

## The Digital Coalface: Ethical Dilemmas of Artificial Intelligence

Douglas Austrom & Carolyn Ordowich

Global Network for SMART Organization Design

Leiden, September 7, 2018

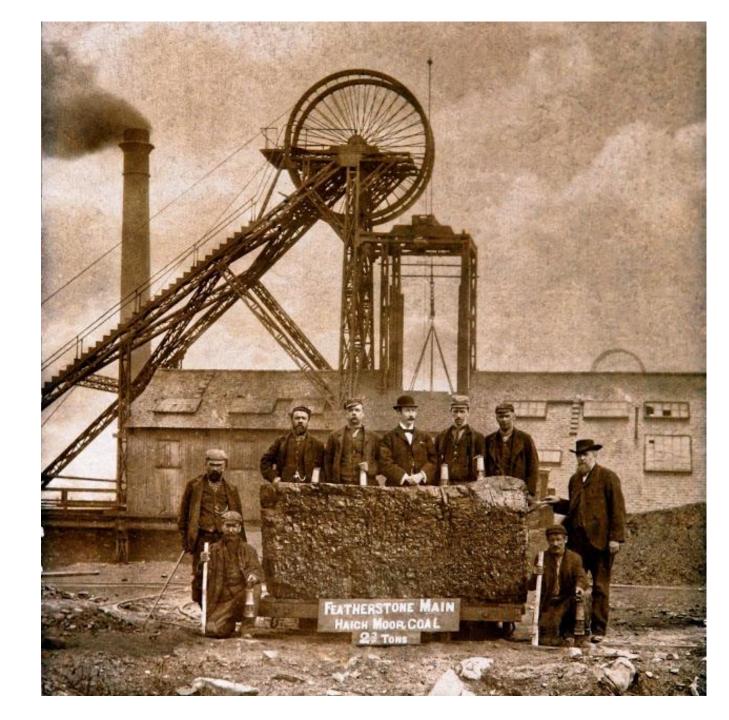
## PURPOSE OF WORKSHOP

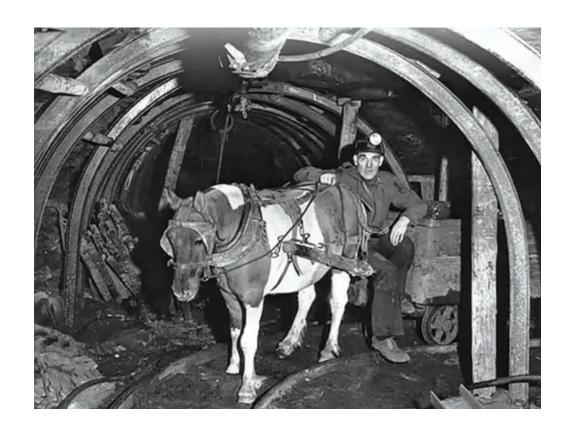
 To raise awareness of, and explore, ethical considerations (challenges, opportunities - both intended and unintended) that arise in the development and application of digital technologies.

 To explore principles already in existence to counteract technodeterminism

• To examine our own awareness of these ethical issues and how to address as a designer

Haigh Moor Colliery

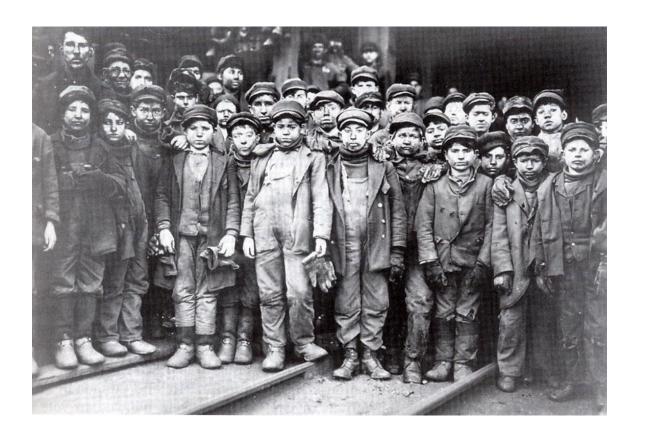




The Harsh Reality







## The **Face** of Coal Mining Today

## The **Faces** of Coal Mining in 1911



## Choice and the Techno-Bureaucratic Imperative

There has arisen in organizational studies a theory and practice that has disestablished the technological imperative from its long reign of unchallenged rule which has created the technocratic bureaucracies that still remain the predominant organizational form in advanced industrial societies.

