

Redefining the Relationship of Man and Machine

By Gerd Leonhard

What are the challenges and opportunities facing society in the next 10 years as a result of an accelerating pace of technological development?

From technology disruption to furthering human happiness

This chapter aims to provide important context framing for the mission-critical business decisions that we will all need to make in the next few years in strategy, business model development, marketing, and HR. Remaining relevant, unique, purposeful, and indispensable in the future is obviously a key objective for every business everywhere, yet technology will no doubt continue to generate exponential waves of disruptions at an ever-faster pace. Soon – once technology has made almost everything efficient and abundant – I believe that we will need to focus on the truly human values of business, i.e. to transcend technology. Successful business will no longer be about running a well-oiled machine; rather it will be about uniquely furthering human happiness.

Exponential and combinatorial: we're at the pivot point

We are witnessing dramatic digitization, automation, virtualization, and robotization all around us, in all sectors of society, government,

and business – and this is only the beginning. I believe these trends will continue to grow exponentially over the next decade as we head towards a world of five to six billion Internet users by 2020, and possibly as many as 100 billion connected devices in the **Internet of Things**, such as sensors, wearables, and trackers.

Beyond any doubt, machines of all kinds – both software and hardware – will play an increasingly larger role in our future, and progressively more intelligent machines will impact how we live our lives at every turn. Netscape's founder turned venture capitalist Marc Andreessen already highlighted this phenomena in a 2011 Wall Street Journal opinion piece entitled *Why Software is Eating the World*³⁶ – a prescient headline that is certain to play out in force in our imminent future.

We are already nearing the pivot point where very few ideas seem to remain in the realm of science fiction for very long. This can be witnessed in areas such as automated, real-time translation (SayHi, Google Translate, Skype Translate) and self-driving and semi-autonomous cars (Google, Tesla, Volvo). The fiction-reality boundary is also being crossed by developments such as intelligent personal agents (Cortana, Siri, Google Now), augmented and virtual reality (Microsoft Hololens, Oculus Rift) and many other recent breakthroughs. Our world is being reshaped by developments that used to only exist in the scripts of Hollywood blockbusters such as *Blade Runner*, *Her*, *Minority Report*, *Transcendence* and *The Matrix*. (On that note, let's be sure not to give these blockbuster movies too much credit as far as realistic foresight is concerned).

Technology: it's no longer about IF or HOW but about WHY

The urgent need for clear man-machine ethics is amplified by the view that we should probably no longer be concerned whether technology can actually do something, but whether it **should** do something. The how is being replaced by the why (followed by who, when and where).

For example, why would we want to be able to alter our DNA so that we can shape what our babies look like? And who should be able to afford or have access to such treatments? What would be the

FUTURESCAPES

limits? In machine intelligence, should we go beyond mere **deductive** reasoning and allow smart software, robots, and artificial intelligence (AI) to advance to adductive reasoning (i.e. to make unique decisions based on new or incomplete facts and rules)? If autonomous machines are to be a part of our future (as is already a certainty in the military), will we need to provide them with some kind of moral agency, i.e. a human-like capacity to decide what is right or wrong even if the facts are incomplete?

“Hellven” challenges

Tremendous scientific progress in sectors such as energy, transportation, water, environment, and food can be expected in the next 10-20 years. I believe most of these achievements will have an overall positive effect on humanity, and hopefully on human happiness (which I would suggest should be the ultimate goal) as well. This would clearly be the **heavenly** side of the coin.

At the same time, on the **hell** side we are now approaching a series of complex intersections at very high speeds. Soon, every single junction we navigate could either lead to more human-centric gains or result in serious aberrations and grave dangers. It has often been said that, “technology is not good or evil – it just is”. It is now becoming clear that the good / bad part will probably be for us to decide, every day, globally and locally, collectively and individually. Clearly, if we assume that machines will be an inevitably large part of that future, we will need to decide both what we want them to be, and perhaps more importantly, what we want to be as humans – and we need to do it soon.

Artificial Intelligence (AI) is the most significant “hellven” challenge

Most technologies, software and hardware alike, are not only becoming much faster and cheaper but also increasingly intelligent. The spectrum of rapid recent advances runs the gamut from the kind of **simple algorithmic** intelligence it takes to win against a chess master, to the advent of thinking machines and IBM’s neuromorphic chips

(i.e. chips that attempt to mirror our own neural networks) and their ambitious **cognitive computing** initiative. Buzzwords such as AI and deep learning are already making the headlines every single day, and this is just the tip of the iceberg. Looking at the investments by the leading venture capitalists and funds, AI has already become a top priority in Silicon Valley and in China, often a certain sign of what's to come.

