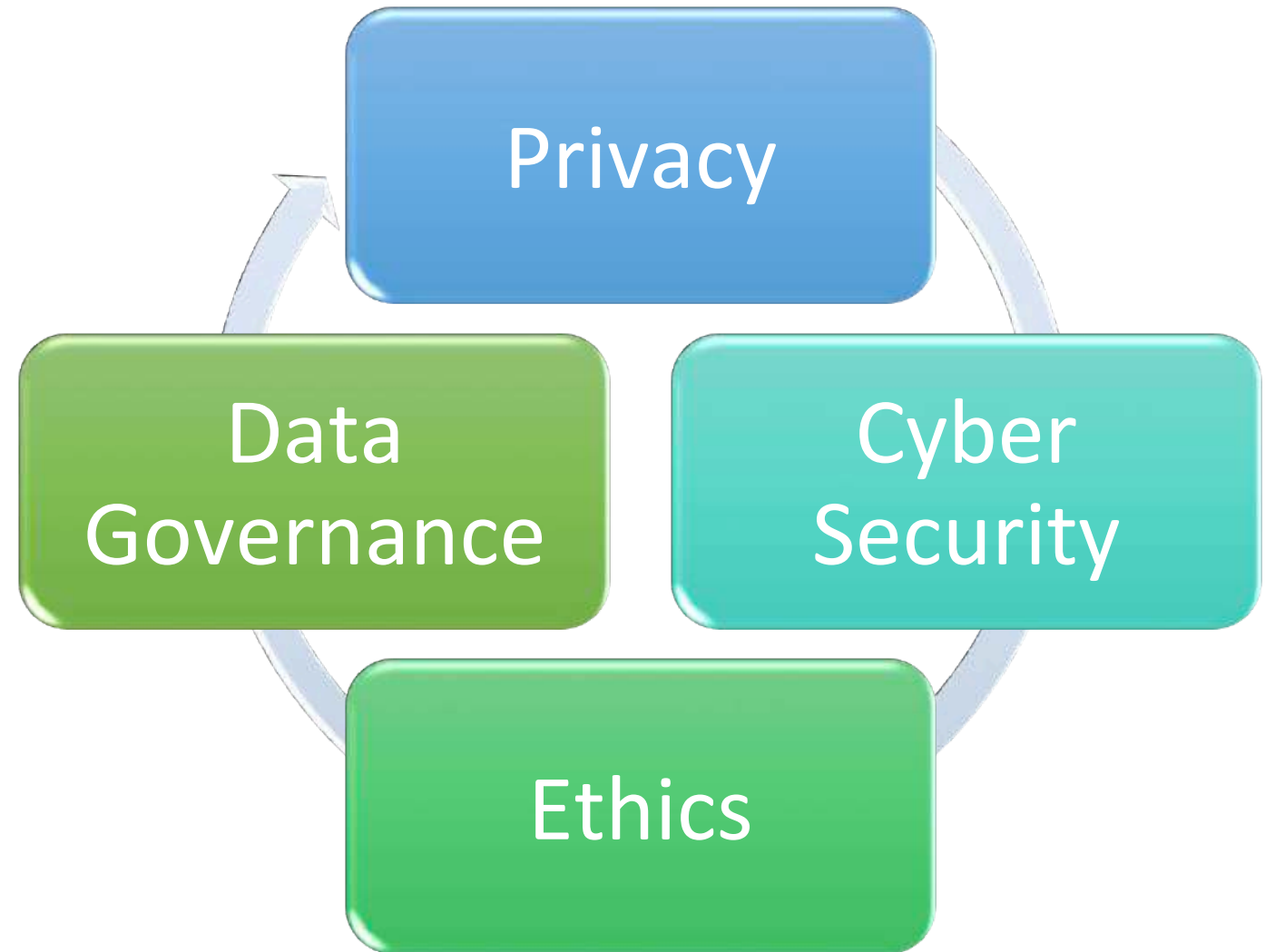
Slides are up on SlideShare:  
<https://www.slideshare.net/carruthk>

# Agenda

1. The world is changing
2. Unforeseen consequences
3. Why ethics & diversity are important
4. What can be done?

How I got here...