Eric Trist, 1973

## Digital Technology will Either Be ...

#### **Constraining**

- Diminish life and human dignity
- Digital Taylorism and an insidious extension of bureaucratic design principles ... pervasive use of cameras, facial recognition, wearables/ implantables ... "Big Brother" monitoring of all our movements
- Reinforce C&C v1 for how we coordinate human endeavors...
  - COMMAND & CONTROL

### Liberating

- Give people and life the potential to flourish as never before
- Positively augment and extend human capabilities ... facilitating horizontal coordination while reducing transaction costs to virtually zero
- Enable **C&C v2** for how we coordinate human endeavors ...
  - CONNECT & COLLABORATE

It's more technically feasible to automate predictable physical activities than unpredictable ones.

#### Technical feasibility of automation, %1

Predictable physical work

78%

For example, welding and soldering on an assembly line, food preparation, or packaging objects

Unpredictable physical work



For example, construction, forestry, or raising outdoor animals

1% of time spent on activities that can be automated by adapting currently demonstrated technology.

McKinsey&Company

## Use Case ... Wikipedia

It's hard to grasp just how important Wikipedia has become for the world, and how vulnerable. It is the fifth most visited website, serving more than 15 billion pageviews per month. It includes nearly 50 million articles, written in almost 300 languages—only 13% in English. It boggles the mind that all of this is created by human volunteers.

The human authorship of Wikipedia is its strength. The deliberative process of the editors ensures that Wikipedia remains robust and tends toward consensus. Just visit Twitter to see what a non-deliberative information platform looks like where bots roam free.

But with human hands come human limitations. As it becomes more and more essential to the world, biased and missing information on Wikipedia will have serious impacts. The human editors of the most important source of public information can be supported by machine learning. Algorithms are already used to detect vandalism and identify underpopulated articles. But the machines can do much more. They can track and summarize information missing from Wikipedia articles. They can even identify articles that are missing altogether, and generate the first draft.

To solve the recall problem of human-generated knowledge bases, we need to superpower the humans.

