Chapter 4
What to Expect From Artificial Intelligence in Business:
How Wise Board Members Can and Should Facilitate Human-AI Collaboration

Peter Verhezen
Antwerp Management School, Belgium

ABSTRACT

We are increasingly living in a digital world, where companies attempt to adapt to a new context of Industry 4.0. The authors believe that artificial intelligence and the use of logarithms will alter the game of competition. Digitization is moving our economy away from "financial capitalism" to "data capitalism," and companies and their boards need to adopt the way they operate and steer the organization to new ecosystems where personalized service becomes part of the new digital strategy. Basically, it is not a battle of AI versus humans, but rather finding a way to enhance the collaboration of AI and humans in organizations. Despite the enormous potential benefits of AI, boards should not ignore the darker side of AI, namely the potential biasedness and sometimes unfairness of algorithms and privacy concerns and the ubiquitous cyberthreats. This is why proper data governance at the board level is needed. The authors suggest that this becomes a critical success factor to be addressed at boards, either as part of the risk management or strategic committee or as a separated digitization committee.

INTRODUCTION

We are increasingly living in a digital world, where companies attempt to adapt to a new context of Industry 4.0. We believe that Artificial Intelligence and the use of logarithms will alter the game of competition. In such a changed environment where digitization is moving our economy away from "Financial Capitalism" to "Data Capitalism", companies and their boards need to adopt the way they operate and steer the organization to new eco-system where personalized service becomes part of the new digital strategy. Basically, it is not a battle of AI versus humans, but rather finding a way to enhance

DOI: 10.4018/978-1-7998-2011-6.ch004

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

the collaboration of AI and humans in organizations. However, despite the enormous potential benefits of AI, boards should not ignore the darker side of AI, namely the potential biasedness and sometimes unfairness of algorithms and privacy concerns, and the ubiquitous cyberthreats. Hence why proper data governance at the board level is needed. We may suggest that this becomes a critical success factor to be addressed at boards, either as part of the Risk Management or Strategic Committee or as a separated Subcommittee – for instance Digitization Committee.

BACKGROUND

Today, a majority of people are communicating via social media leaving digital traces which provides the "new oil" - or data – used by organizations that claim to facilitate the quality of our daily life. Indeed, our world is being dramatically influenced and driven by big data. In 2000, about 25% of all data were digitized, about 18 years later, 97% of all data are digitized in one form or another. In the future, data-rich markets will offer individual choices without the constraints of inescapable cognitive limitations.

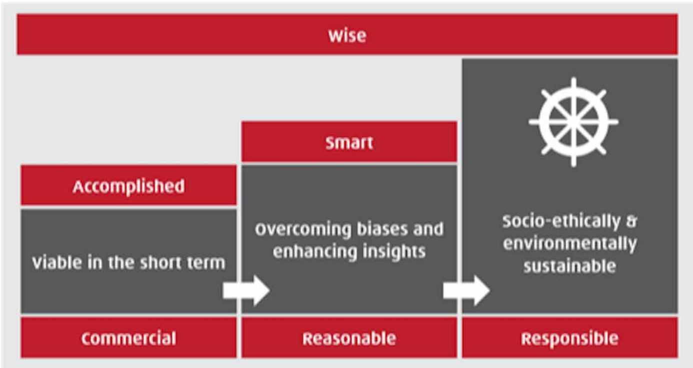
A majority of people in the developed but increasingly also in emerging markets now communicate via the prevailing channels of social media – be it Facebook, WhatsApp, Instagram, Snapchat, Twitter, LinkedIn to name a few, which are all powered by digital data and algorithms. In the future, data-rich markets will offer individual choices without the constraints of inescapable cognitive limitations. Indeed, data are fast combining the new oil. We could easily argue that we are moving from a *finance capitalist system* to a form of *data capitalism*, facilitated by the growing internet traffic or network effect, massive data sets, and the enhanced processing data capacity or analytical power of computer. One of the consequence of this data capitalism lies in the curious shift from causation – as scientists have looked for through appropriate statistical methodology – to correlation where data "speak for themselves", without necessarily understanding the why behind the correlation. Current (narrow) AI applications are not able to generate causal relationships, as UCLA computer scientist and mathematician Judea Pearl argues (2019). AI today remains an algorithmitization of supervised data, allowing to recognize (sometimes complex and ambiguous) patterns in a speedy and efficient manner, whereby this non-causal analysis and mere correlations are often fast and reasonable cheap. But this narrow artificial intelligence cannot answer the why behind any mathematical equation, and AI cannot reflect yet how humans think in terms of general purpose. The human brain observes and continuously searches for causal links from limited data, based on a certain assumed model – which is quite different of how AI algorithms operate today.

But what is Artificial Intelligence (AI)? Is AI the branch of computer science that is concerned with the automation of intelligent behavior? We all can agree in principle that **intelligence** is the ability to deploy novel means to attain a goal – which are extraneous to the intelligence. In other words, intelligence is the ability to accomplish complex goals. The distinction between specialized intelligence and general intelligence helps clarify the difference between the specialized abilities of today's learning machines or (narrow) AI and humans' more general abilities. **Artificial Intelligence** is the overarching science that is concerned with intelligent algorithms, whether or not they learn from data. The adjective *Artificial* refers to intelligent machines that fullfil certain objectives, and more particularly, machines that learn to apply statistical techniques to supervised data that allows them to find patterns in complex big data in a much faster and intelligent manner than humans. This obviously lead to a number of advantages when these data analytics are applied to our current life as in voice or image recognition. **Machine learning** is a subfield of AI devoted to algorithms that learn from data (Husain, 2015; Finlay, 2018). Artificial intelligent machines are “smarter” – faster and better – than humans in terms of certain kinds of specialized knowledge or intelligence. But it remains very specialized knowledge about a certain specific domain and AI cannot compete with humans in terms of general purpose thinking and creative processes for which human executives are so sought after (Brockman, 2019; Ford, 2018).

IBM’s Watson may have beaten the human champion at Jeopardy, but it can’t play any chess game at all. A Tesla car may (sort of) drive autonomously, but that car cannot autonomously pick up a box at the near Carrefour center. Building some form of intelligence out of pure information technology based on digital data without human components is what people normally refer as “artificial intelligence”. IT researchers Davenport and Ronanki (2018) distinguish 3 types of ***Artificial*** Intelligence: (1) *Process Automation* with robotics and cognitive automation (counting today for about 52% of the “AI-applications”), (2) *Cognitive Insights* that allows better predictions which will be our focus (roughly worth 35% of today AI-applications), and (3) *Cognitive Engagement* as in natural language processing chatbots and intelligent assistants (mounting to about 13% of the total AI practices). AI can help is in improving our innovative insights and reduce our biases in our presumed rational thinking, but AI improves our cognitive prediction abilities. “Artificial intelligence” does not actually bring us intelligence but instead a critical component of intelligence, namely *prediction* in a cost-efficient manner (i.e. AI will make prediction cheap). Prediction – i.e. “if-then” logical intelligence - is the process of filling in missing information. Prediction take digital information you have, data, and use it to generate information you don’t have. For instance, statistical regression minimizes prediction mistakes on average which is a powerful tool with relatively small data sets. With the vast amount of personal data, however, digital devices will learn more about us than we may know about ourselves, enabling us “to better predict” and thus allowing us to make better and smarter decisions by reducing uncertainty. Not surprisingly, Google’s CEO Sundar Pichai sees a shift in Google from “searching and organizing the world’s information to AI and machine learning”.

As the cost of prediction continues to drop, we’ll use more of it for traditional prediction problems such as inventory management because we can predict faster, cheaper, and better (Agrawal, 2018). At the same time, firms start using prediction to solve problems that have not historically been thought of as prediction problems. For example, we never thought of autonomous driving as a prediction problem. Traditionally, engineers programmed an autonomous vehicle to move around in a controlled environment, such as a factory or warehouse, by telling it what to do in certain situations—*if* a human walks in front of the vehicle (*then* stop) or *if* a shelf is empty (*then* move to the next shelf). But we could never put those vehicles on a city street because there are too many *ifs*— *if* it’s dark, *if* it’s rainy, *if* a child runs into the street, *if* an oncoming vehicle has its blinker on. No matter how many lines of code we write, we couldn’t cover all the potential *ifs*. Today we can reframe autonomous driving as a prediction problem. Then an AI simply needs to predict the answer to one question: What would a good human driver do? There are a limited set of actions we can take when driving (“*thens*”). We can turn right or left, brake or accelerate—that’s it. So, to teach an AI to drive, we put a human in a vehicle and tell the human to drive while the AI is figuratively sitting beside the human watching. Since the AI doesn’t have eyes and ears like we do, we give it cameras, radar, and light detection and ranging (LIDAR). In a way, machine learning does induce theories – a set of constraints on what the world would be, not necessarily a complete description of it - from data. The AI takes the input data as it comes in through its “eyes” and looks over to the human and tries to predict, “What will the human do next?” The AI makes a lot of mistakes at first. But it learns from its mistakes and updates its model every time it incorrectly predicts an action the human will take. Its predictions start getting better and better until it becomes so good at predicting what a human would do that we don’t need the human to do it anymore. The AI can perform the action itself as long as the infrastructure is available

Figure 1. Wise Decision-Making
Source: Verhezen, P., (2018): Wising Up enable Managers to be commercially savvy, to make reasonable smart decisions, but also to commit to responsible behavior that distinguishes smart from wise leadership:



which means that high-speed connectivity is continuous and of good quality - requiring a 5G network. Businesses are expecting that we are close to a tipping point in materializing this idea of automated cars in geofenced situations and for ships. In non-geofenced circumstances autonomous driving – e.g. non-rectangular logically built cities as you often find in Europe – autonomous driving may face a number of serious technical challenges that may delay full autonomous driving for a while.

But this progress of better "prediction" leaves us with a crucial question: who does own the data and how can these personal data be used by organizations? Indeed, we should acknowledge the fact that today, more data means less privacy; more speed means less accuracy, and more autonomy means less control. More expected accuracy means less *explainability*. There are trade-offs to be decided upon. And let us not forget the simple heuristics that if "data or the application is free, then you are the product". The real customers are those who are willing to pay for access to knowledge about us (derived from these data), so that they can persuade us to purchase a product or influence us. Moreover, "datafication" is not value-neutral either – as we will argue – and using data still requires theory building on which they depend.