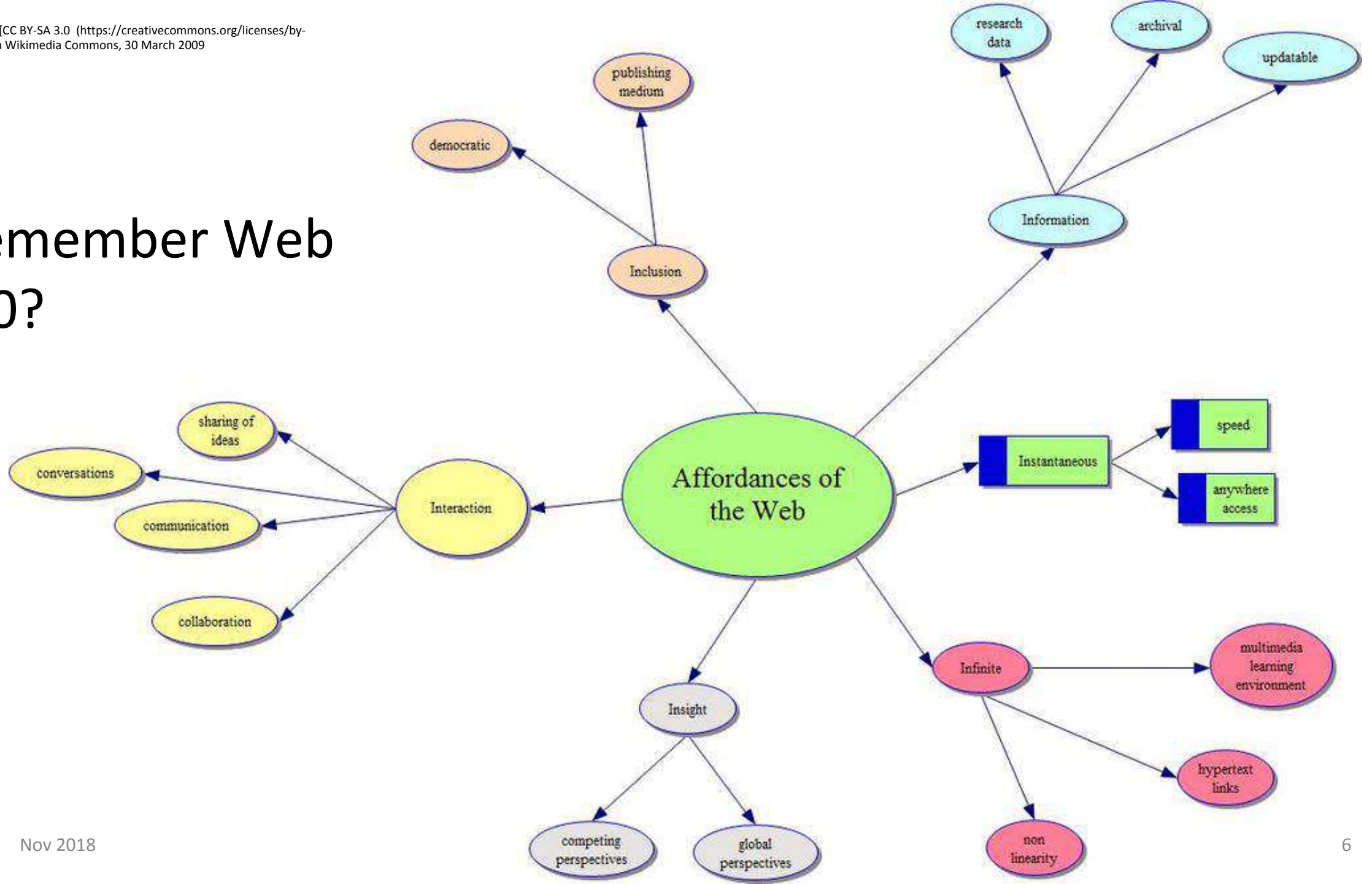
Corporate IT →





The world is  
changing

# Remember Web 2.0?





The world's most valuable resource is no longer oil, but data

David Parkins

Source: <https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource>

# Because software is eating the world...

Marc Andreessen, Why Software Is Eating The World, 20 August 2011, Wall Street Journal,  
<https://www.wsj.com/articles/SB10001424053111903480904576512250915629460>



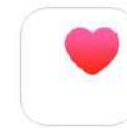


What can I help you with?

Hi, I'm Cortana.



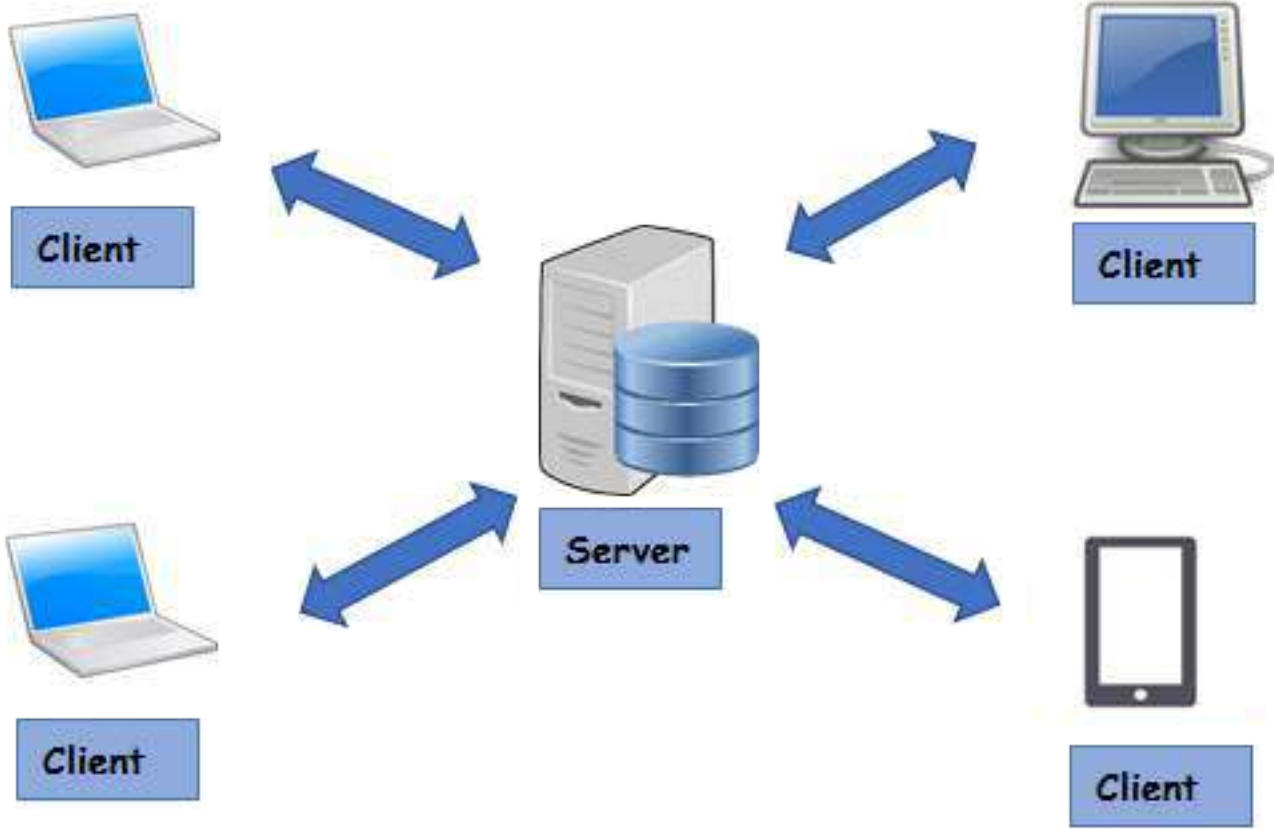
# And hardware is helping



Health.  
An entirely new way to use your  
health and fitness information.



# The old world



Source: <http://toolsqa.com/client-server/client-server-architecture-and-http-protocol/>



RabbitMQ



docker

 serverless



Amazon EC2



redis



AWS Lambda



AWS SNS



Amazon RDS



HashiCorp  
**Terraform**



Amazon RDS



AWS IoT

Source: <https://read.acloud.guru/our-serverless-journey-part-2-908d76d03716>



Unforeseen  
consequences of  
connectedness



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

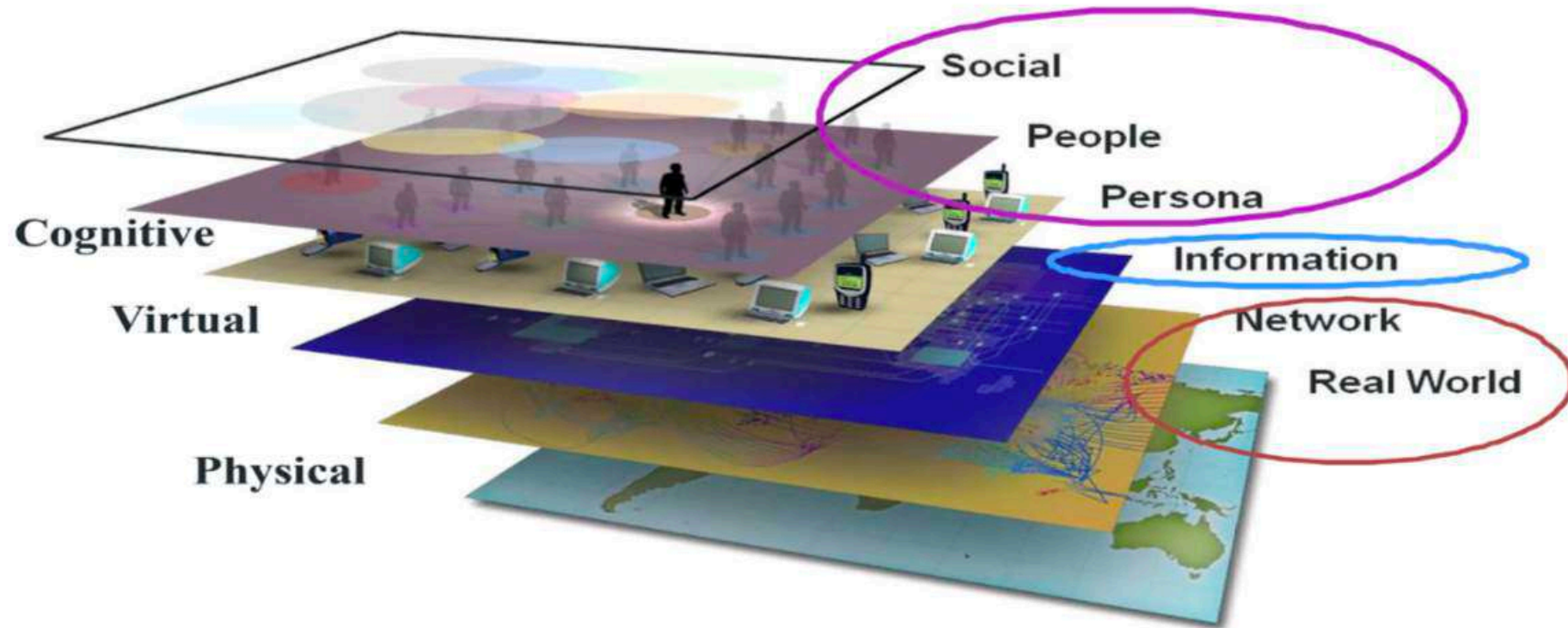
Australians know all about unintended consequences