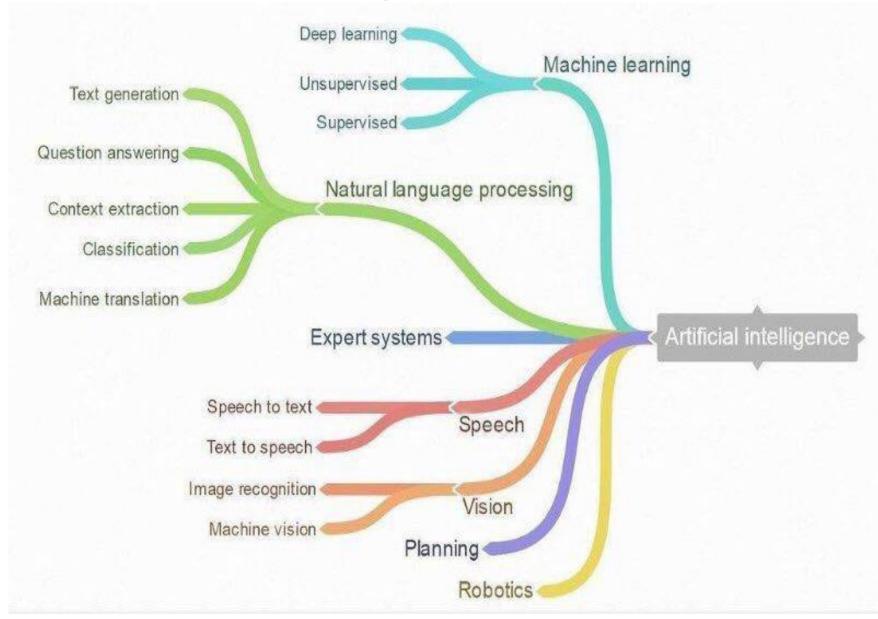


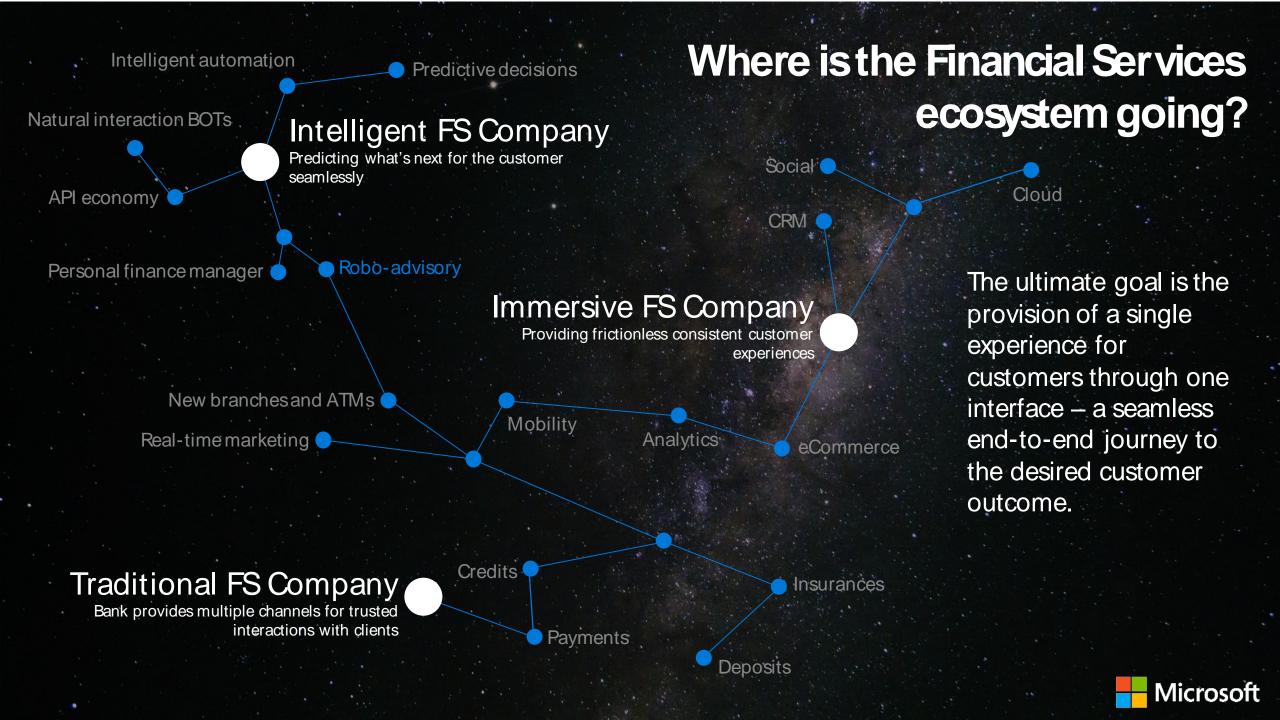
Use Case ...
Workers in
Exoskeleton
Robotic Suits



## Making Sense of Artificial Intelligence

Al is going to be a seismic shift in business — and it's expected to create a \$15.7 trillion economic impact globally by 2030.





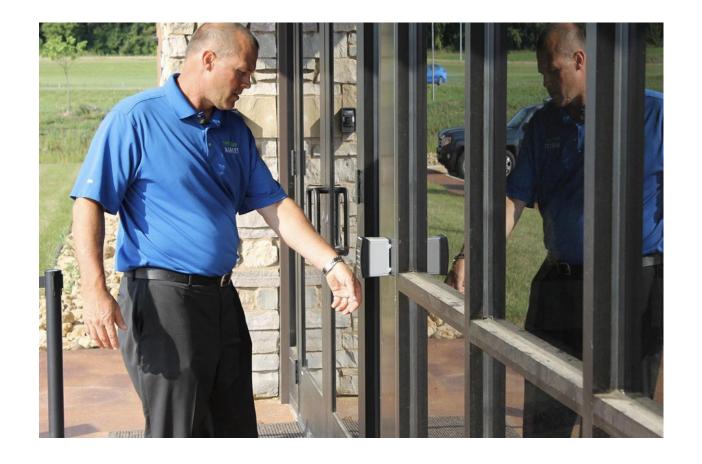


**Intelligent Machines** 

# This company embeds microchips in its employees, and they love it

Last August, 50 employees at Three Square Market got RFID chips in their hands. Now 80 have them.

by Rachel Metz August 17, 2018





#### **Intelligent Machines**

## Who needs democracy when you have data?

Here's how China rules using data, AI, and internet surveillance.