The power of this "dataficated" reality, be it biotech, nanotechnology, robotics, cyber-technology or Artificial Intelligence systems will have a huge impact on everyday life; it records all our movements, human interactions and financial transactions, all stored in the 'cloud'. This new reality within *Industry 4.0* will facilitate our life and possibly enhances it. This transactional improved efficiency will likely result in more sustainable solutions. But at the down side, this *Industry 4.0* will also be managed and possibly manipulated by a few multinational quasi-monopolies, and in worst scenario, the global inter-connectivity could also result in systemic failure.

Businesses need competent and smart employees and managers; increasingly, business is in need of wise leadership to deal with the ambiguities of the enormous global challenges. We believe in the enormous benefits of digital innovative technologies, in particular AI from which deep learning machines is a well-known example. However, wise leadership will need to find an integrated balance between these undeniable positive opportunities by AI and the darker side of AI.

(cf. http://www.verhezen.net/images/papers/VERHEZEN_2018_Amrop_AI_and_Wise_Leadership_181010.pdf *)*

Indeed, we believe that wise decision-making is only possible when managers make smarter decisions, by reducing their biases and potential errors on the one hand, while probing creativity and enhancing environments where innovation can thrive. Our belief is that Deep Learning Machines and Artificial Intelligence as in improving cognitive insights can help humans to make better decisions by reducing errors and by enhancing innovative tools and economic efficiency. Uncertainty, errors, and human biases (i.e. human irrationality) could potentially be significantly reduced by having artificial intelligent machines using the availability of big data to improve predictions or to find for the human eye or brain hidden patterns. Similarly, the tools – be it digital automation or cognitive engagement - that originate from artificial intelligence can be easily incorporated in innovative products and services or employed in robotics to take over repetitive boring or extremely dangerous work.

However, managers and corporate leaders cannot ignore the potential negative implications of artificial intelligence that may jeopardize quite a number of jobs and employment (London, Cui & Whigley, 2019). The social implications could be enormous in certain industries and areas. Moreover, the ethical implications of smart machines could be devastating if not handled properly (Burkhardt, Hohn, Wigley, 2019). At the other hand, the non-financial objectives and intangibles as expressed in the notion of ESG (Environmental, Social-Ethical and Governance objectives) can be improved and assisted by collaborative efforts of smart engineers and managers with artificial intelligence that significantly improve decision-making, taking into account certain stakes beyond mere profitability that are more socio-ethical and ecological sound. Our organizations could benefit from using AI applications in the decision-making process to become more responsible. Having a better understanding of what could happen in the future, based on the power of predictive AI applications, could help to nudge decisions in a certain more sustainable direction. It could inform executives to make a better selection of scarce raw materials, or extend the lifespan of electronics through predictive maintenance, or automate and improve e-waste recycling infrastructure through the combination of image recognition and robotics. Researchers believe that artificial intelligence tools could contribute to a more circular economy by taking advantage of abundant data in 3 major ways: (1) design of circular products, components and materials whereby iterative machine-learning-assisted design processes allow for rapid prototyping and testing, (2) operating of circular business models – product-as-a-service and leasing that increase product circulation and asset utilization through better pricing and demand prediction, predictive maintenance and smart inventory management, and (3) optimizing circular infrastructure where AI can help build and improve the reverse logistics infrastructure required to close the loop on products and materials by improving processes to sort and disassemble products, remanufacture components and recycle materials (Schouteden & Sturges, 2018). Indeed, AI could assist human decision-makers with better models and scenarios for climate change or weather forecasts. In case of natural disasters, AI could assist rescue workers to enhance the [efficiency of] disaster recovery efforts, by mapping and analyzing digitized struck areas, and possibly come up with better or faster rescue solutions.

The use of AI does not necessarily need to result in devastating unemployment as long as wise leaders foresee and prepare themselves and the community for a world where both humans and smart and extremely fast machines will be able to collaborate. And last but not least, corporate governance of organizations embracing big data will require proper data governance and guarantees that negative outcomes with the use of AI – be it cyber breaches or unethical decisions or privacy concerns – will need to be limited.

We here will briefly explain some of the potential benefits of artificial intelligence, after which we also focus on the darker side of deep learning machines powered by AI. Good AI should be distinguished from bad AI. Finally, we will conclude by indicating how smart human decision-makers could collaborate with AI, and why wise leadership and wise boards are needed to monitor and direct the organization towards such beneficial and appropriate collaboration, while not ignoring and preventing or minimize the darker side of AI to materialize.

The Potential Benefits of Artificial Intelligence

Just consider the GPS lead systems that are behind guiding autonomous driving cars or virtual assistants like *Apple's Siri* and *Amazon's Echo*. These tools are reflections of the way we have been thinking and talking. Artificial Intelligence, today, is functioning by brute force using millions of samples, using reinforcement learning based on little pieces to approximate a desired function. *Vanguard*, an investment service company, uses cognitive technology (or a form of AI tool) to provide customers with investment advice at a lower cost than their competitors. Its Personal Advisor Services system automates may traditional tasks of investment, while human advisors focus on higher-value activities.

A majority of people in the developed but increasingly also in emerging markets now communicate via the prevailing channels of social media. Facebook, Instagram, Snapchat, Twitter, LinkedIn to name a few, are all powered by digital data and algorithms. *Amazon*, for instance, looks for unique patterns in the data they receive from customers which reveal the preferences of these customers. Identifying such patterns enables Amazon to statistically deduce customers' wants and needs without having to ask them directly. The data approximately tell what you and I want. The data do not know why we prefer one thing over another; it just "sees" that we choose one over the other, indicating some hidden patterns of our preferences. But that is sufficient for Amazon to feed the preferences-matching algorithm and search for the products [potential] customers are most likely to purchase. Artificial Intelligence seems to be a functional new tool that provides intelligence unavailable before. Or does it? Could it also be that Artificial Intelligence is only the newest fad in town: every organization is eager to put forward the enormous benefits of this newest digital technology, without really knowing how to turn this abundance of data in real changing business models. However, Amazon seems to know what it is doing: it just opened to the public a physical grocery store without check-out lanes or cashiers. You'll fill up your bag, walk out the front door, and get a receipt minutes later for everything that is in your bag: no human interaction is involved in the transaction.

Quite a number of companies are already taking advantage of big data and its prediction power. *Aviva*, a private global insurance company, is now able to predict insurance claims not based anymore on a detailed report of the health of its subscribers - who may have given blood and urine examples costing the company USD 125 per person for the analysis – but on credit reports and consumer marketing data which cost only USD 5 per person on average. Those data on the life style of people taking an insurance now function as a proxy to predict the health of these customers.

Banks can detect credit card fraud by looking at anomalies, and the best way to find them is to crunch all the data – big data - rather than a sample. The card network uses information or data about past fraudulent (and nonfraudulent) transactions to predict whether a particular recent transaction is anormal and possibly fraudulent, and preventing actual fraud or future illegal transactions.

When a mall operator uses advanced analytics to select tenants, optimize mall layout, and determine rents, its revenues can rise by 20 percent, according to a McKinsey survey. With its ability to enable personalization and customization at scale, AI can be a powerful differentiator for consumer-facing businesses. It improves precision and speed-to-market, and increases and enhances the potential for interactions, engagement, and transactions. Companies such as *Spotify* playlists, the *Facebook* newsfeed combine human and computer expertise to create new services and enable people to discover and engage with content and brands in new ways.

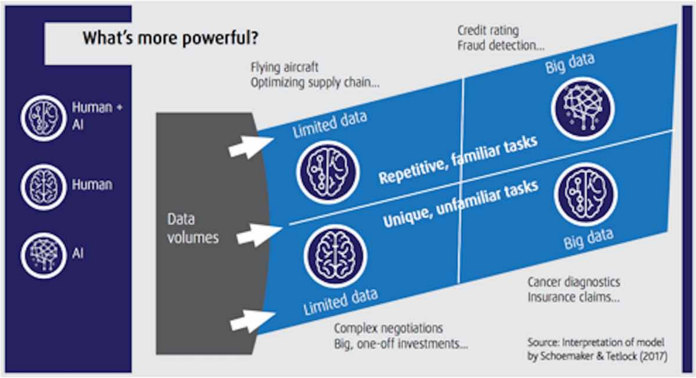
A few organizations use AI for personalization better than the listed fashion online organization *Stitch Fix* or the movie streaming giant *Netflix*. These organizations basically personalize their offers to individual customers by applying a sophisticated algorithm that is using continuous conditional probability calculations - which is the chance that one thing happens, given that some other thing has already happened. Conditional probability is how AI systems express judgments in a way that reflects their partial knowledge. And personalization algorithms run on conditional probabilities, all of which must be estimated from big data sets in which you as an individual are the conditioning event. Real problems are framed in terms of conditional probability (if-then logic) to solve them. Computers do not understand why you are watching a particular movie, but they are great at crunching data, i.e. tabulating vast databases of subscribers' movie-watching histories from a ratings matrix to estimate conditional probabilities of individual movies' preferences - discovered organically by AI. In other words, digital economy is about suggestions and thus conditional probability, translated into prediction, rather than search.

In AI, recognizing a pattern means fitting an equation to data through *supervised learning* (i.e. the algorithm detects interesting features in data where the end game is a predictive model of some sort). The big breakthrough in AI was the introduction of the use of neural networks for estimating prediction rules from data. The misleading notion of "neural network" is a complicated equation with a lot of parameters that is capable of describing very complicated patterns in data. These neural networks work incredibly well across a range of prediction tasks, from language to images to video. However, despite the initial excitement for AI using *reinforced learning*[1] – because this kind of learning resembles more how we humans think through trial and error – the number of successful reinforcement learning based solutions in (narrow) AI is tiny compared to the number of supervised learning ones - over 95% of the AI cases (Finely, 2018).

Other examples are governments that understand having sensors affixed to bridges and buildings to watch for signs of wear and tear could and is preventing potential disasters to occur. The cost of collecting and analyzing the data that indicate when to take early action is lower than the cost of an outage. Note that these predictive analytics may not explain the cause of the problem (the why), it may only indicate that a problem exists (the what). *General Electric* and *Rolls Royce* both have implemented big data analytics into their commercial jet engine business to predict more accurately when to replace expensive parts or when to optimally start maintenance of the jet engines, allowing those firms to apply new business models by leasing or renting power to the airplanes instead of merely selling engines.

Indeed, computers have significantly improved at image and voice recognition and speech synthesis. Computers can now detect tumors in radiographs earlier than most humans. Medical diagnosis and personalized medicine will improve substantially. Transportation by self-driving cars – where transportation is transformed into a prediction problem - will keep us safer, on average. And hopefully, we can sort out the ethical challenges regarding the use of this new kind of digitized intelligence.