# Weaponization of social media



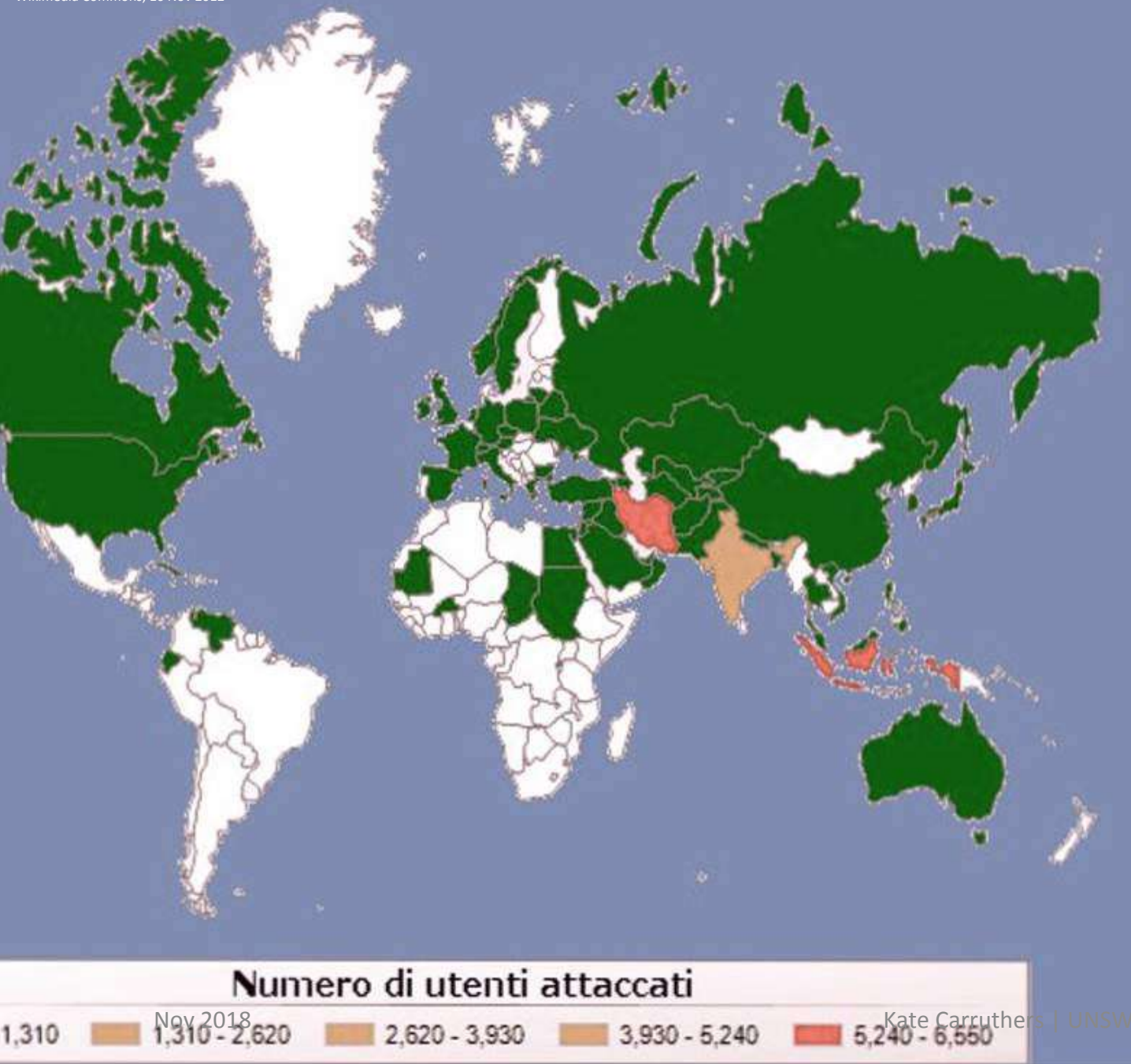
- “ The internet is changing war and politics
- War and politics are changing the internet
- Terrorists livestream their attacks
- “Twitter wars” produce real-world casualties
- Viral misinformation alters not just the result of battles, but the very fate of nations
- The result is that war, tech, and politics have blurred into a new kind of battlespace that plays out on our smartphones

Singer, Peter Warren, and Brooking, Emerson T.. *Likewar: the Weaponization of Social Media*. Eamon Dolan/Houghton Mifflin Harcourt, 2018.



*Figure 2.1. The information environment*<sup>36</sup>

Nissen, Thomas Elkjer. *The Weaponization Of Social Media: Characteristics of Contemporary Conflicts*. Royal Danish Defence College, 2015. <https://www.stratcomcoe.org/thomas-nissen-weaponization-social-media>



"This [Stuxnet] has the whiff of August 1945..."