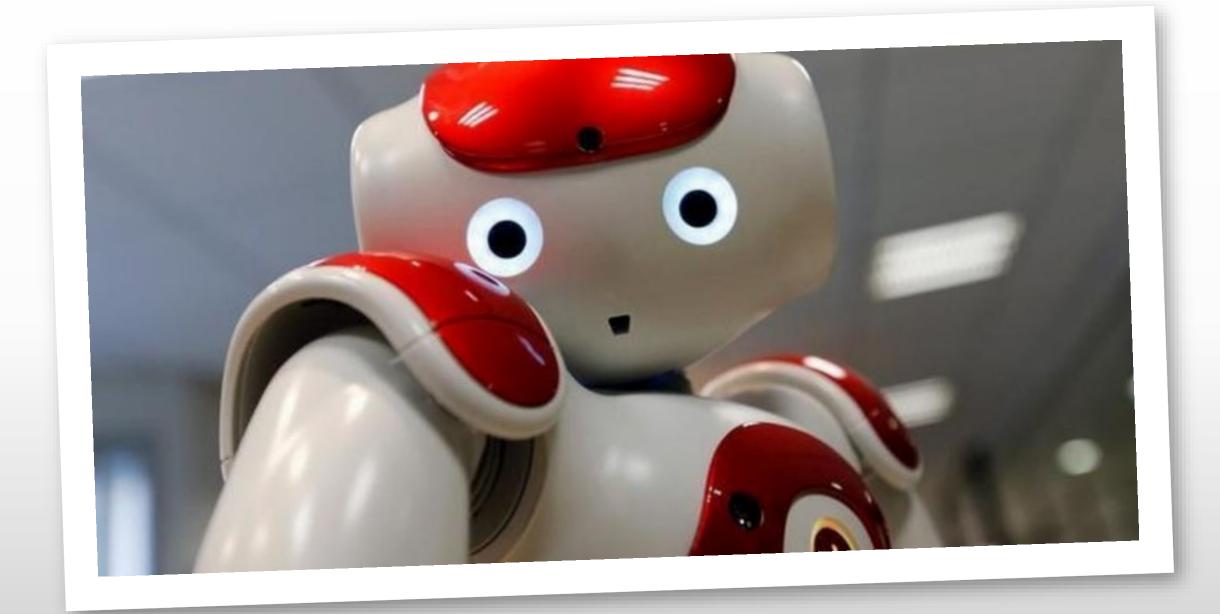
by Christina Larson August 20, 2018

People in Beijing are always under the watchful eye of Mao—and myriad surveillance cameras.











https://www.youtube.com/watch?v=maC2D4KZTyE

## Not with a Bang, but ...

#### One Convenience at a Time

Smartphones insinuated themselves into our lives. Now, think about the iPhone. Ten years ago, smartphones barely existed. Five years ago, they were mediocre phones, maybe good music players with short battery lives and not much else because the Wi-Fi was so bad everywhere you couldn't do much with them. Today, everybody uses them to do everything. No one made that decision. No one said, "OK, now we're going to let iPhones change our lives, disrupt our dinner conversations, and change the way we conduct business meetings." It happened one convenience, one cost saving at a time, and it changed our societies.

That, to me, is how future globalization will occur: one convenience at a time, one job at a time—not being replaced in every variety of office. Nobody will ever decide to have a job apocalypse in which we replace all the service-sector workers or all the doctors or all the lawyers. But it's already happening in media. It's happening in law. It's happening at the low end of medicine. And I think we're getting close to the holy-cow moment.

Richard Baldwin, Graduate Institute of International & Development Studies and Centre for Economic Policy Research

## How They Capture Our Attention ...

... isn't always best for our well-being.

- Snapchat turns conversations into streaks, redefining how our children measure friendship.
- Instagram glorifies the picture-perfect life, eroding our self-worth.
- Facebook segregates us into echo chambers, fragmenting our communities.
- YouTube auto plays the next video within seconds, even if it eats into our sleep.

These are NOT neutral products. They are part of a system **designed to addict us.** 



## Why and How Things Are Different Now



#### **Artificially Intelligent**

No other media drew on massive supercomputers to predict what it could show to *perfectly* keep you scrolling, swiping or sharing.



#### 24/7 Influence

No other media steered two billion people's thoughts 24/7 – checking 150 times per day – from the moment we wake up until we fall asleep.



#### **Social Control**

No other media redefined the terms of our social lives: self-esteem, when we believe we are missing out, and the perception that others agree with us.



#### **Personalized**

No other media used a precise, personalized profile of everything we've said, shared, clicked, and watched to influence our behavior at this scale.

## Choices, Choices

From	То
End-users being disengaged bystanders	Individuals negotiating with companies on equal terms
Opacity	Transparency
Unawareness	Agency
No control over how data is used	Ownership and profit sharing when data is used
Lack of awareness of the underlying values	Global agreement on the values that should enshrined

## Digital Technology will Either Be ...

#### **Constraining**

- Diminish life and human dignity
- Digital Taylorism and an insidious extension of bureaucratic design principles ... pervasive use of cameras, facial recognition, wearables/ implantables ... "Big Brother" monitoring of all our movements
- Reinforce C&C v1 for how we coordinate human endeavors...
  - COMMAND & CONTROL

### Liberating

- Giv people and life the potential to flourish as never before
- Positively augment and extend human capabilities ... facilitating horizontal coordination while reducing transaction costs to virtually zero
- Enable **C&C v2** for how we coordinate human endeavors ...
  - CONNECT & COLLABORATE

## Moving from 'Can We' to 'Should We'

Technologists, business people and organization designers need to ask key user-centered questions before launching new business models, new products or new ways of organizing.

They must understand technology's impact on their context by asking themselves ...

- Will this technology result in overall good?
- What might be some unintended consequences of this technology?
- What are the social and ethical impacts of the technology?
- Will this technology augment human intellect, disrupt it, or substitute for it?
- How could this technology be used negatively against users?

## strategy+business

**AUTUMN 2018 ISSUE** JULY 31, 2018

# The Future of Artificial Intelligence Depends on Trust

If it is to drive business success, AI cannot hide in a black box.