At the same time, almost every single major information and communications technology (ICT) company already has several initiatives in this man-machine convergence arena. Google and Facebook are busy acquiring small and large companies in a wide range of AI and robotics-related fields. They clearly realize that the future is not just about **big data**, mobile, and **connected everything**. They see the next horizon as embedding capability to make every process, every object, and every machine truly **functionally intelligent**, albeit not (yet) humanly intelligent as far as social or emotional traits are concerned. But maybe this is just a question of when rather than if?

Just imagine what AI could do to our everyday activities such as searching the web (as we call it today), and you can get a glimpse of what's at stake here. In the very near future, who will bother with typing a precise two-word search phrase into a box when **the system** already knows everything about you, your schedule, your location, your likes, your connections, your transactions, and much more? Based on the situational context, your **external brain** i.e. the **AI in the cloud** will already know what you need, before you even think of it, and will propose the most desirable actions as easily as today's Google maps propose walking directions. Hellven, once again, depends on your standpoint.

IBM, the creator of Watson Analytics, a leading commercially available AI product, appears to be betting the farm on this future. IBM is investing billions of dollars into neurosynaptic chips and cognitive computing – designed to emulate the human neural systems with the intention of creating a **holistic computing experience**, i.e. computing that feels as natural as breathing. Computing is no longer outside of us – a thought both scary and exhilarating. Apart from

FUTURESCAPES

IBM, Google is working on its own **Global Brain** project and the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland is pushing the EU's hotly contested **Human Brain project**. China's Baidu has also signaled its ambitions to discover the holy grails of AI by hiring top-level researchers in that field including Stanford's Andrew Ng, and by opening up a Silicon Valley AI center. The list goes on. Clearly, man-machine convergence is on top of the global agenda and investors smell enormous profits.

But: machines don't have ethics

The AI gold-rush has only just started, and this is probably a very good time to be more concerned about whether Silicon Valley's leading venture capital firms have enough foresight to consider more than their financial returns. After all, it is they who are funding commercial applications of man-machine technologies that might have potentially catastrophic side effects on humanity. In my view, the issue of how man and machine will inter-relate in the future should not be viewed from a profit-only perspective. Machines don't have ethics and neither does money. The coming combination of these forces that operate beyond and above human values strikes me as even more dangerous.

Some futurist colleagues predict that we will soon reach a point where the capacity of **thinking machines** will exceed that of the human brain; a point that Ray Kurzweil, scientist and author of *How to Create a Mind*, calls The Singularity, with 2029 as the likely ETA. At this point, if not earlier, even larger and deeply wicked problems will emerge. For example, if we maintain that technology does not (and will not) have ethics, it would probably be downright stupid for anyone to expect that any current or future software program, machine, or robot would be able to act based on human morals, values, or ethics. Thus, the **morals of machines** will emerge as a major factor in the future of humanity, and the issues around what I call **Digital Ethics** (see below) will quickly become more essential as technology spirals into the future.

Every algorithm will need a “humarithm”

I coined the **humarithm** neologism in 2012 – as a wordplay that riffs off algorithms – because I believe that the chains of logic, formulas and **if this then that** rules urgently need to be paralleled with corresponding systems of ethics, values and assumptions, and new if we believe this we must do that rules. I believe that every time we offload a task to an algorithm (a machine) we will also need to think about what kind of humarithm we need to offset the side-effects, i.e. how to best deal with the unintended consequences which are certain to arise.

For example, we may eventually come to the conclusion that commercial airliners can indeed be better piloted by software and robots than by human beings; most research already indicates that this is indeed the case. But if so, we must certainly think about how the passengers will feel about traveling inside a large metal tube that is steered entirely by a robot. This may well be a typical case of where efficiency should not trump humanity.