Figure 2. Collaboration between humans and machines
Source: Interpretation by Verhezen, P. (2019) of model by Schoemaker & Tetlock (2017)



Now that everything becomes increasingly "datafied", it looks like we can measure most aspects of human life, and with those powerful machine-learning techniques, we now can build ecosystems (Lohr, 2015). Well-known examples are weather- and traffic-prediction models which are being extended to predict the global climate and plan city growth and renewal. MIT professor Sandy Pentland and his team for instance, researches to how human behavior and ecosystems interact; this "social physics" looks at patterns of cultural behavior and develops mathematically accurate predictions how people make decisions. At present, organizations are trying to influence conscious processes and explicit knowledge. Yet, Pentland's research indicates that sociometric data show that unconscious processes and tacit knowledge are potentially even more important in determining the behavior of organizations. Let us summarize our understanding of using AI to make smarter decisions so far:

The progress in AI will in many cases be exponential rather than linear. Already the progress in a wide range of applications (e.g., vision, natural language, motion control) over the last 12 months was faster than in the 12 months prior, according to McKinsey and the Boston Consulting Group. The level of investment is increasing rapidly. The quality- adjusted cost of sensors is falling exponentially. And the amount of data being generated is increasing exponentially. In most cases, when AIs are properly designed and deployed, they're better predictors than humans are. And yet we're often still reluctant to hand over the reins of prediction to machines. For example, there have been studies comparing human recruiters to AI-powered recruiters that predict which candidates will perform best in a job. When performance was measured 12, 18, and 24 months later, the recruits selected by the AI outperformed those selected by the human recruiters, on average. Despite this evidence, human recruiters still often override the recommendations provided by AI system when making real hiring decisions. Where Artificial Intelligence have demonstrated superior performance in prediction, companies must carefully consider the conditions under which to empower humans to exercise their discretion to override the AI. The organizations that will benefit most from AI will be the ones that are able to most clearly and accurately specify their objectives. Remember that today AI is at its best within very specific domain of expertise, assistance or automation. We're going to see a lot of the currently fuzzy mission statements become much clearer. The companies that are able to sharpen their visions the most will reap the most benefits from AI. Due to the methods used to train AIs, AI effectiveness is directly tied to goal-specification clarity. What makes AI so powerful is its ability to learn. Normally we think of labour as being learners and of capital as being fixed. Now, with AI, we have capital that learns. Companies need to ensure that information flows into decisions, they follow decisions to an outcome, and then they learn from the outcome and feed that learning back into the system. Managing the learning loop will be more valuable than ever before.

The way one can make individuals and groups smarter, the way one can make a more “humanized AI”, will work only if feedback is truthful. In other words, data must be grounded on truth. However, manipulative advertising, propaganda and “fake news” destroy the usefulness of social sampling and data in general. We need data that we all can trust and we also need fair, data-driven assessment of public norms, policy and government based on trusted data about current conditions. Only under those circumstances, the individual and especially societies’ overall fitness and intelligence can improve or can be “trusted”.

Obviously, data-driven markets offer compelling advantages, and innovation and progress should not be stifled by irrational emotional fears or too stringent regulations. But the shortcomings and ethical challenges should not be ignored, especially the concentration of data and the possible systemic failure. And in case real artificial general intelligence would become a reality, a “matrix”-like intelligence, we really should be concerned about the malicious consequences of such super powerful machine-related intelligence. What interests us here is the importance of transparency of information and its algorithms to reduce potential information asymmetry. In other words, can data governance control artificial intelligence?

The Darker Side of Artificial Intelligence

We can assume that in most cases, the aim to use data and algorithms can be a force of progress and good use. Algorithms are used to help us better understand the world. Algorithms underlying artificial intelligence are only as good as the big data input. Indeed, algorithms are programmed to collect and categorize a vast amount of data points in order to identify patterns in a user’s online behavior that could allow recommendations and more precise predictions. The algorithmic identity in any application gets more complex with every social media interaction, the clicking or likewise ignoring of advertisements, and the financial status as derived from online payments. Huge amounts of digitized data are now available. The more people share their personal information and preferences on social media – and people feel empowered to do so -, smart entrepreneurs will definitely take advantage and initiate new algorithms that embrace the enormous amount of data in cyber-space, and commercialize them in one form or another, or in best scenario initiate new insightful patterns that could help common good. Would it be too farsighted to put the argument forward that the intentional accumulation of “like it” on *Facebook* is nothing else as the product of an algorithm that rewards attention-seeking and shock value? What could undermine the benefits of AI? However, some entrepreneurs like Elon Musk, the CEO of *Tesla* and *SpaceX*, and the late Cambridge physicist Stephen Hawking claim that general AI and smart machines may become “our biggest existential threat” as a species. We discuss a few of those negative issues that is mostly inherent to big data analytics.

Biasedness

Facebook's algorithm decides what information to show us on the basis of the choices we already have made. This filter algorithm used by Facebook may create a filter bubble or echo chambers, even for initially unbiased people. The filter model picks up small initial differences and exaggerates them until the other side of the argument is lost. And we do not mention even the spreading of untrue rumors that become fake news which has become a source of constant entertainment. As in a kind of post-truth world.

The popular *Tinder* application for instance uses algorithms to romantically link people together is an example of what can be described as "amplified biasedness" by the machine learning. Tinder is one of the fastest growing social networking apps on a global scale with users in 190 countries swiping 1,6 billion pictures and generating around 20 billion matches every day. This location-based dating application plays a game-changing role in the dating world. However, we should not ignore how the biases of Tinder algorithms is a reflection of our society and how we analyze and perceive humans. Despite the personal swiping choices we make in finding a romantic partner, this online dating application seems to be reinforcing racial prejudices. Depending on how an algorithm is programmed – and Tinder's "magic" black box is not revealing how it functions -, the users' online behavior and the set of data it is given to process the intended matching process, certain cultural aspects will be highlighted, visualized and prioritized while others are left out or rendered invisible. This kind of algorithms are not value-free and reflect the cultural and individual preferences and human biases as in a darker shadow, not exactly expected from a cold presumably objectively calculating machine.

It should not surprise us that the specific workings of algorithms remain rather elusive, as developers and data scientists rarely provide the coding of the underlying programs in the name of technological neutrality and objectivity and in order to preclude unnecessary competition. But we can derive some basic features of the Tinder application: since each user expresses individual preferences, the system provides personalized recommendation which are obtained through collaborative filtering and algorithmic calculations. Tinder's "algorithm of desire" all boils down to ranking people according to their desirability – based on "skill levels". Nobody wants to be rejected. Tinder complied with this psychological insight by keeping the left swipes unknown to the users. Similarly, the right swipes are kept secret as well, and sometimes matches are not shown to slow down the very desirable people – the "winner should not take all" [desirable…] - to give people with lower raking a chance, and thus keep Tinder in the game. Not exactly neutral, is it. But psychologically, you could describe it as smart but not necessarily wise business thinking.

Humans are now constantly bombarded with personalized recommendations based on our online behavior and data sharing on social networks as *Facebook, Twitter, Amazon, Spotify* and *Netflix*. Machine-learning algorithms paired with AI is designed to develop in a manner – attempting in mimicking the human process of learning (seeing, remembering and creating a pattern in one's mind) – that allows Tinder's AI-paired algorithm for instance to develop its own point of view on people. The AI system does not know why it is recommending a particular match, but has strategically learned to develop a "thinking" (i.e. finding correlating patterns) that could resemble human intuition. The system identifies languages and words that share a common context which could potentially indicate similarities, potentially resulting in swipes that are clustered together reflecting perceived preferences through these embedded vectors of the participants' likes. Unfortunately, such algorithms also reflect the darker side of our culture: embedded biasedness. Apparently, studies reveal that Black women and Asian men are potentially marginalized and possibly discriminated in such online dating environments. With all the dare consequences. If initially several Caucasian matches were “successful” for instance, the algorithm will continue on the same biased trajectory. Confirming a “statistical commonality” according to gender, class or race in supplying a meaning for those categories, will be “learned, analyzed” and conceptualized by the algorithm. Not exactly the most neutral manner to advertise your “assets”. Admittedly, the data points remain hidden in the black box and cannot be overridden by any external critical remark or research, but it reinforces our suspicion against presumed cold speedy machines and its algorithm that are advertised to be neutral and objective. The opposite is true.

Specific biases in the used criteria and variables in these algorithms are either unexamined or remain unconscious and unaware by the data designers, enhancing our point that we should be worried to blindly trust these algorithms. And here we face a paradox: machine learning AI pretends to be neutral and provide better decision-making options whereas in reality the underlying criteria and variables of these algorithms – often based on detecting personal preferences through behavioral patterns to come up with recommendations - are nothing else as a mirror to our societal practices, potentially even reinforcing existing biases. Indeed, societal biased garbage in, biased garbage out. The game of speedy and more precise predictions is not so objective as being proclaimed by the owners of these apps. Even if we or those owners have the best intentions, those intentions too could be easily (socially or personally) biased.

Unethical Use of Data and Fairness

Datasets could easily be turned into unethical use as *Facebook* has shown during the US elections in 2016. *Cambridge Analytica* used a personality model – the big five personality traits test - in its promotional material to send specific political tailored messages and ads catered to people and undecided voters with a specific profile to influence the US elections. What about Google possible decision to drop “do not evil” in its mission statement mid 2018, allowing the possible road to be opened for a watered-down version of a Google search engine – aptly labelled Dragonfly which blocks or self-censures sensitive topics regarding human rights for instance – to consider a re-entry into China’s huge consumer market. Moreover, artificial intelligence is more like advertising intelligence where big corporations have got better at collecting consumer data, filter and package them and sell them back to these consumers in the form of recommendations.

We can easily argue that the code of algorithms is not value-neutral – it contains many judgments about who we are, who we should become, and how we should live. In case we would be asked to choose a software solution, will we be subtly influenced to buy from a particular online vendor and will we be affected by the vendor’s (subconscious) prescriptive norms and values. What if these values are less than benevolent? The business ethicist Edward Freeman highlights this ethical conundrum by asking what will happen when a self-driving car under certain unfortunate circumstances – where an accident cannot be avoided – will need to make a (algorithmic) decision by making a choice about whether to sacrifice its occupants or risk (possibly fatally) harming passengers in other cars or pedestrians. How to implement robo-ethics or address ethical challenges with respect to AI? “How to guide developers to write this code”, Freeman is asking. Keeping in mind that facts are distinct from values, we can conjecture that from an evolutionary perspective, our genes and memes attempt to survive. If we attempt to bring ethics into machine learning, it leads to a whole series of Trolley problems. At what number of people in line for death should the computer or GPS system decide to shift a moving trolley to one person? It remains an ethical challenge for most humans to make a fair and just decision. Obviously, we could also turn the whole Trolley problem by questioning who has given pedestrians access to the rails. These kind of judgments about moral and ethical choices are just as important as they always have been.