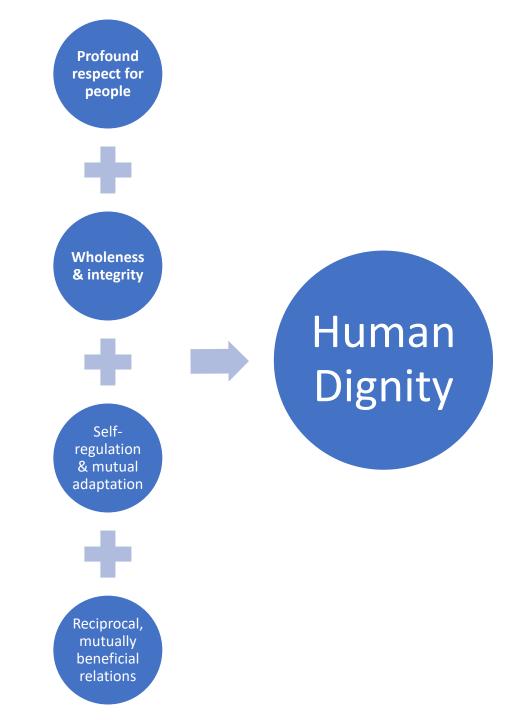
	Ethics	Morals			
What are they?	The rules of conduct recognized in respect to a particular class of human actions or a particular group or culture.	Principles or habits with respect to right or wrong conduct. While morals also prescribe dos and don'ts, morality is ultimately a personal compass of right and wrong.			
Where do they come from?	Social system - External	Individual - Internal			
Why we do it?	Because society says it is the right thing to do.	Because we believe in something being right or wrong.			
Flexibility	Ethics are dependent on others for definition. They tend to be consistent within a certain context, but can vary between contexts.	Usually consistent, although can change if an individual's beliefs change.			
The "Gray" A person strictly following Ethical Principles may not have any Morals at all. Likewise, one could violate Ethical Principles within a given system of rules in order to maintain Moral integrity.		A Moral Person although perhaps bound by a higher covenant, may choose to follow a code of ethics as it would apply to a system. "Make it fit"			
Origin	Greek word "ethos" meaning "character"	Latin word "mos" meaning "custom"			
Acceptability	Ethics are governed by professional and legal guidelines within a particular time and place	Morality transcends cultural norms			

## Defining Moment for AI Ethics is ... NOW!

The decisions we make now are going to sit at the core of our models for years and continue to evolve, continue grow and continue to learn. So we need to set them on **a firm ethical foundation** so that as they grow through the years they'll continue to reflect our values.

Darin Stewart, VP Gartner Research

## STS First Principles

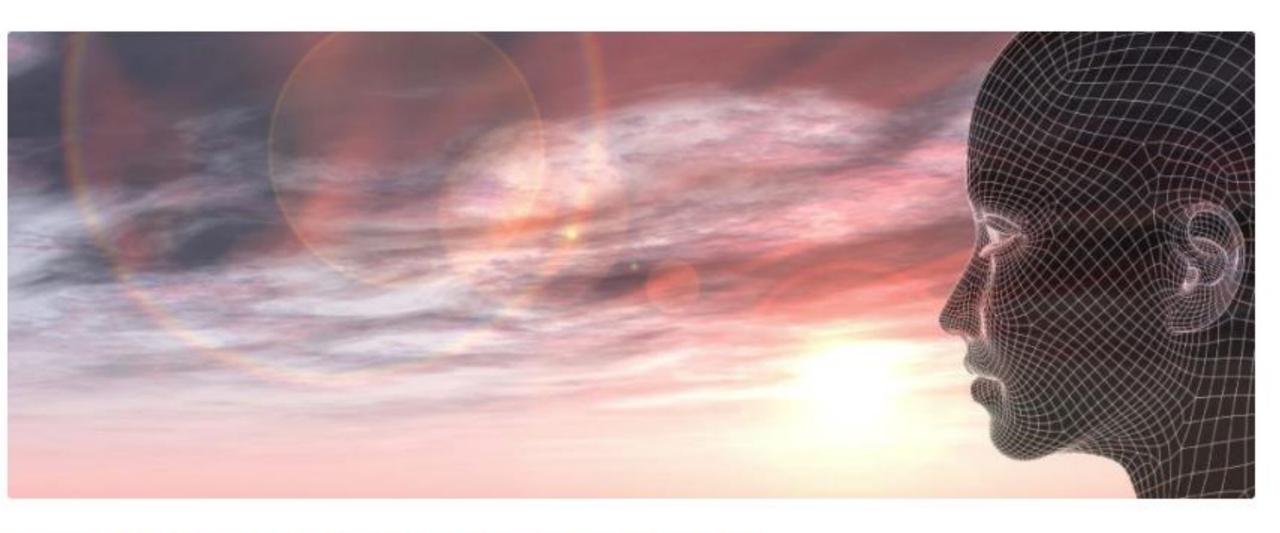


## Google's Believes that AI Should ...

- Be socially beneficial.
- Avoid creating or reinforcing unfair bias.
- Be built and tested for safety.
- Be accountable to people.
- Incorporate privacy design principles.
- Uphold high standards of scientific excellence.
- Be made available for uses that accord with these principles.