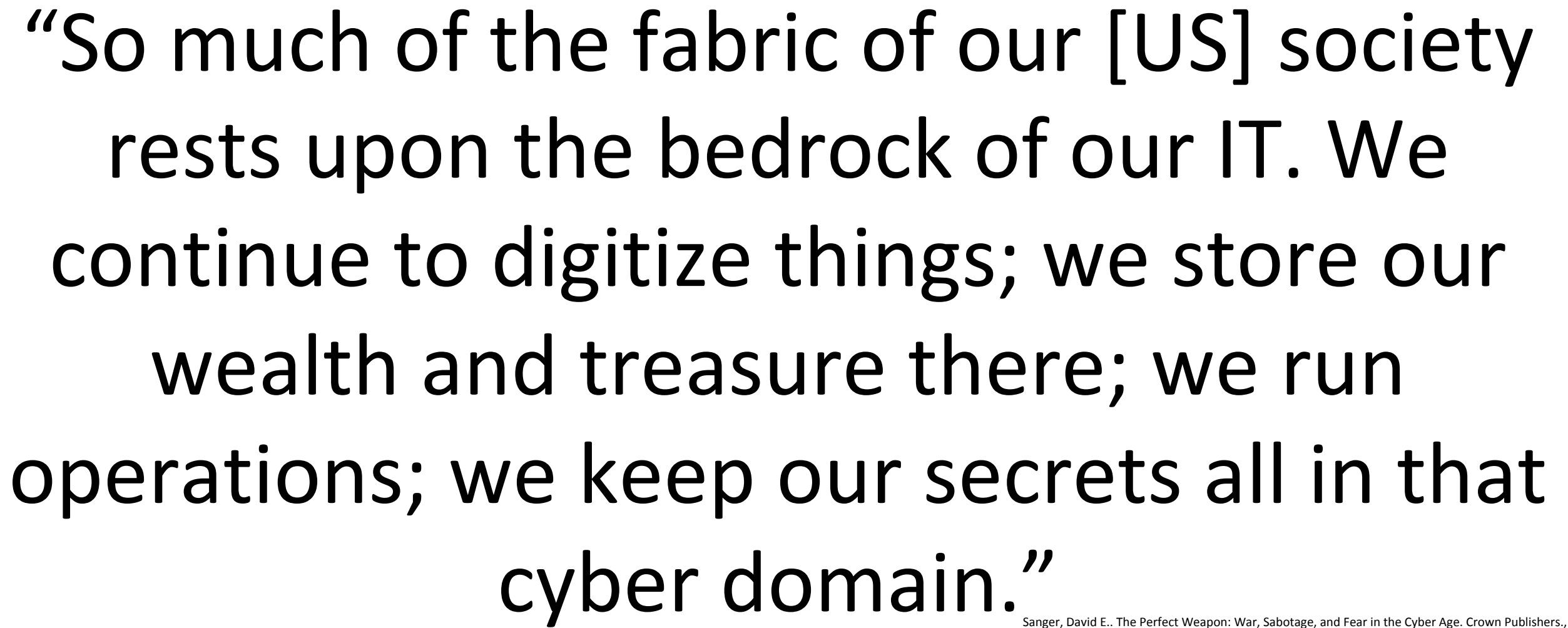
"Someone, probably a nation-state, just used a cyber weapon in a time of peace...to destroy what another nation could only describe as their critical infrastructure."

"That's a big deal. That's never happened before"

Paul D. Shinkman, Former CIA Director: Cyber Attack Game-Changers Comparable to Hiroshima, US News, 20 Feb 2013, <https://www.usnews.com/news/articles/2013/02/20/former-cia-director-cyber-attack-game-changers-comparable-to-hiroshima>

Gen. Michael Hayden, former NSA & CIA Director





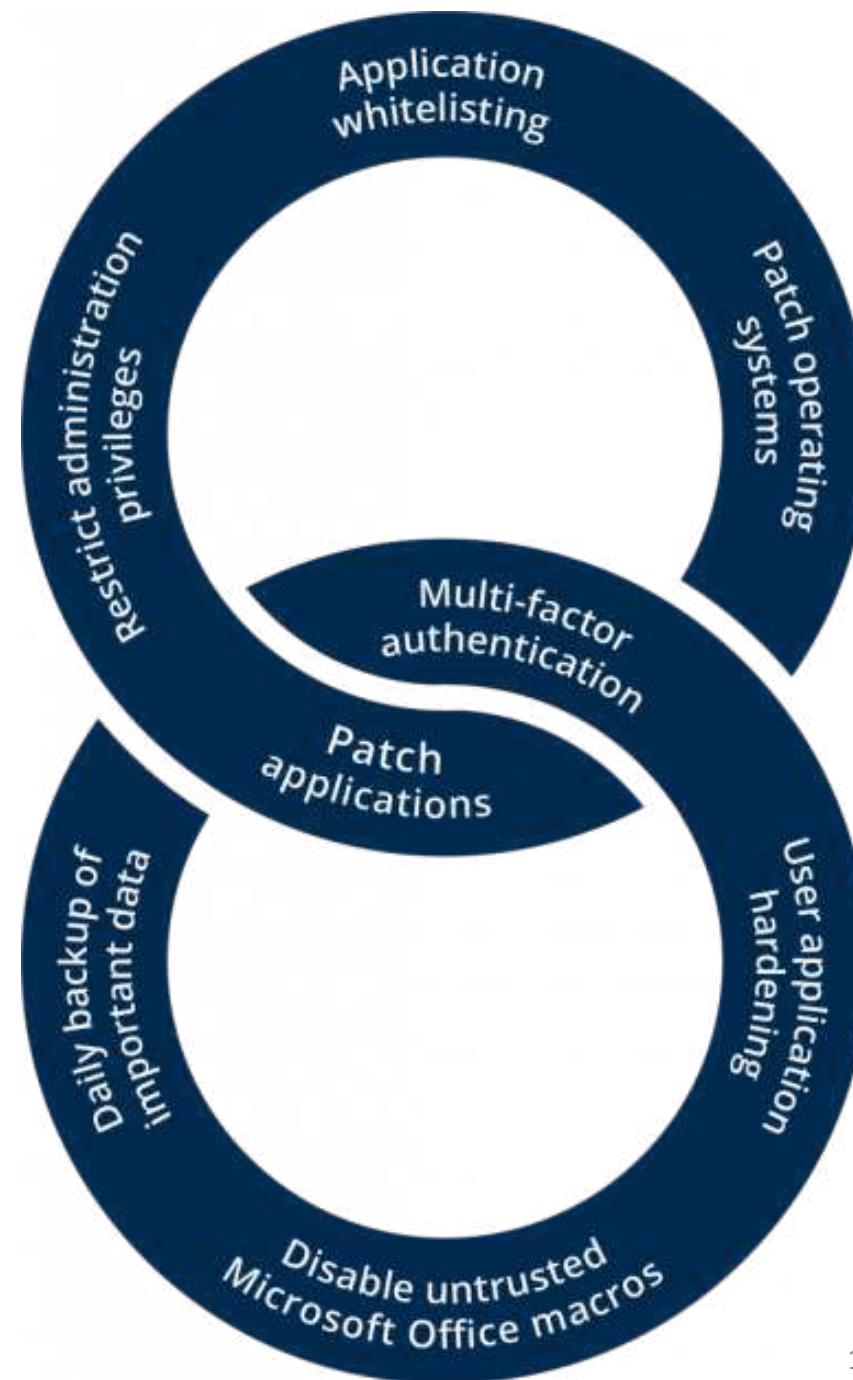
“So much of the fabric of our [US] society rests upon the bedrock of our IT. We continue to digitize things; we store our wealth and treasure there; we run operations; we keep our secrets all in that cyber domain.”

Sanger, David E.. The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age. Crown Publishers., 2018.