Who’s serving who? The trap of machine-thinking

In my view, the issue is less likely to be the cookie-cutter, dystopian Hollywood plot that we have watched dozens of times, i.e. the elimination of humanity by AIs. The much bigger concern is that we as humans, might soon be forced to effectively behave more like – or even become – machines in order to remain productive or useful in a machine-age economy. Just imagine a world where you simply cannot compete or even keep up without some kind of wearable augmented reality (AR) or virtual reality (VR) device, or without an implant, or other mental or physical augmentations. Given that many of us are already utterly dependent on our mobile devices, and often feel alone or incomplete without them, these scenarios may become reality a lot faster than we think. Which university professor would not want to have the world’s knowledge available instantly in the lecture theater using a Wikipedia-app controlled via a contact lens or an unobtrusive brain-computer-interface (BCI)? Which doctor would not want IBM’s Watson Analytics VR-display to provide him with real-time medical information and thereby protect him from malpractice lawsuits?

FUTURESCAPES

Once these technologies are cheap, easy to use, and ubiquitous their utter convenience will be extremely tempting.

The real question for now is probably not if and when the machines will attempt to control, replace, or even eliminate us. The more fundamental and timely concern is whether and how we can remain truly human in a world that is quickly becoming a kind of global-brain-machine. A machine comprised of super-intelligent software, hardware, and processes, with human traits being increasingly removed from the equation because they are simply slowing things down too much. Imagine a world without serendipity, boredom, mistakes, mystery, and surprise. A world in which everything has become efficient, optimized, hyper-connected, intelligent, and real-time? In that world, what will happen to us humans, the limited wetware, the eight to nine billion people who may inhabit the planet in the next 20 years?

If this strikes you as a wicked problem, consider that this gigantic man-machine operating system (OS) might in fact be what some of the leading global technology companies are already striving for. LinkedIn is busy building a global **economic graph**; a kind of proprietary OS for work, jobs, and HR that uses Big Data and AI to predict hiring trends and training needs (among many other things). Facebook already has its own global OS for social and commercial relationships, and Google has its Global Brain project (amid its numerous investments in AI, robotics and deep learning companies). Clearly, the future is already here.

The future of work and jobs: moving towards the right-brain

We now have to face the distinct possibility that as machines rapidly become more capable of doing what we used to do, particularly with our left brains, we will probably need to become more human and increasingly less like machines. Ironically, this is completely the opposite of what traditional MBAs looked like, avoiding emotions, limiting imagination, and sticking to schedules and plans. If you believe that non-algorithmic i.e. emotional or subconscious factors such as trust, purpose, ethics, and values will remain at the core

of human societies in the foreseeable future, this will clearly put a much stronger emphasis on the right brain. Education, training and learning will be changed forever as a consequence and we are already seeing the tip of that iceberg emerging.

What if – in the near future – many routine business activities or operations are actually handled by algorithms and intelligent agents acting on our behalf? In sectors such as procurement, logistics, or telecommunications network management it might well be possible to have intelligent, self-learning software, and reasoning robots take care of 50-100 percent of the frequent and repetitive tasks, within a decade from today. This would obviously lead to huge increases in efficiency and potentially massive costs savings, bringing much lower prices for consumers but also a crushing commoditization for those companies (and people) that currently provide these services. Clearly, human operators cannot and should not compete here – the only way for us is to move up the food-chain, i.e. above the API (application program interface).

Who will have stewardship and control?

Some urgent questions arise as we enter the age of man-machine convergence: who will actually have stewardship of these issues? Who is in charge of what is ignored, allowed, or sanctioned? Should it be trusted to the likes of Defense Advanced Research Projects Agency (DARPA) or the United Nations? If this is not just about technology and business but also about ethics, values, and culture, who would have authority over these matters? What will happen to our collective, cultural and social concerns, i.e. those beyond the commercial agenda? How will our social contracts change because of this, and will these achievements make us happier?

The challenges of unintended consequences

In my view, unintended consequences of exponential technological progress are by far the biggest challenge that we will need to tackle in this coming age of smart machines – and hopefully not *Our Final Invention* as James Barrat presents in his brilliant 2013 book³⁷.

FUTURESCAPES

Exponential technological advancements are certain to have a myriad of unintended yet even inter-related and combinatorial consequences. In many cases I believe, these must be considered more seriously before we proceed.

Some will prove to be rather harmless and more easily remedied such as using smartphones while driving becoming an increasing cause of accidents. Others may have potentially catastrophic outcomes – such as AIs that could learn how to fix and augment themselves, leading to a so-called **AI explosion** and **superintelligence** that could spell the end of humanity as we know it (again, as Hollywood likes to depict so deftly).