## Al Applications Google Will Not Pursue

- Technologies that cause or are likely to cause overall harm. Where there is a material risk of harm, we will proceed only where we believe that the benefits substantially outweigh the risks, and will incorporate appropriate safety constraints.
- Weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people.
- Technologies that gather or use information for surveillance violating internationally accepted norms.
- Technologies whose purpose contravenes widely accepted principles of international law and human rights.



### State of California Endorses Asilomar Al Principles

August 31, 2018 / by The FLI Team

Artificial intelligence has vast potential, and its responsible implementation is up to us.

## Small Group Deliberation – 25 min

• In your small group discuss ...

What are the ethical principles that you believe we as organization designers need to incorporate in our work in order to challenge techno-determinism and the Digital Coalface?

- Identify your top 3-5 ethical principles
- Prepare to share with the total group

## **Group Reports**

- 3 minutes per group
- Each group listens to other groups for similarities and tick those on your flipchart which match what you heard from other groups
- Whole group to reflect on "common ground"

#### TABLE OF DISRUPTIVE TECHNOLOGIES

De Digital footprint eraser	Ps Personal digital shields	Human head transplants	HC Human cloning & de-extintion	Da Distributed autonomous corporations	Sp Space solar power	EL Space elevators	Fully immersive virtual reality (VR)	Artificial consciousness	Qt We can't talk about this one
Conversational machine interfaces	Life-expectancy algorithms  82 DE	Sa Stratospheric aerosols	Br Battlefield robots	Ad advisors & decision -making machines  85 DE	Ab Al board members & politicians  86 EA	Invisibility shields 87 SP	Ph Factory photosynthesis 88 SP	Th Transhuman technologies 89 HA	Te Telepathy  90 HA
SS Planetary-scale spectroscopy 71 SP	I p   Implantable phones   72 MI	He e-tagging of humans 73 DE	Mp Male pregnancy & artificial wombs	DNA data storage	GV Genomic vaccines 76 SP	Qs Quantum safe cryptography 77 DE	Cp Cognitive prosthetics	Data uploading to the brain  79 HA	Reactionless drive
Gh Predictive gene-based healthcare  61 DE	Automated knowledge discovery	Rs Autonomous robotic surgery 63 EA	Emotionally aware machines MI	Humanoid sex robots MI	Bh Human bio-hacking	Me Internet of DNA	Thought control machine interfaces  68 MI	Dream reading & recording HA	Whole Earth virtualisation  70 DE
Mega-scale desalination 51 SP	Sw Self-writing software 52 EA	Mm Public mood monitoring 53 DE	Pb Programmable bacteria	Peer-to-peer energy trading & transmission  55 DE	La Lifelong personal avatar assistants 56 MI	Sd Smart dust 57 DE	Low-cost space travel	PC Planet colonization  59 HA	Shape-shifting matter  60 SP
Mc Medical tricorders  41 DE	Sf Smarf flooring & carpets  42 DE	Dt Diagnostic toilets	Se Smart energy grids	Bf Algal bio-fuels  45 SP	Op Human-organ printing  46 SP	BS Artificial human blood substitute	New materials 48 SP	Fusion power 49 SP	Self-reconfiguring modular robots  50 SP
Distributed ledgers 31 DE	Precision agriculture	Autonomous vehicles 33 EA	Intention decoding algorithms  34 MI	Drone freight delivery 35 EA	App Autonomous passenger aircraft 36 EA	Fp 3D-printing of food & pharmaceuticals 37 SP	Sr Swarm robotics	4-dimensional materials  39 SP	Ze Zero-point energy
Robotic care companions  M	Sc Smart controls and appliances 22 DE	Contured meat  23 SP	Pelivery robots & passenger drones	Autonomous ships & submarines 25 EA	Resource gamification	Water harvesting from air	Broadcasting of electricity  28 SP	Bp Bio-plastics	Beam-powered propulsion SP
Cryptocurrencies 11 DE	Concentrated solar power  12 SP	Pp Predictive policing	Micro-scale ambient energy harvesting  14 SP	Airborne wind turbines  15 SP	Avatar companions  16 MI	Metallic hydrogen energy storage	Smart glasses & contact lenses  18 HA	Pellution eating buildings  19 SP	Force fields 20 SP
Sn Smart nappies	Deep ocean wind farms 2 SP	Va Vertical agriculture 3 SP	We Wireless energy transfer SP	Balloon-powered internet  5 SP	Powered exoskeletons  HA	Computerized shoes & clothing  7 DE	Vacuum-tube transport  8 SP	Sj Scram jets 9 SP	Asteroid mining  10 SP