Even if a data set is accurately reflecting historical facts, it does not mean that these data are [ethical] fair, especially if it can proven that history itself was not necessarily fair. We should question whether an algorithm is fair, whether AI is doing things that humans believe are ethical. Bringing ethics into AI, one needs a human-in-the-loop approach as in an "open algorithm", not a black box.

An example of unfair use of data here is "predictive policing or profiling". Researchers like David Sumpter (2018) argue that commercial software that is widely used to predict recidivism is no more accurate or fair than the predictions of people with little to no criminal justice experience. At best, algorithms may match the accuracy of humans in this exercise, but just much faster. So while these models are far from perfect, they can be useful speedy tools. Admittedly, studies by Professor Philip Tetlock from the Wharton School found out that the average "expert" was 'roughly as accurate as a dart-throwing chimpanzee'. However, experts who were able to continuously include new info/data in their probabilistic reasoning, created bell-shaped curves in their head and drastically improved their predictions. But still it is hard to beat the collective wisdom. And here algorithms basically reflect the collective data on which that wisdom could be derived from.

There is not necessarily an equation of fairness, and the normative notion cannot be fully derived from descriptive logic. This brings us to David Hume's is-ought problem whereby there is a deep gap between what is (a scientific objective reality) and what ought to be – i.e. an ethical question of how we want to live and what kind of society do we want. Hume's dichotomy between "is" and "ought" implies that what ought to be cannot be directly derived from what is, and therefore aspiring ideals cannot really be bridged by algorithms. On the contrary, algorithms may reinforce "old ethical habits and norms". Only conscious and mindful humans – who hopefully can be considered wise decision makers – can put forward what kind of society we want to live in, what kind of life we want to strive for. Moreover, research by Ernest Fehr and others with the Ultimatum tests for instance indicate that the notion of fairness seems to be inherent to human thinking. What society we would like, what we consider as 'fair', our aspirations are norms set to make it a better world. Using the factual data of what is in the world does not make it a good "prescriber", only a good "recommendation engine in perpetuity". No change should be expected from such (commercial) thinking, unless we explicitly bring in these values and norms that aspire for a better and different future. Science does not provide the answers to normative [ethical] questions.

The Dictatorship of Data & Paralyzing Privacy

We can easily fall into the trap of the fetish of quantification and data. However, the quality of underlying data can be poor or even biased. It can be mis-analyzed or used in a misleading manner. And worse, data can also fail to capture what it purports to quantify. And consequently, we may attribute a degree of truth to the data which it does not deserve. Many thinkers have argued that creative brilliance does not depend on data[2].

The increasing reliance on data may also lead to the risks of a "tyranny of algorithms" where unelected data scientists and data experts are running the world. The incredible power of *Google*, *Amazon*, *Facebook*, *Apple*, *Microsoft*, *Baidu*, *Alibaba*, *Tencent* and others cannot be overstated. They currently control the data, and thus they control AI. Can we trust these organizations that they do the right thing, always? Not quite. The shadow of Big Brother seems to start looming over the use of our social media.

Moreover, the internet has made tracking easier, cheaper and more useful. However, the internet and big data also threatens our privacy. The *Cambridge Analytica* debacle – which used data from Facebook to influence the US elections and possibly the Brexit vote – shows that through access of personal data, companies and individuals (having access to these data) can influence human's behavior through personalized messages and advertising in a way never seen before. We believe that individuals should own and control access to their personal data, instead of the application providers. Moreover, in non-democratic states, or even in nominally democratic ones, government know things about their citizens that was considered fiction during Orwell's *1984* time. And obviously, the prospect of AI for malicious military purposes remains frightening.

Moreover, research demonstrates that rating just 6 obscure movies (out of the top 500) could identify a Netflix customer 84% of the time; if one knew the date on which a person rated the movies, the accuracy rate apparently increased to an incredible 98%. Hence why in an era of big data, the three core strategies long used to ensure privacy – (1) individual notice and consent, (2) opting out, and (3) anonymization – have lost much of their effectiveness. It is obvious that privacy is under attack from all sides. To what extent should the power of the internet and AI firms be clipped and constrained to secure the privacy of the individual?

Probability and Punishment

Big data threaten to imprison us – perhaps literally – in probabilities. For instances, the use of big data to conduct "predictive policing" may seem to be sensible, but it also stigmatizes certain socio-racial groups further. Using big data analysis to select what streets, groups, and individuals to subject to extra scrutiny, simply because an algorithm pointed them as more likely to commit a crime. For instance, US Homeland security's FAST (future, attribute screening technology) and other foreign Western government agencies try to identify potential terrorists by monitoring individual vital signs, body language and other psychological patterns. If these data and analyses are misused, it can lead not only to discrimination against certain groups but also to "guilt" by association. Punishing people before they do something bad negates the very idea of the presumption of innocence – the principle upon which our legal system as well as our sense of fairness is based. We should acknowledge that thinking bad things is not illegal; doing them is. Guilt is only possible when someone actually committed an (illegal) act or crime, not thought of it, or may have a higher probability to potentially commit a crime because of a specific social-economic background.

Predictive analytics predicated on mechanical objectivity comes at a price. Indeed, in the courtroom, objectivity, trade secrets and judicial transparency may pull in opposite directions. Mechanical objectivity is not the same as ethical thinking. Nor is such objectivity necessary reflecting the essence of scientific thinking or discovery. And such probability thinking also deprives us from a free will and erodes the fundamental notion of human dignity.

What Kind of Intelligence do Smart Computers Have? Is Singularity Near?

The arrival of neural networks made computer even more intelligent. Neural nets, whose basic design was directly inspired by our brain's architecture, have scored some spectacular successes in game playing, and patter recognition. Face recognition in *Apple*'s iPhone X uses neural networks to uniquely identify its owner's face. Convolutional neural networks, a form of regression model used to predict, is used for face and voice recognition. This principal component analysis uses the data to classify people, rather than relying on our preconceptions. Such component analysis and similar mathematical approaches underlie most of the algorithm used to classify behavior. Machine-learning, or self-learning or deep learning attempts mimicking our human brain - learning by itself – enabling the computer device and algorithm to learn "automatically" without us telling them what to do, or what patterns to pay attention to. Computers seem to become smarter by the day. Will AI out-smart us?

Whether AI will make humans subservient or obsolete, or whether AI will become a beneficial enhancement of our abilities to enrich our lives, the effective outcome of these different scenarios remains hard to predict at this point. *Tesla* uses neural networks in its car vision system to warn about potential collisions. *Google* has made great progress in the quality of its translations. The structure of these neural networks means that the algorithms are good at identifying objects in picture, putting together sounds to make up words and recognizing what to do in a game, but not yet beyond those tasks. When *IBM's DeepBlue* was able to defeat the world Chess champion Kasparov in 1997, it was considered a considerable step to the computer becoming "intelligent". However, while a computer could win a high profile chess game, and 20 years later, *IBM's Watson* won from the champion in a more complicated Go-game, which was an impressive engineering feat, but it is a highly specialized algorithm. It took AI less than three years to find solutions to beat the human champion because the human brains don't have the processing power to consider so many moves ahead. *DeepMind*'s Alpha Zero – bought by Google - works by playing hundreds of millions of games itself, pruning mistakes that led to losses, and elaborating on strategies that lead to wins. Such systems involving generative adversarial network techniques that generate and observe data. However, at the other hand, let us not ignore the fact that it is still proving difficult to get a robot arm to pick up a cup of water. These 'smart' machines are smart only in their specific domains. Present-day data engineering is still far away from matching the power and versatility of neurons and their synapses of our brain. Artificial Intelligence cannot yet make their own plan – as a conscious being can.

Current AI machine-learning algorithms are, at their core, rather simple and straightforward. Some may even describe these computers as dumb fast machines on steroids, using "stupid little neurons" as basis. AI is doing descriptive statistics in a way that is not science and would be almost impossible to make into science. At this point, AI is doing descriptive statistics in a way that is not science and would be almost impossible to make into science. At their core, the current algorithms are "dead simple stupid"; they work by brute computing force – "if you use reinforcement learning of credit-assignment feedback, you can get those little pieces to approximate whatever arbitrary function one wants", according to MIT social physicist Sandy Pentland. Despite the remarkable advances in computing, the hype about Artificial General Intelligence (AGI) – i.e. a general intelligence computer that will think like a human and possibly develop consciousness – smacks to science fiction, according to Venki Ramakrishnan, the 2009 Nobel Laureate for Chemistry, an idea that is supported by many other researchers. And he is not the only skeptical scientist regarding AGI. We do not have sufficient neuroscientific knowledge yet to understand what exactly consciousness[3] is, or how we remember a phone number, or the reasons why we suddenly loose memory, how exactly neurons interact. Deep learning machines and AI cannot answer the "*why*" question yet. We have no idea which parts of the brain – if the brain at all – are responsible for human consciousness. It seems that we tend to underestimate the complexity and creativity of the human brain and how amazingly general it is, compared to any digital device we have developed so far.

When Max Kurzweil – who quipped the notion of 'singularity' in his 2005 publication - declared that computers will become as intelligent as humans by 2035 – based on the power of exponential improvements as we have seen over the last 30 years - the spirit was out of the Pandora box. Max Tegmark (2016) at MIT refers to the view among AI experts that AI systems will probably (over 50%) reach overall human ability (AGI) by 2045, and very likely (with 90% probability) by 2075. From reaching human ability - singularity – it will move on to superintelligence in 2100 (75%). Again, there is no scientific proof for these inferences. Informed guesses at best. At such a point of singularity, computers will become as and likely more powerful than human intelligence; humans should progress to a stage of becoming "trans-human" – a cyber-human (electronically enhanced) or neuro-augmented (biological-genetically enhanced) human, or a *homo deus* - in order to remain relevant in a world where we may compete with Artificial Generally Intelligent machines. To some singularity is seen as an opportunity whereas others emphasize the dangers. It is true that most discussions around AI is focused on a narrow weak interpretation of AI (as in machines controlled by humans), and limited attention on the potential dramatic transformations that Artificial General Intelligence (AGI) may bring. Oxford philosopher Nick Bostrom coined the notion of "superintelligence" that may see humankind as a potential threat. However, as MIT Prof Brynjolfsson expressed it, any future depends on the choices we make. The near or further future is not different: "We can reap unprecedented bounty and freedom, or greater disaster than humanity has ever seen before". The future, however, will be an ever more demanding struggle against the limitations of our brain and intelligence. Singularity, on the other hand, is the point at which computers becomes as smart as us. With a rapidly changing ecology of intelligence and rapid evolution of machine-learning, we may need to consider the probability and advantages of an evolution towards cyborgs and superminds, above and beyond the *homo sapiens*.