Drones also make for a very good example here. There is certainly some logic in augmenting or even replacing postal and delivery services with drones in urban areas or even in places that lack infrastructure such as in Africa. If we were actually to pursue this however, we are certain to face a slew of unintended consequences which may well void most benefits we may otherwise derive from it. Consider issues such as these drones providing the perfect means for real-time surveillance, or the likelihood of citizens acquiring weapons or other means of disabling those drones that have become a nuisance to them.

Such wicked problems may well become the default in the very near future. How do we harness the positive outcomes of these new technologies without creating monsters on the flip side?

Opportunities and challenges driven by abundance

Through exponential technological progress we will soon reach new levels of man-machine relationships. This clearly has the potential to solve many challenges that are subject to bold scientific endeavors – such as energy, food, water, and the environment. Connecting everyone and everything (the Internet of Everything) will generate very powerful network effects that – among many other things – will allow us to further perfect crowd-sourcing, crowd-funding, and crowd/peer-creation. While this in itself is hardly a panacea either, it does enable business models that were impossible before, generating increased abundance at an even faster pace and quickly

challenging our economic logic to the core – as we are already seeing in the debates about Uber and AirBnB. What would be the purpose of increasing consumption if almost everything is abundant? When the price of most goods goes towards zero because they can be reproduced instantly, why do I need to work for a living, and what would my money still buy me?

On the flip side, the challenges of actually reaching abundance driven by a highly evolved **man-machine OS** will be numerous. For example, we will very likely see potentially dramatic job losses – technological unemployment – on a global scale and especially in the BRICs/CIVETS. This could result in social unrest, increased crime, and terrorism born out of sheer hopelessness. There is also the quite real threat of creating a truly perfect real-time surveillance network where nothing, not even your thoughts, would remain private. Finally, there is the dramatic and global rise of **machine thinking** (see above). This is accompanied by an increasingly popular mindset that treats human idiosyncrasies like story-telling, mystery, boredom, contemplation, and imagination as mostly wasteful and inefficient, and wishes to make an algorithm out of everything. The combined effect of these challenges could certainly be considered a kind of hell that would rival George Orwell's worst fears.

Digital ethics are becoming crucial as man and machine converge

To be ready for this coming age of intelligent machines and increasing man-machine convergence I believe we urgently need to start debating and crafting a global **Digital Ethics Treaty**. This would delineate what is and what is not acceptable under different circumstances and conditions, and specify who would be in charge of monitoring digressions and aberrations. No small feat clearly, but maybe the process and the result could be similar to the guidelines that came out of the 1975 Asilomar Conference on Recombinant DNA – a framework that seems to have guided the development of biotechnology deftly and effectively for the last 35 years.

I believe a **Digital Ethics Treaty** will soon prove as important as

FUTURESCAPES

the nuclear non-proliferation treaties (NPT) that are already in place, and that have indeed proven to be enforceable (if not entirely without friction). So here are some admittedly still fairly raw rules I would like to propose for inclusion in such a treaty:

- We should not allow humans to actually become technology (in the sense of fundamental augmentation of the human body or mind).
- We should not allow humans to be effectively governed by intelligent technologies.
- We should not allow the fundamental altering of human nature and the **manufacturing** of new creatures with the help of technology (such as large scale genetic manipulation).
- We should not allow robots and intelligent machines to upgrade, fix or alter themselves.
- We should not allow the open or inadvertent discrimination of humans that chose not to use technology to increase their efficiency or competitiveness.
- We should not require or allow robots to make ethical decisions, i.e. to become sentient or develop some kind of moral agency.

“The best way to predict the future is to create it”, (to quote Peter Drucker, Abraham Lincoln and Alan Kay).

The bottom line is that if you are running a business or an organization today, you will probably encounter these man-machine-convergence challenges very soon, or maybe it will be more of a **gradually then suddenly** event for you.

Either way, the future of how we relate to and intertwine with machines is being defined at this very moment, and it raises some fundamental questions:

- *How are you shaping the debate about the future direction of your business – are you building towards the human future or the machine future?*
- *Heaven or hell – how should society seek to shape evolution and*

ethical governance of technological innovation and the boundaries between man and machine?

- *What sort of organizational structures, strategies and business models will we need to survive and thrive in a world of abundance, declining prices and high technological unemployment?*