#### A dashboard of 100 wonderful, weird (and possibly worrying) ways the world might change in the foreseeable future

#### Example of organizations active in each area

- 1 Monit (South Korea), Abena Nova (Denmark), Siempre Secos (Spain)
- 2 Statoil (Norway), Siemens (Germany), Volturn (US), UMaine (US)
- 3 Green Skies Vertical Farms (US), Aero Farms (US), Neo Farms (Germany), Urban Crop Solutions (Belgium)
- 4 WiTricity (US), Powermat (Israel), Apple/Power By Proxi (US), Qualcomm (US), Mojo Mobility (US), Mopar (US), Fulton Innovation (US)
- 5 Google/Alphabet (US)
- 6 ReWalk (US), Rex Bionics (US), SuitX/US Bionics (US), Ekso Bionics (US), Lockheed Martin (US)
- 7 Google/Alphabet (US), Samsung (Korea), Hexoskin (Canada) Owlet (US), Komodo Tech (Canada), Shiftwear (US), Lechal (India), OM Signal (Canada)
- 8 The Boring Company/Elon Musk (US), China Aerospace Science and Industry Corporation (China)
- 9 Reaction Engines (UK), NASA (US), Boeing (US), Lockheed Martin (US), Airbus (France)
- 10 Deep Space Industries (US), Planetary Resources (US), Made in Space (US)
- 11 Bitcoin (Japan), Ripple (US), Litecoin (US)
- 12 Solarreserve (US), Abengoa (Spain), North China Power Engineering (China), Shanghai Electric (China), Zhejiang Supcon Solar (China), NWEPDI (China)
- 13 PredPol (US), ECM Universe (US)
- 14 Pavegen (UK), ECEEN (China)
- 15 Google/Alphabet (US), Joby Energy (US), Altaeros (US), Kitegen (Italy), Enerkite (Germany)
- 16 Pullstring (US), Amazon (US), Alphabet/Google (US), Nintendo (Japan), Invisible Girlfriend/Boyfriend (US)
- 17 NASA (US)
- 18 Alphabet/Verily (US), Amazon (US), Vuzix (US), Everysight (Israel)
- Elegant Embellishments (Germany), iNova (Spain), Studio Roosegaarde (Netherlands), Prosolve 370e (Germany)
- 20 Dstl (UK), Boeing (US)
- 21 Softbank (Japan), AIST (Japan), Blue Frog Robotics (France), Care-o-bot (Germany), Riken/Sumitomo Riko (Japan), Mayfield Robotics (US)
- 22 Amazon (US), Google/Alphabet (US), Philips (Netherlands), Samsung (South Korea), Dyson (UK), Miele (Germany), iRobot (US)
- 23 Impossible Foods (US), Memphis Meats (US), Super Meat (Israel), Finless Foods (US), New Harvest (US)
- 24 Wing/Alphabet (US), Starship Technologies (UK), Volocopter (Germany), eHang (China), Piaggio (Italy)
- 25 Leidos (US), Boeing (US), Rolls Royce (UK)
- 26 Joulebug (US), Waterpebble (UK)
- 27 Permalution (US), Sun to Water (US)
- 28 Powercast (US)
- 29 NatureWorks (US), Gruppo MAIP (Italy), Genomatica (US), Green Dot Bioplastics (US)
- 20 NIACA (LIC)