According to these AI enthusiasts, the real risk with AGI is not malice but competence. Admittedly, a super-intelligent machine should be extremely good at accomplishing its goals. As long as these goals are aligned with ours, no problem. In case they are not, big trouble can be expected. Hence why the importance of bringing the "ought" or ethical dimension in the equation without further due. As a number of AI experts admit, postponing ethical critical thinking on AI until after goal-aligned AGI is built would be irresponsible and potentially disastrous. A super-AI machine lacking a moral compass would be like an unguarded projectile on steroids (that could seriously harm us). Safety engineering and ethical thinking is more than needed. We would not send humans to the moon without all precautions that safety engineering could think of. Similarly, we should not build super-computers or machine superintelligence (although we are still decades away for such an achievement, if at all) without in-built safety mechanisms to guide actions in an appropriate manner.

Intelligence tests of AI should build on Alan Turing famous "imitation game" test. A computer passes the Turing test, or imitation game if it can fool a human, during a question and answer session, into believing that it is, in fact, a human being. We are a long way from achieving this feat. Humans are very good and creative in connecting the dots of different frameworks, that can result in new innovative thinking or inventions. Our current algorithms are not yet very good at doing so.

According to the philosopher Daniel Dennett, Alan Turing could not foresee the uncanny ability of superfast computers to sift mindlessly through big data if which the internet provides an inexhaustible supply, and find probabilistic patterns in human activity that could be used to express "authentic"-seeming responses into the output for almost any probe human would attempt to decipher whether the computer is smart enough to fool us as authentic human. In other words, in such a case, the computer would outsmart and fool the human, without necessary being able to become more intelligent than humans. Hence why Dennett describes a plausibly multidimensional "computer-agent" more like an amygdala or cerebellum than a real mind; at best such a computer could be defined as a special-purpose subsystem that could play an enormous supporting role, but not "remotely up to the task of framing purposes and plans and building insightfully on its conversational experiences" that could resemble the general intelligence of a human.

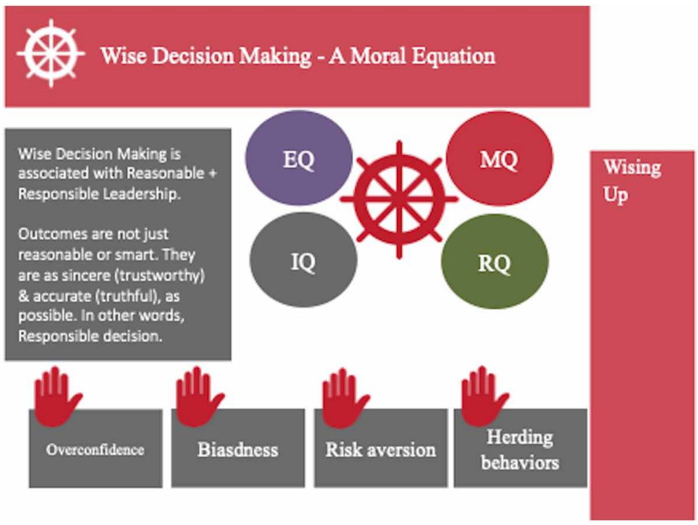
Moreover, current machine learning systems operate almost exclusively in a statistical or model-blind mode. In that sense, current systems remain rather opaque and focus on the what (if) question to execute a specific task. The Oxford quantum physicist, David Deutch – who conceptualized the notion of quantum computing - believes with the late philosopher Karl Popper that human-level intelligence and thinking tout court lie in the ability of creative criticism, interleaved with creative conjectures allowing humans to learn one another's behaviors, including language and extracting meaning from one another's utterances. The power of AI may be impressive but the G of AGI still remains elusive. It is that aspect of general creativity that leads to innovation that is a truly human characteristic. Add the ability to ask normative and thus ethical questions and to feel empathy and compassion (a form of emotional intelligence), and humans still have some distinctive competitiveness over "intelligent" machines.

Crucial in understanding the limitations of current (narrow) AI is the fact that our world knowledge is often expressed in causal terms, something that is not really done by current neural networks or deep (referring to the different layers like in neural brain network) learning machines. These machines are currently more about finding statistical regularities in complex patterns, but not really organizing that as objects that can have various kinds of causal impacts on other objects. Artificial General Intelligence would be the kind of intelligence that is contextualized, situationally aware, nuanced, multifaceted and multidimensional while also having the flexible learning abilities that humans have, which goes beyond supervised learning of big data (where computers excel over humans), unsupervised learning, reinforcement learning, virtual learning and other various kinds of learning. Not feasible in the near and medium time ahead – despite the enormous efforts of some genius researchers!

As mentioned a few times, computers are not able to answer the *why-question* which implies self-conscious and understanding causal networks. Today, Artificial Intelligence uses big data to operate or to conform to the (Bayesian) logic of probability and logic of proportions. Causal thinking altogether definitely escapes the current AI algorithms. A Bayesian network can only tell us how likely an event is, given that we observed another information; but it cannot answer counterfactual questions. Bayesian networks is basically the mathematical transformation of information or conditional probabilities without causality (e.g. given that I know, what is the chance that x occurs). It only identifies associations between variables, that may indicate correlation or regression, but it does not answer of possible causation. The computer scientists and mathematician Judea Pearl argues that the human brain is not just wired to do probability problems but to do causal problems – something the computer cannot (yet) (2019).

Our intuitive causal reasoning – we explicitly or implicitly use models - often clashes with the logic of probability and statistics in general. Whenever we see patterns, we look for a causal explanation. AI algorithms driven computers, on the other hand, are good at optimizing a specific task, but currently unable to understand cause-and-effect relations. AGI will need causal models, as it should become a

Figure 3. Including a moral, ecological and social sustainability dimension in decision-making
Source: Verhezen, P., (2018): Normative ethical questions (MQ) is inherently a human feature, besides emotional intelligence.



machine that can reflect on its actions and learn from past mistakes. UCLA Professor Judea Pearl (2019) believes that the algorithmitization of counterfactuals – he introduced the necessity of causal thinking in the AI-field - will be a major step toward understanding these why questions and making consciousness and agency a computational reality. Well, I guess that it all depends how we define and understand consciousness and mind in humans and how we ever may or may not be able to transfer these to computers.

Some researchers like Pearl believe it is possible to build a "moral causal thinking robot' in the distant future. We still assume that some form of consciousness and self-awareness is necessary to be able to think morally and care for others. Till then, a moral robot remains science fiction.

Boards Steering Towards Humans Collaborating with Artificial Intelligence, While Emphasizing Strict Data Governance

Can humans still compete with fast computers or are we losing the battle with powerful machines? Should we trust computers? How should one address the numerous risks and address the darker side related to Big Data and AI, while still be able to benefit from the speed and potential accuracy of AI (be it cognitive insights, cognitive engagement or digital automation)? Human society is a network just like the neural nets trained for deep learning, but currently the "neurons" in human society are still a lot smarter. Some artists have described the current form of AI as "artificial stupidity". But we should not underestimate the enhancing "creativity" of AI either.

We here focus on some *governance solutions* and *smart decision making* that are related to (1) privacy concerns, (2) biases in predictive recommendations, and (3) the audit of black box algorithms on the one hand. That some jobs will disappear with the increasing applications and growing predominance of AI cannot be denied. However, we recommend **wise corporate leadership** to accommodate these new innovative tools in the form of AI and deep learning machines, and stimulate collaboration between smart humans and intelligent machines where applicable and possible.

1. Governance that requires more Accountability and Responsibility

In order to reduce the risks of the darker side of AI, we suggest a more stringent data governance for those who use data provided by people who often have been unaware of the potential misuse of their privacy. We therefore suggest to enhance the accountability of organizations using these data to their commercial benefit. Secondly, we believe that recommendations in sensitive areas as justice or the medical field should not undermine the basic principle of human dignity and the principle that people should not be judged unless strong factual proof is available. And finally, we also believe that more transparency is needed to reduce the risks related to these AI "black boxes", by having gatekeeping third party certification that could enhance the trust in AI and its use of big data.

a. From Personal Privacy to Firm's Accountability

Although consumers and individuals should provide their consent in the use of their personal data, we suggest to be less focused on consent as such, but make the data users – i.e. the organizations using or selling those data – accountable for what they do with the data. It concretely means that the burden of responsibility is shifted from the public (and the aggregation of individuals who may have consented on the use of their data) to the effective users of data who actually benefit most of these secondary data use.

Another method is "differential privacy" which implies that the data are deliberately blurred so that a query of large data set only reveal approximate and not real results. The EU's General Data Protection Regulation (GDPR) has imposed strict rules to protect privacy. These privacy principles have been copied by many countries across the world. In addition, EU regulation is convinced that digital firms cannot lock out competition, such as *Microsoft*, *IBM*, and lately *Google* and *Facebook* have attempted to do for which they all were fined. And we do not even focus on the Chinese AI related organizations which are less bridled by privacy regulations; since the Chinese communist party here play a role that is questioned in the West. The rule in Western democracies imposes equal treatment of anyone, including rivals, who may use their platforms. The weekly *Economist* claims that that European regulations want consumers to control their privacy and how their data are monetized. The ability of European consumers to "switch creates competition that should boost choice and raise standards." The challenge is to make GDPR less clunky, and to avoid Europe to become a tech enclave, cut off from the mainstream, and allowing the (geopolitical) rivalry between US and China to create tech giants who will continue to act in an oligopolistic fashion.

b. Propensity of recommendation versus Innocent till proven guilty

Allowing governments to take predictive action because certain big data analysis suspecting someone is not good enough since it undermines the human dignity and the foundations that one is innocent till proven guilty. Not the other way around.

Indeed, the more we may switch to holding people accountable for their possible actions based on data-driven interventions and predictions with the intention to reduce the risk in society, the more we categorize or stigmatize certain groups, and the more we devalue the ideal of individual responsibility which we believe remains a fundamental human right.

Moreover, by pushing people to take certain decisions based on recommendations basically denies people's responsibility for their actions since it destroys their fundamental freedom to choose their own behavior, though admittedly often (unconsciously) influenced by peers and social pressure.