Cyber is the new  
battlefield

& everything is cyber now

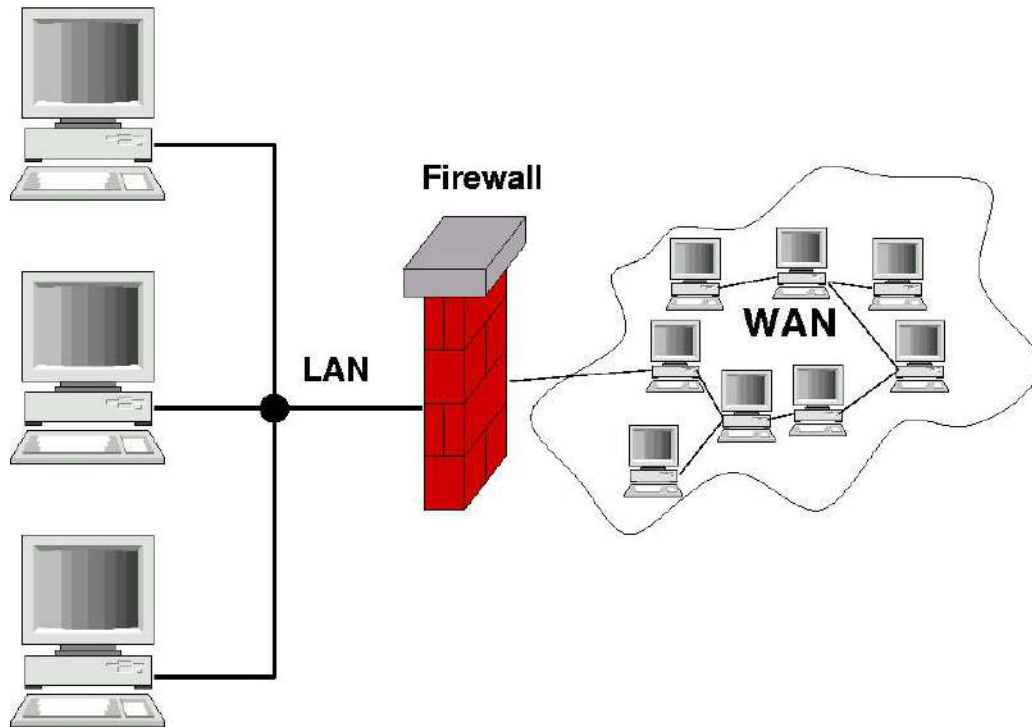
# Even the essential 8 will not protect us



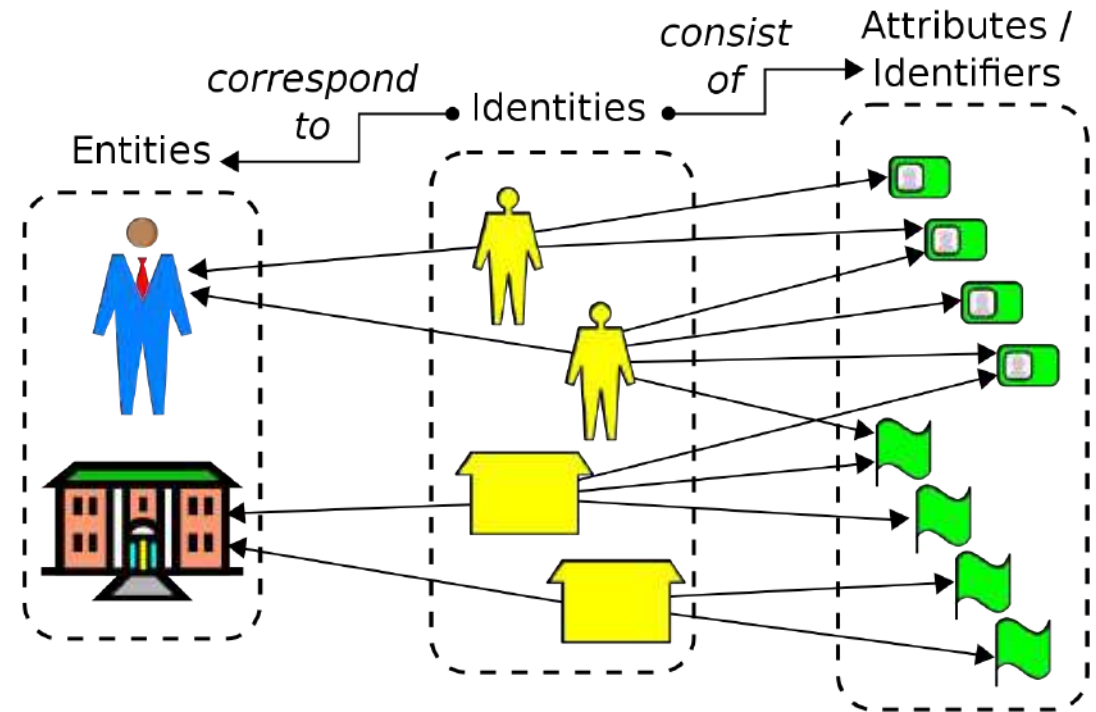
<https://www.huntsmansecurity.com/solutions/cyber-security-compliance/asd-essential-eight/>

# Boundary is no longer at your organisation's firewalls - Identity is the boundary

## Old world



## New world



By Harald Mühlböck [GFDL (<http://www.gnu.org/copyleft/fdl.html>) or CC-BY-SA-3.0 (<http://creativecommons.org/licenses/by-sa/3.0/>)], via Wikimedia Commons

By Audun Jøsang [CC BY 3.0 (<https://creativecommons.org/licenses/by/3.0/>)], via Wikimedia Commons

# Combine weaponized connectivity with capitalism



**Ewen Shearer**  
@SheepOverboard

On our destiny...

"Big Brother will not be a fascist paranoid kleptocracy, but a sociopathic AI-driven corporate panopticon, having wrested the reins from our grubby hands to impose the cold logic of coded imperatives upon our free will – eventually to discard us."

8:37 pm · 22 Oct 2018

Source: <https://sheepoverboard.com/trojans-cusp/>



**Justin Warren** ✓  
@jpwarren

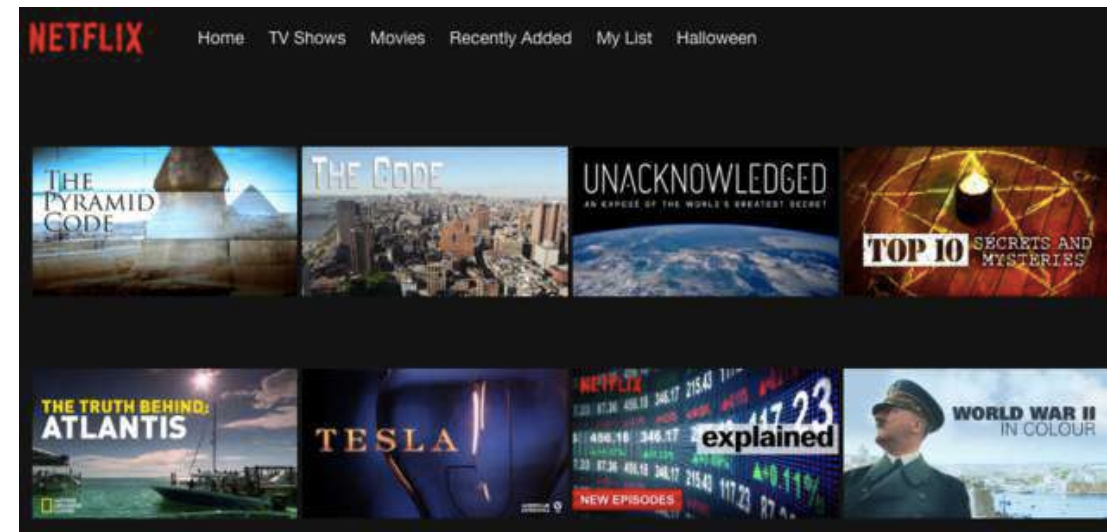
Replying to @paul\_shetler, and @kcarruthers

In the future, all capitalism is surveillance capitalism.