LATER

31 Everledger (UK), Stampery (Spain), Brickblock (Germany), Slock.it (Germany)

- 32 Blue River Technology (US), Hortau (Canada)
- 33 Google/Waymo (US), Voyage (US), Nvidia Automotive (US), most major auto-makers
- 34 Amazon (US), Google/Alphabet (US), Philips (Netherlands), Samsung (South Korea), Dyson (UK), Miele (Germany), iRobot (US)
- 35 Google/Alphabet (US), Amazon (US), Flirtey (US)
- 36 Airbus (France), Boeing (US)
- 37 FabCafe (Japan), NASA (US)
- 38 SRI International (US)
- 39 Stratasys (US), Autodesk (US)
- AN NIACA ILIC
- 41 Basil Leaf Technologies (US), Dynamical Biomarkers Group (US/Taiwan), Scanadu (US)
- 42 Starwood Hotels (US), MariCare (Finland), Scanalytics (US), Futureshape (Germany)
- 43 Flowsky (Japan), Scanadu (US)
- 44 Tesla (US), ABB (Switzerland), Siemens (Germany), IBM (US), Itron (US)
- **45** Synthetic Genomics/ExxonMobil (US), Global Algae Innovations (US), Algenol (US)
- 46 Organavo (US), Envision TEC (Germany), RegenHU (Switzerland), Cellink (Sweden), Seraph Robotics (US)
- 47 Hb02 Therapeutics (South Africa), Biospace (US)
- 48 For example Vantablack by Surrey NanoSystems (UK)
- 49 ITER (EU/France), Tokamak Energy (UK), Alphabet/ Google/Tri Alpha Energy (US), General Fusion (Canada), Helion Energy (US), Lockheed Martin (US)
- 50 Festo (Germany)
- 51 Israel Desalination Enterprises Technologies (Israel), Acciona (Spain), Fluence Corporation (US)
- 52 Microsoft (US), Google/Alphabet (US), Open AI (US)
- 53 Open Utility/Essent (UK/Netherlands), Knowelsys (China)
- **54** Gingko Bioworks (US), US Naval Research Laboratory (US), US Army Research Lab (US), Darpa (US)
- 55 Open Utility (UK/Netherlands), Power Ledger (Australia), L03 energy (US), Energy Web Foundation (Switzerland)
- 56 Konami Corp (Japan), Mitsuku (UK)
- **57** MOOG (US), Darpa (US)
- 58 Space X/Elon Musk (US), Blue Origin (US), Virgin Galactic (UK), Rocket Lab (US), Axiom Space (US), SpaceIL (Israel), Firefly Aerospace (US)
- 59 Space X (US), UAE Mars Mission (UAE), NASA (US)
- 60 Intel (US)
- 61 Kite Pharma/Gilead Sciences (US), 23andMe (US), Phenogen Sciences (US), Regeneron (US), Veritas Genetics (US)
- 62 IBM (US)
- 63 Intuitive Surgical (US), Verb Surgical/Alphabet/Johnson & Johnson (US), Da Vinci Surgery (US)
- 64 IBM (US), Toyota (Japan), Mimosys (Japan), Persado (US), Joy AI (US)
- 65 Realbotix (US), True Companion (US)

- 66 BioTeq (UK), Grindhouse Wetwear (US), Dangerous Things (US), see also The Eyeborg Project and the
- Cyborg Foundation
  67 Alphabet/Google Genomics (US), Amazon (US), Illumina (US), Oxford Nanopore Technologies/Metrichor (UK)
- 68 CTRL-Labs (US), Emotiv (US), Neuralink (US), maybe Facebook (US)
- 69 No example found
- 70 Improbable (UK)
- 71 European Organization for Astronomical Research in the Southern Hemisphere (European consortium of 16 countries)
- 72 No example found
- 73 Epicenter (Sweden) and Three Square Market 32M (US)
- 74 No example found
- 75 Twist Bioscience (US)
- 76 Vaccinogen (US), EpiVax (US), IBM (US),
- 77 Alphabet/Google (US), KETS (UK), IDQ (Switzerland), Isara (Canada)
- 78 Darna (US)
- 79 Kernel (US), Neuralink/Elon Musk (US), 2045 Initiative (Russia), Darpa (US), General Electric/Braingate (US), possibly Facebook (US)
- 80 NASA (US), Cannae (US)
- 81 Apple (US), Amazon (US), Alphabet/Google (US),
- 82 No example found
- 83 CIA (US)
- 84 Lockheed Martin (US), QinetiQ (UK), Boston Dynamics/Softbank (US/Japan)
- 85 Woebot (US), Pefin (US), LV (UK)
- 86 Deep Knowledge Ventures (Hong Kong), Tieto (Finland)
- 87 BAE Systems (UK), Toyota (Japan). NB. Big difference between optical camouflage and bending light to make things disappear
- 88 Breakthrough Energy (US), RIPE (US), Joint Centre for Artificial Photosynthesis (US)
- 89 SENS Research Foundation (US), Methuselah Foundation/Peter Thiel (US)
- 90 Facebook (US). Neuralink/Elon Musk (US)
- 91 Suicide Machine (Netherlands), Just Delete Me (US)
- 92 No example found
- 93 Turin Advanced Neuromodulation Group (Italy)
- 94 Sooam (South Korea), Revive and Restore (US)
- 95 No example found
- 96 Rebeam (US), Solaren Corp (US)
- 97 Thoth Technology (Canada)
- 98 Improbable (UK), HelloVR (US), Magic Leap (US), Microsoft (US). See also Mind Maze (US), Facebook (US) and possibly Apple (US)
- 99 Possibly Alphabet/Google (US)
- 100 As it says, we can't say

## In the Future ....

Will we will look back at today as a turning point towards humane design ... when we moved away from technology that extracts attention and erodes society, towards technology that protects our minds and replenishes society?