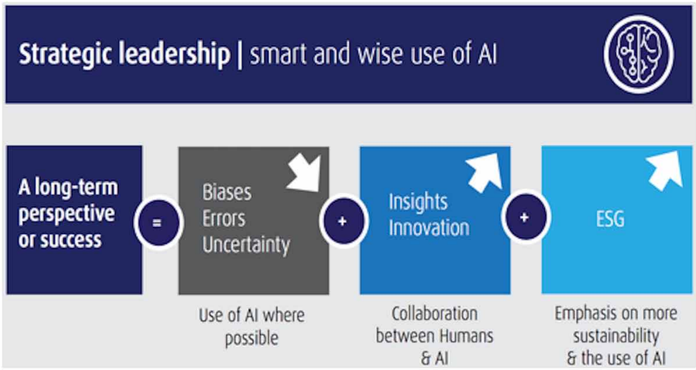
c. Audit of Algorithm or more transparency on Black Box

So far, most AI proponents have emphasized the positive practical consequences of artificial intelligence. However, the real danger lies in the fact that scientists do not really understand the "black box" behind the algorithm. The new machine-learning programs may have recognized patterns via deep neural networks and subsequent practical useful conclusions. But we have no idea how the computer or algorithm came up with the inference or conclusion. And the larger the data set, the more difficult it will become to understand and to analyze – even with the help of computers – these predictive conclusions. Does this mean that we will become completely dependent on the computer? It reminds me to the incredible and *hineinsight* misplaced faith in the practical wisdom of mathematical quants when calculating the "exact" value of these collateralized debt obligations – securitized risk tools – which partially caused the global financial crisis in 2008, bringing us to the brink of a real and complete global disastrous meltdown.

If we solely rely on the accuracy and "objectivity" of the conclusions prepared by black box algorithms, the more we ignore the reasons behind. Instead we strongly argue to bring back the notion of explainability, the why beyond the mere what. Explainability should not be sacrificed as the mimetic lamb in the name of more accuracy. In other words, we suggest that trust in AI can only be achieved by increasing more disclosure about the algorithm and the underlying prediction system. Hence why the introduction of "algorithmist" who acts as a reviewer of big data analysis and predictions in an impartial and purely confidential manner, could be sensible. In other words, a form of external certification is needed to create trust by a third party endorsing the reliability, replicability and accuracy of the AI algorithm. For instance, in the case of AI driving cars, AI diagnosing patients and in AI "robo" investing, some form of certification should become mandatory. And although ideation, creativity and innovation are often described as thinking outside the box, it does not justify secretive black boxes for nobody seems to be accountable.

We also could enhance transparency by using the new Blockchain technology - a P2P database in the absence of any central authority governing its use – currently applied for the rather volatile cryptocurrency Bitcoin. The central ledger system allows all participants to see every transaction by every other participant and encryption ensures no [chain of] block can be forged. In its essence, Blockchain is a disintermediation of trust with the use of mathematical guarantees. Our rather perceived subjective notion of trust is expressed as a computationally guaranteed property of the Blockchain system. In other words, blockchain provides an example of how a human ideal of trust is translated into mathematics and encrypted algorithmic code.

Figure 4. Wise leadership: Integrating Human Intelligence and Artificial Intelligence
Source: Verhezen, P. (2019) as published in http://www.verhezen.net/images/papers/VERHEZEN_2019_Wise_Decision_Making_and_AI_2nd_Paper.pdf



2. Future Collaboration between Humans and AI-learning machines

Digital tools and artificial intelligence in particular will enable humans to make better decisions when properly applied. We may not fully understand or control our destiny, but at least we have a chance to bend it in the direction of our own values which we feel are worth living for. The future is not just something that will happen to us – completely deterministically – but is likely something that we can and hopefully will build.

Nonetheless, the AI revolution that we can expect will be on scale of a real industrial revolution, hence why the notion of "*Industry 4.0*" was quipped. There will be beneficiaries but unfortunately also some losers in the process. Those jobs that need social interaction won't disappear immediately, those which are repetitive and can be optimized using data will be replaced by smart computers. Humans and AI are collaborating to improve five elements of business processes: (1) *Flexibility* (as in Robotics in Auto-manufacturing, Software to improve Product design, Software development estimates), (2) *Speed* (as in Fraud detection, aggregate patient data assisting in cancer treatment, Video analytics that enhance public safety), (3) *Scale* (as in Automated applicant screening in Recruitment, the use of bots in improving customer service, monitoring systems), (4) *Decision-Making* (as in diagnostic applications in equipment maintenance, real time Robo-advisors in financial services, Disease prediction) and (5) *Personalization* (as in wearable AI devices that improves the guest experience, wearable sensors to improve health care, and AI analytics in retail fashion). It seems that human-machine collaboration enables organizations to interact with employees and customers in new more effective ways.

Automation that eliminates a human from a task does not necessarily eliminate them from a job. One cannot deny that computerization, robotization and digitization have been responsible for considerable losses in blue-collar jobs in the last two decades. The physical labor that is at risk of being replaced by computers include tellers, cashiers, garment factory workers, fruit harvesters, assembly line inspectors and labor. Receptionist, bartenders, caterers may survive in the medium term but over a longer period, also their jobs will be at risk for being replaced by very smart machines. Presumably, elderly home caretakers, hair stylists, physical therapists will be able to keep their jobs which require a high level of dexterity within an unstructured environment. For the time being, aerospace mechanics, taxi drivers, plumbers and house cleaners may sustain their job for a little while longer.

However, lately the white-collar jobs are also under siege. Accountants, many legal and medical professionals, insurance adjusters, financial analysts and stockbrokers, travel agents, personal tax preparer, basic translators, and telemarketers all may disappear within years as result of ever more sophisticated machine-learning programs. Criminal defense attorneys, CEOs, psychiatrist, PR directors and social workers will likely keep their job because of the required social interaction and creativity or strategy based cognitive work required. Scientists, medical researchers and artists likely will sustain their job as well in the medium term.

Despite all the pessimistic forecasts of what jobs would be technically possible to do with machines, the actual job losses or resulting unemployment levels will likely much smaller. Some estimates forecast about 9% of jobs in the USA and Europe are at risk for being automated. PWC researchers find instead that 38% of jobs in the US *could* be at risk of automation by the early 2030s. The actual replacement is likely much lower, around 10-15%.

However, AI algorithms will be to many white-collar workers what tractors were to farmhands: a tool that significantly increases the productivity of each worker and thus shrinking the total number of employees required. Algorithms already exist; those ambidexterious robots still need to be invented.

In the long run, the biggest effect of automation and robots driven by algorithms is likely to be on workers in developing nations that currently rely on low-cost labor for their competitive advantage. Foxcon is taking a lead in automating and replacing blue workers in China. The question then becomes whether these Chinese factories in the global supply chain remain low cost and competitive. Probably, the low-cost advantage may disappear.

It will be an interesting but also challenging question how autonomous car will share the road with pedestrians, human-driven vehicles and other autonomous cars. How to combine the human values when they might be in conflict with the navigation system of autonomous cars? We need to think about this "robo-ethics" and make sure that these AI-vehicles take our human nature into account, so that they are well coordinated and well aligned with our desires and values system to increase our quality of life.

Moreover, as MIT professor Erik Brynjolfsson and research scientist Andrew McAfee argue about the paradox of robotics progress – known as the Moravec's paradox: contrary to traditional intuition, high level reasoning requires very little computation, but low-level sensorimotor skills require enormous computational resources. Robots and artificially intelligent computers are good at making precision welds on assembly lines, or at fast calculating the ideal distance in self-driving cars through its GPS system, but they still "can't tie their own shoes".

Robots and their AI behind are great at performing dirty, dull, monotonous and dangerous jobs that no one else wants to do. In terms of risk management, artificial intelligence out-maneuvers human intelligence in the field of *known knowns* such as fraud detection, medical diagnosis, bail decisions. In contrast to machines, humans are sometimes extremely good at prediction with little data. Cognitive superiority of humans, however, is still very valid in the area of ideation, large frame pattern recognition, and complex forms of communication. *Apple's Siri* and *Amazon's Alexa* can answer questions and can control devices around your home, but can they appreciate a joke? Not (yet) at this time.

As executives, making decisions, one needs to focus on the consequences (which one is able to know) rather than the probability (which one likely does not know). This is the idea of making executive decisions under uncertainty. AI deep learning machines may be able to assist (smart) decision-makers to provide some clue in terms of conditional probabilities and reduce this fundamental uncertainty.

Table 1. Comparison of AI and Humans (compiled by author)

	Little data available	*Big data available*
Known Knowns (Repetitive tasks)	Humans + AI (strong case for collaboration)	AI > Humans (better use AI = *Big Data Analytics and automation of repetitive tasks*)
Unknown Knowns *(incomplete information = randomness)*	Humans > AI (weaker case for collaboration, though *creativity of humans* may outsmart computers)	AI ≅ Humans (weaker case for collaboration though *computer power may augment human intelligence*)
Known Unknowns (Unique Decision)	Humans > AI (humans should take the lead, use *Heuristics*)	Humans + AI (strong case for collaboration where *AI augments the decision-power of humans*)
Unknown Unknowns *(Black Swans)*	Possibly Humans >AI ?? although too many unknown variables *(unclear what can be "expected")*	Possibly AI > Humans ?? *(unclear what can be predicted)*

Computers may be able to provide some lacking information (based on finding or revealing for humans hidden patterns) to improve decision-making. Unfortunately, with big data analytics the fundamental uncertainty may seem to vanish under (statistical) averaging, but randomness may therefore not have completely disappeared, despite the feeling that we are better informed through big data. However, in practice, this may be good enough (knowing how or practical technè versus the more elusive knowing that or episteme) to provide functioning AI driven tools.

In the face of *known unknowns* (rare events such as predicting earthquakes), humans make better decisions than machines. With *unknown knowns*, prediction machines appear to provide a very precise answer that could be very wrong. And in case of *unknown unknowns*, the black swans, Nassim Taleb convincingly argues that we cannot predict truly new events from past data, and likely both humans and machines fail. Nonetheless, in some instances, AI is able to uncover unknown unknowns as *GNS Healthcare* applies machine-learning AI to find overlooked relationships among data in patients' health records, enabling GNS to uncover a new drug interaction hidden in in unstructured patient notes.

However, we believe that machine learning and prediction can enhance the productivity of human decision-making by providing an initial prediction that humans can combine with their own assessments, such as checking the creditworthiness of loan applicants. Secondly, machine learning can provide a second opinion or facilitate monitoring (of patients in hospitals for instance). With a reliable diagnosis from an image, patients can forego an invasive biopsy. Advances in AI and machine learning mean less need to "satisficing" and more "ifs" and more "thens". In other words, AI allows for more complexity with less risk, transforming decision-making by expanding options. Machine-learning techniques are increasingly good at predicting missing information, including identification and recognition of items in images. This kind of pattern recognition to predict disease is what radiologists do. Prediction machines may be able to reduce uncertainty, but they won't always eliminate it.