6:54 am · 26 Oct 2018

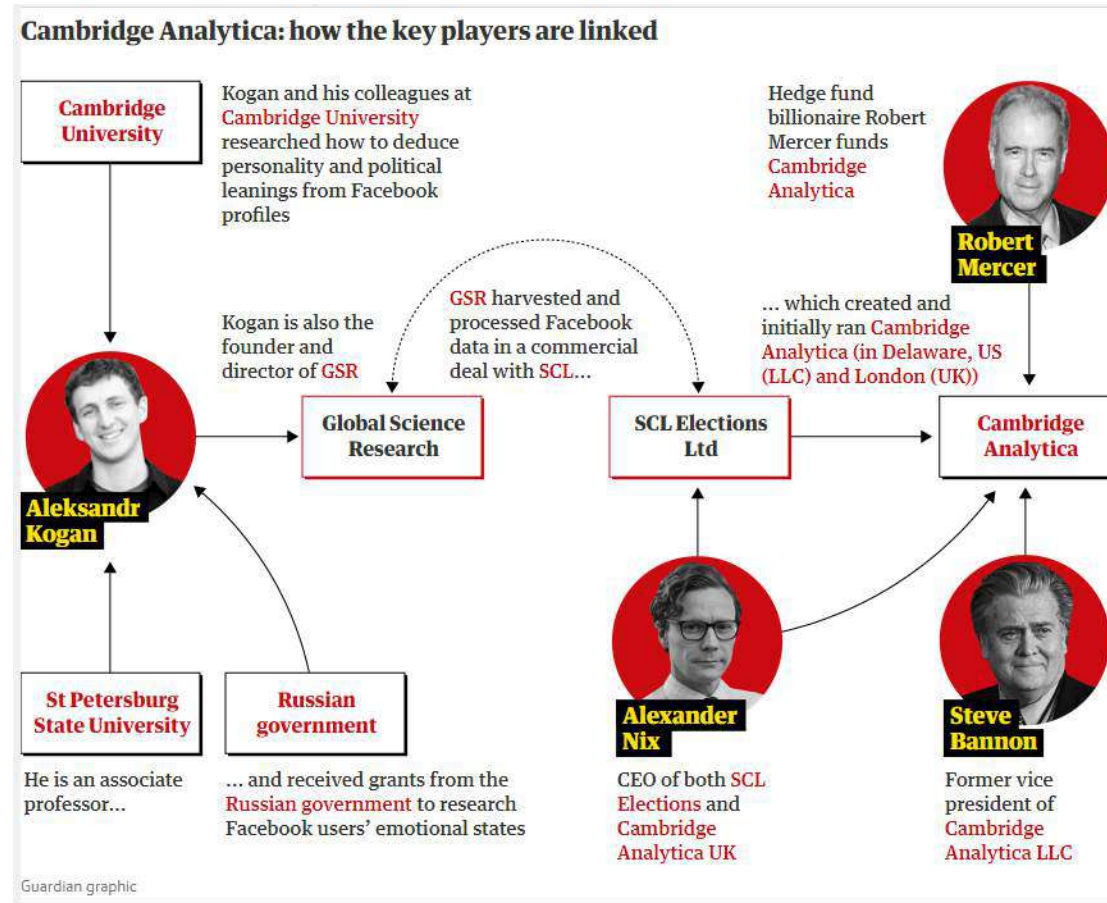
Source: <https://twitter.com/jpwarren/status/1055548155410804736?s=11>

“Netflix pushes content based on whether the company’s algorithm thinks it will make us click and not necessarily whether that content will be good or bad for us.”



Why Netflix Should Scare You More Than It Does, Huffpost, 11 Oct 2018, Todd Van Luling,  
[https://www.huffingtonpost.com.au/entry/netflix-scared-bad\\_us\\_5bbcd832e4b01470d055d4b3](https://www.huffingtonpost.com.au/entry/netflix-scared-bad_us_5bbcd832e4b01470d055d4b3)

“Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach”



<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>



# Why ethics and diversity are important



Moral principles that govern a person's behaviour or the way in which they conduct an activity...

**“We ask ethical questions whenever we think about how we should act. Being ethical is a part of what defines us as human beings.”**

The Ethics Centre, Sydney

# Ethics of:

- **Data** - how we generate, record & share data
- **Algorithms** - how we interpret data via artificial intelligence, machine learning and robots
- **Practices** - devising responsible innovation and professional codes to guide this emerging science

Luciano Floridi, Mariarosaria Taddeo, What is data ethics?  
Phil. Trans. R. Soc. A 2016 374 20160360; DOI: 10.1098/rsta.  
2016.0360. Published 14 November 2016

“Some of the key findings are **intuitive**: participants prefer to save people over animals, the young over the old, and more rather than fewer. Other preferences are more **troubling**: women over men, executives over the homeless, the fit over the obese.”

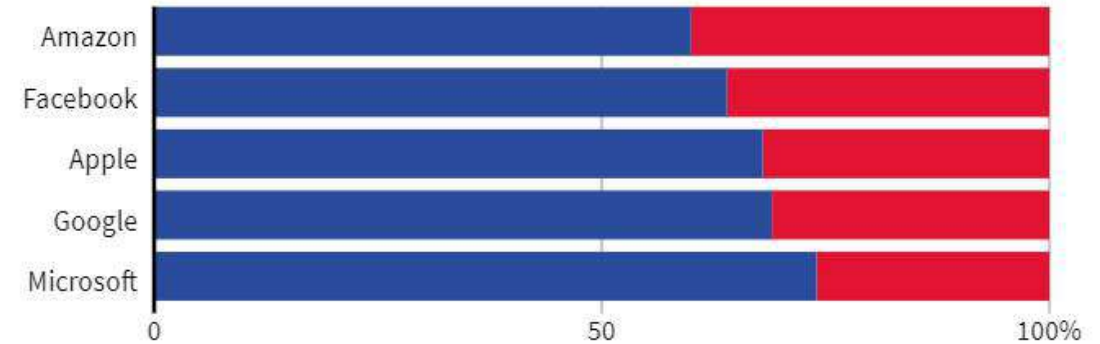
Seth Lazar, Colin Klein, Why we need more than just data to create ethical driverless cars, The Conversation, 25 October 2018, <https://theconversation.com/why-we-need-more-than-just-data-to-create-ethical-driverless-cars-105650>