Intelligent machines driven by AI use their specialized intelligence to solve parts of the problem; humans use their general intelligence to do the rest. In that sense, smart AI machines can help engage and coordinate large groups of people to become more effective and efficient.

Most likely, AI will shift human resources management toward the relational and away from the transactional. More crucially, the arrival of AI will lead to an increasing importance of human judgment (supported by AI-predictions). Prediction by AI and judgment by humans are complements: as the use of prediction increases, the value of judgment (or decision) rises. The (human) judgment uses (AI) predictions to make a smarter decision. Thinking about a network – as our brain functions – is analogous to thinking about entire ecosystems. How would you guide ecosystems to grow in a good direction? Artificial intelligence use methods to "learn" from large quantities of data where computers can recognize patterns and derive conclusions from these "insights". Big data is not use if one does not turn it into knowledge that can be applied in concrete cases. However, human learning and curiosity does not occur in isolation, but is always embedded in traditions and accumulated wisdom of past generations. Until scientists will have solved the basic paradox of learning, the best AI will be unable to compete with a four or five years old.

CONCLUSION

As long as we are aware of the potential threats from AI and take all possible measures to reduce those risks, including being aware of the dangers of complete dependency on technological digitization of our world view, we should be able to remain creative and innovate for a better and more fair world. The darker side of AI – be it the biases of weak AI, privacy violations, and overreliance on data of an AI driven Black Box, as well as the existential threats of a strong AI – cannot be ignored. We should be prepared for the danger of ideas, ideologies and institutions that allow information to feed collective decisions and understanding that may be contrary to an open and less dogmatic perspective of the world. If intelligence is the ability to deploy novel means to attain a goal, then we should allow some competitive forces to drive evolution. At this point, the progress of machine learning, particularly multilayered artificial neural networks, is not resulting yet in achieving general intelligence, but is mainly restricted to specific problems of mapping well-defined inputs to well-defined outputs. Likely the real danger lies not in the machine itself but in the way humans may use it. Instead of focusing on singularity, Berkeley scientist Ken Goldberg claims, we should embrace **multiplicity** – a hybrid view of how new technologies and humans might collaborate in partnership toward meaningful human solutions. Qualities like intuition, empathy, creativity remain crucial human qualities. Hence why a more holistic approach by humans could be blended with the precision that machines provide.

Artificial intelligence is becoming good at many human jobs such as diagnosing diseases, translating languages, providing customer service, and it is improving fast since 2013. Obviously, there is a fear that AI will ultimately replace human workers throughout the economy. But that does not need to be a necessity. Never before have digital devices and machine tools (Internet of Things is currently estimated to reach 20 billion devices) been so responsive to us. This kind of IoT technology may radically alter how works gets done and who does it. However, the impact may be even larger when AI technology will complement and augment human capabilities, not replacing them.

Admittedly, we do not understand yet the neural architecture of the human brain that through evolution was trained to run well. But evolution goes slow and we start to understand the constraints and limitations of our brain power. However, to achieve general intelligence, or human-level intelligence, learning machines need to ask normative questions that are guided by a kind of blueprint of reality, a model of society in which we aspire to live in. At this point, no learning machine is able yet to answer causal "what if" questions. The machine does not have a "purpose"; AI only aims to achieve some specific objectives, not the design of those – be objectives or purpose at this point in time. Data science only facilitates the interpretation of data and connect them with and to reality. No matter how big the data get and how skillfully they are manipulated by data scientists, such machine learning and AI remain quite opaque in their learning. AI research has so far focused on systems that are better and much faster at making decisions, but not necessarily at making better decisions (which would require a normative or value-laden stance). Machine's decisions may be ineffably stupid in the eyes of humans if the objective or utility function of the machine is not appropriately aligned with our human aspirations and values. Unless, an artificial general intelligence would emerge that has its own conscious to make their own related goals, utilities and values to be followed – almost completely independent of ours, which remains "science-fiction" today. However, if a weak or strong AI is meant to serve humankind, then we need to think about AI and its purpose in relationship to how we see a fairer and more just world, which is an ethical question. Addressing such enormous challenges will require top notch engineering, computer science, legislation and likely above all, moral and wise leadership to guide us through.

What makes us different from AI, is our unique history which gives us our notions of purpose and goals. There is no meaningful sense in which there is an abstract notion of purpose. It is likely embedded within our own history and traditions. Although there may not be a genuine demarcation line between intelligence and mere computation, we still see different kinds of intelligence beyond the brute force of rational computational logic thinking, including making sense and giving meaning on an experience. Finding a purposeful meaning – based on human values and human insights - to one's life or thinking while using working together with smart machines, is the best scenario we should strive for. Hence why wise beyond smart decision-making is not just a luxury but a necessity. Wisdom is breadth and a broader and often longer term framing.

At this point, deep learning machines are not achieving any form of general artificial intelligence, but have become good at predicting based on big data. The key question is not whether AI will bring benefits, but how those benefits will be distributed and how we can limit the dark side of AI. Will some companies and countries have a huge competitive advantage? And will this advantage be translated in a vicious or rather virtuous stream of decisions?

We have argued that [weak] AI may have contributed to better and decentralized matching processes but it has occasionally also created new forms of information asymmetries through centralized data in the hands of a few tech giants. Hence why some forms of transparency of the [use of] data and information [by algorithms] is necessary, but likely difficult to achieve. So, the alternative is specific regulations wanting to ensure competitive markets that mandate the [progressive] sharing of data. The future of our economy lies in the clever exploitation of our informational surplus which we can achieve in our data-rich markets. AI and big data can enable better human coordination in the market which makes us more sustainable. The old adage of "trust but verify" is still valid. However, we cannot and should not ignore the darker side of artificial intelligence that could easily undermine the trust in these new technology. The future lies in a beneficial collaboration between human general intelligence with artificial specific intelligence and deep learning machines. The human advantage lies in the ability to ask metaphysical (why) questions and address ethical concerns. Only humans can feel empathy and mindful compassion towards other beings – which seem to constitute ourselves as social beings (based on emotions inherent to most higher-level mammals). In addition, our neocortex allows us to think rationally or reasonable and make links between unexpected patterns that result in innovative and insightful improvements of tools.

And the use of intelligent artificial tools could improve organizations' products and services positively impacting our quality of life.

REFERENCES

Agrawal, A. (2018, Apr.). The economics of Artificial Intelligence. *McKinsey Quarterly*.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines. The Simple Economics of Artificial Intelligence*. Boston, MA: Harvard Business Review Press.

Alexandre, L. (2017). *La Guerre des Intelligences: Intelligence Artificielle versus Intelligence Humaine*. Paris: JC Lattès.

Alpaydin, E. (2016). *Machine Learning. The new AI*. Cambridge, MA: The MIT Press.

Arbib, M. A. (Ed.). (2003). The Handbook of Brain Theory and Neural Networks. Cambridge, MA: The MIT Press.

Ariely, D. (2008). *Predicably Irrational. The Hidden Forces that Shape our Decisions*. New York: HarperCollins.

Baer, T. & Kamalnath, V. (2017). *Controlling machine-learning algorithms and their biases.* McKinsey & Company.

Barrat, J. (2013). *Our Final Invention. Artificial Intelligence and the end of the Human Area*. New York: Dunne Books.

Batra, G., Queirolo, A., & Santhanam, N. (2017). *Artificial Intelligence: The time to act is now.* McKinsey & Company.

Baxter-Kellman, R. (2015). *Het Membership Model. Wat je van Netflix, Spotify en Starbucks kunt leren*. Amsterdam: Business Contact.

Berns, G. (2008). *Iconoclast. A neuroscientist reveals how to think differently*. Academic Press.

Borison, A., & Hamm, G. (2010). Prediction Markets: A new tool for strategic decision making. *California Management Review*, *52*(4), 125–141.

Bostrom, N. (2014). *Superintelligence. Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press.

Bourton, S., Lavoie, J., & Vogel, T. (2018). Will Artificial Intelligence make you a better leader? *The McKinsey Quarterly*.

Brafman, O., & Brafman, R. (2010). *Click. The Magic of Instant Connections*. New York: Random House.

Brockman, J. (Ed.). (2019). *Possible Minds. 25 Ways of Looking at AI*. London: Penguin.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age. Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: WW Norton Company.

Bughin, J., & van Zeebroeck, N. (2018, July). Artificial Intelligence: Why a Digital Base is critica. *McKinsey Quarterly*.

Burkhardt, R., Hohn, N. & Wigley, C. (2019). *Leading your organization to responsible AI.* McKinsey Analytics.

Chalmers, D. (1996). *The Conscious Mind. In search of a Fundamental Theory*. New York: Oxford University Press.

Christensen, C. M., Hall, T., Dillon, K., & Duncan, D. S. (2016). *Competing against Luck. The Story of Innovation and Customer Choice*. New York: HarperCollins Business Publ.

Davenport, T. H., & Ronanki, R. (2018, Jan.). Artificial Intelligence for the Real World. Don't start with moon shots. *Harvard Business Review*, 108–116.

Deheane, S. (2014). *Consciousness and the Brain: Deciphering How the Brain codes our Thoughts*. London: Penguin.

Domingos, P. (2015). *The Master Algorithm. How the quest for the ultimate learning machine will remake our world*. London: Penguin.

Erisman, P. (2015). *Alibaba's World. How a remarkable Chinese company is changing the face of global business*. London: Palgrave MacMillan.

Finger, Z. S. (2001). *Origins of Neuroscience: A history of explorations into Brain Function*. Oxford, UK: Oxford University Press.

Finkelstein, S., Whitehead, J., & Campbell, A. (2008). *Think Again. Why good leaders make bad decisions and how to keep it from happening to you*. Boston: Harvard Business Press.

Finlay, S. (2018). *Artificial Intelligence and Machine Learning for Business. A no-nonse guide to Data-driven Technologies*. London: Relativistic Publications.

Ford, M. (2018). *Architects of Intelligence. The truth about AI from the people building it*. Birmingham, UK: Packt Publishing.

Furr, N., & Shipilov, A. (2019, July). Digital doesn't have to be Disruptive. The best results can come from adaptation rather than revinvention. *Harvard Business Review*, 94–103.

Gazzaniga, M. (2011). *Who's in charge? Free will and the science of the brain*. New York: HarperCollins.

Gazzaniga, M. (2015). *Tales from both sides of the Brain. A life of neuroscience*. New York: HarperCollins.

Geissbauer, R.; Schrauf, S. & Pillsbury, S. (2018). Digital Champions. *Strategy + Business*.

Gigerenzer, B. (2014). *Risk Savvy. How to make good decisions*. Lane Books.

Gilovich, T., & Ross, L. (2016). *The Wisest One in the Room. How to harness Psychology's Most Powerful Insights*. London: OneWorld Publishing.

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision-Making. *Annual Review of Neuroscience*, *30*(1), 535–574. doi:10.1146/annurev.neuro.29.051605.113038 PMID:17600525

Grant Halverson, H. & Rock, D. (2015). Beyond Bias. *Strategy + Business*, (80).

Gratton, L. (2011). The Shift. The future of work is already here. London: William Collins (Harper Collins).

Gupta, S. (2018). *Driving Digital Strategy. A guide to reimagining your business*. Boston, MA: Harvard Business School Press.

Heath, C., & Heath, D. (2013). *Decisive. How to make better decisions*. London: Randomhouse.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin-Pearson.

Kahneman, D. (2011, Nov.). Beware the inside view. *McKinsey Quarterly*.

Kahneman, D. & Klein, G. (2010, Mar.). Strategic Decisions: When Can You Trust Your Gut? *McKinsey Quarterly*.

Kahneman, D., Lovallo, D., & Sibony, O. (2011, June). Before You Make That Big Decision. *HBR*, 50-60.

Kahneman, D., Rosenfield, A.M., Gandhi, L., & Blaser, T. (2016). Noise. Inconsistent Decision-making. *HBR*, 38-46.

Kanter, R. M. (2009). Supercorp. How Vanguard Companies Create Innovation, Profits, Growth, and Social Good. New York: Academic Press.

Kanter, R. M. (2011, Nov.). How Great Companies Think Differently. *Harvard Business Review*, 66–78.

Kevles, B. H. (1997). *Naked to the bone: Medical Imaging in the Twentieth Century*. New Brunswick, NJ: Rutgers University Press.

Kiel, F. (2015a). *Return on Character. The Real Reason Leaders and Their Companies Win*. Boston: Harvard Business Review Press.

Kiel, F. (2015b, Apr.). Measuring the Return On Character. *Harvard Business Review*, 20–21.

Kizza, J. M. (2017). *Ethical and Social Issues in the Information Age*. Cham: Springer. doi:10.1007/978-3-319-70712-9

Klein, G. (2013). *Seeing What Others Don't. The Remarkable Ways We Gain Insights*. London: Nicholas Brealey Publishing.

Kurzweil, R. (2006). *The Singularity is near. When humans transcend biology*. London: Penguin.

Kurzweil, R. (2012). *How to create a Mind. The secrets of human thoughts revealed*. London: Viking.

Lee, K.-F. (2018). *AI Superpowers China, Silicon Valley, and the New World Order*. New York: Houghton Mifflin Harcourt Publishing Company.

Libert, B., Beck, M., & Wind, J. (2016). *The Network Imperative. How to survive and grow in the Age of Digital Business Models*. Cambridge, MA: Harvard Business Review Press.

Lieberman, M. D. (2013). *Social. Why our brains are wired to connect*. New York: Penguin Randomn House.

Likens, S. & Kersey, K. (2019, May). Automating trust with new technologies. *Strategy + Business*.

Lindstrom, M. (2008). *Buyology: Truth and Lies about Why we Buy*. New York: Broadway Books.

Lohr, S. (2015). *Data-ism. The revolution transforming decision-making, consumer behavior, and almost everything else*. New York: HarperCollins.

London, S., Chui, M. & Whigley, C. (2019). *The ethics of Artificial Intelligence.* McKinsey & Company.

Lozano, J. M. (2010). *The Relational Company. Responsibility, Sustainability, Citizenship*. Bern: Peter Lang. doi:10.3726/978-3-0353-0326-1

Luca, M., Kleinberg, J., & Mullainanthan, S. (2016, Jan.). Algorithms need managers, too. Know how to get the most out of your predictive tools. *Harvard Business Review*, 96–101.

Luger, G. F., & Stubblefield, W. A. (1998). *Artificial Intelligence. Structures and Strategies for Complex Problem Solving*. Harlow, MA: Addison Wesley Longman.

Makridadis, S., Hogarth, R., & Gaba, A. (2010). *Dance with Chance. Making Luck Work For You*. London: Penguin.

Malone, T. W. (2018). How Human-Computer ‘Superminds’ are Redefining the Future of Work. *MIT Sloan Management Review*, 34–41.

Manyika, J., Chui, M., & Lund, S. (2017). *What’s now and next in Analytics, AI and Automation*. McKinsey & Company.

Nanjok, N., Le Fleming, H. & Srivatsay, N. (2018). Digital Technology and Sustainability: Positive Mutual Enforcement. *Strategy + Business.*

Penrose, R. (2016). *The Emperor’s Mind: Concerning Computers, Minds and the Laws of Physics*. Oxford, UK: Oxford University Press.

Pernick, R., & Wilder, C. (2008). *The Clean Tech Revolution. Discover the top trends, technologies, and companies to watch*. New York: HarperCollins.

Polson, N., & Scott, J. (2018). *AIQ. How Artificial Intelligence works and how we can harness its power for a better world*. London: Bantam Press-Penguin.

Priest, D., Krishnamurthy, & Blanter, A. (2018, Apr.). The New Automation is Smart, Fast and Small. Emerging digital tools and techniques are reinventing large-scale IT initiatives, one process at a time. *Strategy + Business*.

Rao, A. & Cameron, E. (2018, July 31). The future of Artificial Intelligence depends on Trust. *Strategy + Business.*

Rees, M. (2018). *On the Future. Prospects for Humanity*. Princeton, NJ: Princeton University Press.

Rifkin, J. (2014). *The Zero Marginal Cost Society. The internet of things, the collaborative commons, and the eclipse of capitalism*. New York: Palgrave MacMillan.

Satel, S., & Lilienfeld, S. O. (2013). *Brainwashed. The Seductive Appeal of Mindless Neuroscience*. New York: Perseus Books Books.

Schoemaker & Tetlock. (2017). Building more intelligent enterprises. *MIT Sloan Management Review*, *58*(3), 27-37.

Schouteden, C., Sturges, D, Chow, E., Meloni, M., Engineer, N., Lange, A., … Saleh, T. (2018). *Artificial Intelligence and the Circular Economy. AI as a tool to accelerate the Transition.* Ellen MacArthur Foundation + Google + McKinsey & Company.

Schwartz, J.; Thomson, J. & Kleiner, A. (2016, Dec.). The Neuroscience of Strategic Leadership. Research shows how leaders can take the high road less traveled. *Strategy + Business*.

Schwarz, B., & Sharpe, K. (2010). *Practical Wisdom. The Right Way to Do the Right Thing*. London: Penguin Books.

Shermer, M. (2008). Five ways Brain Scans mislead us. *Scientific American, Novembner*, *5*, 1412–1414.

Siegel, J. D. (2010). *Mindsight. Transform your brain with the new science of kindness*. London: One-World Publications.

Siegel, J. D. (2017). *Mind. A Journey to the Heart of Being Human*. New York: Norton & Company.

Silver, N. (2012). *The signal and the noise. Why so many predictions fail – but some don't*. New York: Penguin Group Publishing.

Sumpter, D. (2018). *Outnumbered. From Facebook and Google to fake news and filter-bubbles – the Algorithms that control our lives*. London: Bloomsbury Sigma.

Taleb, N. N. (2005). *Fooled by Randomness. The hidden role of Chance in Life and in the Markets*. New York: Random House Trade Paperbacks.

Taleb, N. N. (2007). *The Black Swan. The Impact of the Highly Improbable*. London: Allen Lane – Penguin.

Taleb, N. N. (2012). *The Antifragile*. Academic Press.

Taleb, N. N. (2018). *Skin in the Game. Hidden Asymmetries in daily life*. New York: Random House.

Tegmark, M. (2016). *Life 3.0. Being Human in the age of Artificial Intelligence*. London: Penguin.

Tenbrunsel, A. E., Diekmann, K. A., Wade-Benzoni, K. A., & Bazerman, M. H. (2010). The Ethical Mirage: A Temporal Explanation as to Why We Are Not As Ethical as We Think We Are. *Research in Organizational Behavior*, *30*, 153–173. doi:10.1016/j.riob.2010.08.004

Tononi, G. (2012). *Phi. A Voyage from the Brain to the Soul*. New York: Pantheon Books.

Verhezen, P. (2013). Managerial Wisdom in Corporate Governance. In M. Thompson & D. Bevan (Eds.), *Wisdom and Complexity in Organizations*. London: CEIBS & Palgrave. doi:10.1057/9781137002655.0016

Wilson, J., & Daugherty, P. R. (2018, July). Collaborative Intelligence: Humans and AI are joining forces. *Harvard Business Review*, 114–123.

Zarkadakis, G. (2015). *In our Image. Savior or Destroyer? The History and Future of Artificial Intelligence*. New York: Pegasus Books.

Zhang, T., Gino, F., & Bazerman, M. H. (2014). Morality Rebooted: Exploring Simple Fixes to Our Moral Bugs. *Research in Organizational Behavior*, *34*, 63–79. doi:10.1016/j.riob.2014.10.002

ENDNOTES

[1] Reinforcement learning is "an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learning and unsupervised learning. It differs from supervised learning in that labelled input/output pairs need not be presented, and sub-optimal actions need not be explicitly corrected. Instead the focus is finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge). The environment is typically formulated as a Markov decision process (MDP), as many reinforcement learning algorithms for this context utilize dynamic programming techniques. The main difference between the classical dynamic programming methods and reinforcement learning algorithms is that the latter do not assume knowledge of an exact mathematical model of the MDP and they target large MDPs where exact methods become infeasible". Cf https://en.wikipedia.org/wiki/Reinforcement_learning

[2] Even academic research itself is driven by citation indices and impact, because they were easy to calculate. In the process, they became the currency of science. In other words, an algorithm dictates in the academic career path who gets tenure track and who does not. Welcome to the tyranny of data that apparently drives the life of many.

[3] Guilia Tononi argued to "quantify" consciousness, denoted by the Greek letter 'Phi', measuring how much different parts of a system know each other. The consciousness theory become known as the integrated information theory (IIT) which logically postulates that computers cannot have a real consciousness. This theory has been challenged by David Chalmers and AI expert Murray Shanahan.