# “Amazon scraps secret AI recruiting tool that showed bias against women”

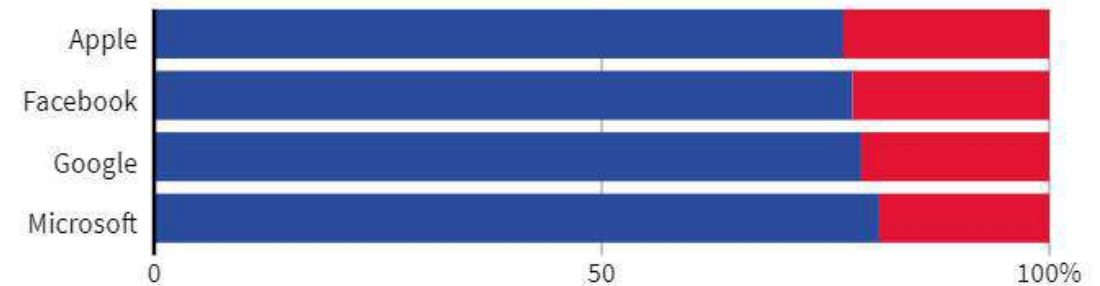
Nov 2018

## GLOBAL HEADCOUNT

■ Male ■ Female



## EMPLOYEES IN TECHNICAL ROLES



Note: Amazon does not disclose the gender breakdown of its technical workforce.

Source: Latest data available from the companies, since 2017.

By Han Huang | REUTERS GRAPHICS

Amazon scraps secret AI recruiting tool that showed bias against women, Jeffrey Dastin, 10 Oct 2018, Reuters

<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

# Amazon's sexist hiring algorithm could still be better than a human

November 1, 2018 10.28am GMT

Lavanchy, Maude. Amazon's sexist hiring algorithm could still be better than a human, November 1, 2018, The Conversation.  
<https://theconversation.com/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human-105270>



Algorithmic bias isn't just sexist or racist

“An algorithm which simply optimizes cost-effectiveness in ad delivery will deliver ads that were intended to be gender-neutral in an apparently discriminatory way, due to crowding out.”


[Lambrecht, A](#) and Tucker, C E (2018) *Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads*. Management Science. ISSN 0025-1909 (In Press)

1 in 4 statisticians say they  
were asked to commit  
scientific fraud

Min Qi Wang, Alice F. Yan, Ralph V. Katz, Researcher Requests for Inappropriate Analysis and Reporting: A U.S. Survey of Consulting Biostatisticians, *Annals of Internal Medicine*, October 2018, DOI 10.7326/m18-1230

# VW engineer sent to the clink for three years for emissions-busting code

James Liang gets 40 months on the cooler and \$200,000 fine

By [Kieren McCarthy](#) in [San Francisco](#) 25 Aug 2017 at 19:44 120  SHARE ▼



McCarthy, Kieren. VW engineer sent to the clink for three years for emissions-busting code: James Liang gets 40 months on the cooler and \$200,000 fine, The Register, 25 Aug 2017, [https://www.theregister.co.uk/2017/08/25/vw\\_engineer\\_gets\\_3yrs\\_for\\_emissionbusting\\_sw/](https://www.theregister.co.uk/2017/08/25/vw_engineer_gets_3yrs_for_emissionbusting_sw/)



- For years we have assumed that programming is a neutral act
- We also assumed that the default human being was a white male
- We also assumed that software would not have real world impacts

1. Technology has no ethics. People demonstrate ethics.
2. Technology inherits the biases of its makers.
3. Therefore we need diversity and formal mechanisms to reduce bias.

Inclusion is the process that begets  
diversity

Build inclusive practices and  
diversity will come



What can be  
done?

# Things we can do

- New approaches to:
  - Design & development of hardware & software
  - Managing the emerging info sec threat landscape
  - Regulation of data security and privacy
- New practices for designers, developers, and business people

Traditional approaches

Privacy by design

Security by design

Codes of ethics

Guidelines

# Privacy by Design

1. **Proactive** not Reactive; **Preventative** not Remedial
2. Privacy as the **Default Setting**
3. Privacy **Embedded** into Design
4. Full Functionality – **Positive-Sum**, not Zero-Sum
5. End-to-End Security – **Full Lifecycle Protection**
6. **Visibility** and **Transparency** – Keep it Open
7. **Respect** for User Privacy – Keep it **User-Centric**

Source: <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>

# OWASP Security by Design Principles

1. **Minimize** attack surface area
2. Establish **secure defaults**
3. Principle of **Least privilege**
4. Principle of **Defence in depth**
5. Fail **securely**
6. **Don't trust** services
7. Ensure **Separation of duties**
8. Always **avoid** security by obscurity
9. Keep security **simple**
10. Fix security issues **correctly**



Open Web Application Security Project [https://www.owasp.org/index.php/Security\\_by\\_Design\\_Principles](https://www.owasp.org/index.php/Security_by_Design_Principles)



“Public agencies urgently need a practical framework to assess automated decision systems and to ensure public accountability.”

Source: Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability, Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, April 2018, <https://ainowinstitute.org/aiareport2018.pdf>

## KEY ELEMENTS OF A PUBLIC AGENCY ALGORITHMIC IMPACT ASSESSMENT

1. Agencies should conduct a self-assessment of existing and proposed automated decision systems, evaluating potential impacts on fairness, justice, bias, or other concerns across affected communities;
2. Agencies should develop meaningful external researcher review processes to discover, measure, or track impacts over time;
3. Agencies should provide notice to the public disclosing their definition of “automated decision system,” existing and proposed systems, and any related self-assessments and researcher review processes before the system has been acquired;
4. Agencies should solicit public comments to clarify concerns and answer outstanding questions; and
5. Governments should provide enhanced due process mechanisms for affected individuals or communities to challenge inadequate assessments or unfair, biased, or otherwise harmful system uses that agencies have failed to mitigate or correct.

Source: Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability, Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, April 2018, <https://ainowinstitute.org/aiareport2018.pdf>

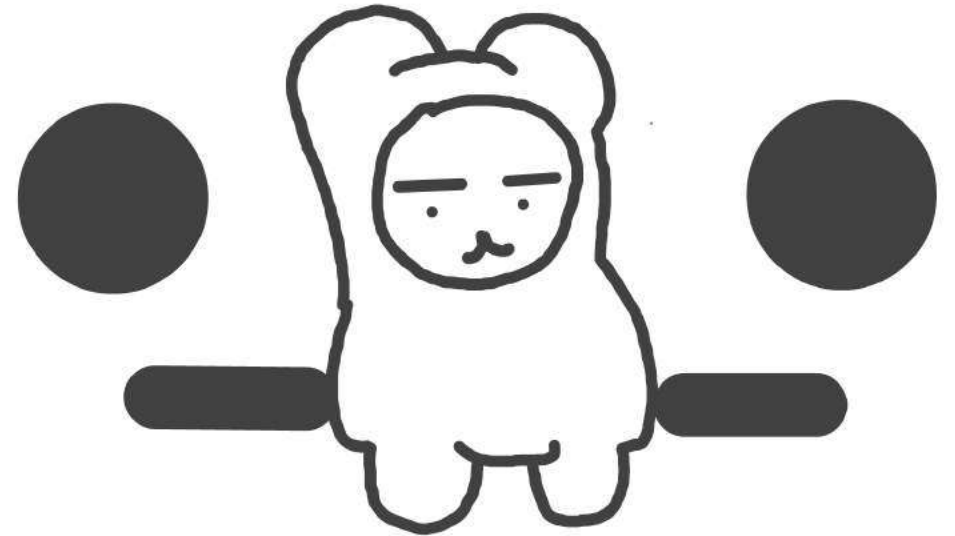
# House of Lords Artificial Intelligence Committee 2017

The Lords' report proposes five main principles for an AI code:

1. Artificial intelligence should be developed for the **common good and benefit of humanity**
2. Artificial intelligence should operate on **principles of intelligibility and fairness**
3. Artificial intelligence **should not be used to diminish the data rights or privacy** of individuals, families or communities
4. All **citizens have the right to be educated** to enable them to flourish mentally, emotionally and economically alongside artificial intelligence
5. The autonomous power to hurt, destroy or deceive human beings **should never be vested in artificial intelligence**

Source: Report of Session 2017-19 - published 16 April 2017 - HL Paper 100  
<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm>

“Governor Jerry Brown recently signed [S.B. 1001](#), a new law requiring all “bots” used for purposes of influencing a commercial transaction or a vote in an election to be labeled.”



© 2016 Maximus Renegade .... Licensed under [CC-BY](#).

Victory! Dangerous elements removed from California's bot-labeling bill, Jamie Williams & Jeremy Gillula, 5 October 2018 <https://www.eff.org/deeplinks/2018/10/victory-dangerous-elements-removed-californias-bot-labeling-bill>



## [Code of Practice for consumer IoT security](#)

HTML



## [Code of Practice for consumer IoT security](#)

PDF, 483KB, 24 pages



## [Consumer guidance for smart devices in the home](#)

HTML

### **Secure by Design**

The [UK] Government's Code of Practice for Consumer Internet of Things (IoT) Security for manufacturers, with guidance for consumers on smart devices at home.

Secure by Design, 7 March 2018,  
UK Department for Digital, Culture, Media & Sport.  
<https://www.gov.uk/government/publications/secure-by-design>

# UK government guidance











1. No default passwords
2. Implement a vulnerability disclosure policy
3. Keep software updated
4. Securely store credentials and security-sensitive data
5. Communicate securely
6. Minimise exposed attack surfaces
7. Ensure software integrity
8. Ensure that personal data is protected
9. Make systems resilient to outages
10. Monitor system telemetry data
11. Make it easy for consumers to delete personal data
12. Make installation and maintenance of devices easy
13. Validate input data

<https://www.gov.uk/government/publications/secure-by-design/code-of-practice-for-consumer-iot-security>

# Ethics Canvas

Project Title:

Date:

<p><b>Individuals affected</b></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> <b>1</b></p>	<p><b>Behaviour</b></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> <b>3</b></p>	<p><b>What can we do?</b></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> <b>9</b></p>	<p><b>Worldviews</b></p> <p>Discuss how the general perception of somebody's role in society can be affected by the project.</p> <p> <b>5</b></p>	<p><b>Groups affected</b></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> <b>2</b></p>
<p><b>Relations</b></p> <p>Discuss problematic differences in individual behaviour such as differences in habits, time-schedules, choice of activities, etc</p> <p> <b>4</b></p>		<p><b>Group Conflicts</b></p> <p>Discuss the impact on the relationships between the groups identified, e.g. employers and unions</p> <p> <b>6</b></p>		<p><b>Problematic Use of Resources</b></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> <b>8</b></p>
<p><b>Product or Service Failure</b></p> <p>Discuss the potential negative impact of your product or service failing to operate as intended, eg technical or human error, financial failure/ receivership/acquisition, security breach, data loss, etc.</p> <p> <b>7</b></p>			<p><b>Problematic Use of Resources</b></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> <b>8</b></p>	

The Ethics Canvas is adapted from Alex Osterwalder's Business Model Canvas. The Business Model Canvas is designed by: Business Model Foundry AG. This work is licensed under the Creative Commons Attribution-Share Alike 3.0 unported license. To view a copy of this license, visit <https://creativecommons.org/licenses/by-sa/3.0/>. To view the original Business Model Canvas, visit <https://strategyzer.com/canvas>.

# Recap

- Data is the new oil
- Software is eating the world
- Hardware is helping it do so
- Weaponization of social media shows us power of unintended consequences
- Cyber is now everyone's problem
- We need to build in ethics & build diversity in teams



Contact me: [k.carruthers@unsw.edu.au](mailto:k.carruthers@unsw.edu.au)

SlideShare: <https://www.slideshare.net/carruthk>



# Some resources

- Luciano Floridi, Mariarosaria Taddeo. What is data ethics? [Phil. Trans. R. Soc. A](#) 2016 374 20160360; DOI: 10.1098/rsta.2016.0360. Published 14 November 2016
- Digital Enlightenment Forum: Digital Ethics. Workshop Report. (2016, March 1). Retrieved August 16, 2017, from <https://digitalenlightenment.org/sites/default/files/users/14/Digital%20Ethics%20Workshop%20Report%20v2.pdf>
- Maggiolini, Piercarlo. A deep study on the concept of digital ethics. (2014). *Revista de Administração de Empresas*, 54(5), 585-591. <https://dx.doi.org/10.1590/S0034-759020140511>
- [Lambrecht, A](#) and Tucker, C E (2018) *Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads*. *Management Science*. ISSN 0025-1909 (In Press)
- Lavanchy, Maude. Amazon's sexist hiring algorithm could still be better than a human,. November 1, 2018, The Conversation <https://theconversation.com/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human-105270>
- Reisman, Dillon, Schultz, Jason, Crawford, Kate, Whittaker, Meredith. *Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability*, April 2018, <https://ainowinstitute.org/aiareport2018.pdf>
- McCarthy , Kieren. VW engineer sent to the clink for three years for emissions-busting code: James Liang gets 40 months on the cooler and \$200,000 fine, The Register, 25 Aug 2017, [https://www.theregister.co.uk/2017/08/25/vw\\_engineer\\_gets\\_3yrs\\_for\\_emissionbusting\\_sw/](https://www.theregister.co.uk/2017/08/25/vw_engineer_gets_3yrs_for_emissionbusting_sw/)
- Singer, Peter Warren, and Brooking, Emerson T.. *Likewar: the Weaponization of Social Media*. Eamon Dolan/Houghton Mifflin Harcourt, 2018.
- Sanger, David E.. *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age*. Crown Publishers., 2018.



# Some resources

- [AI Now Institute](#) at New York University
- [Causeit Data Ethics](#)
- [The BIG Data Ethics Cheat Sheet](#), Hackermoon
- [Guidelines on Ethical Research](#)
- IEEE AI and Ethics in Design course  
<http://innovationatwork.ieee.org/new-course-program-now-available-ai-and-ethics-in-design/>
- Tufts University course [Ethics of AI, Robotics and Human-Robot Interaction](#)
- [Digital Ethics Lab - Oxford Internet Institute - University of Oxford](#)
- [Georgetown University, Kennedy Institute of Ethics, Ethics Lab](#)
- [British Sociological Association](#)

