Chapter 6
Risks and Side Effects of Data Science and Data Technology

Clemens H. Cap

Abstract In addition to the familiar and well-known privacy concerns, there are more serious general risks and side effects of data science and data technology. A full understanding requires a broader and more philosophical look on the defining frames and on the goals of data science. Is the aim of continuously optimizing decisions based on recorded data still helpful or have we reached a point where this mind-set produces problems? This contribution provides some arguments toward a skeptical evaluation of data science. The underlying conflict has the nature of a second order problem: It cannot be solved with the rational mind-set of data science as it might be this mind-set which produces the problem in the first run. Moreover, data science impacts society in the large—there is no laboratory in which its effects can be studied in a controlled series of experiments and where simple solutions can be generated and tested.

1 Introduction

Data science has been defined by Braschler et al. (2019) as the *unique blend of skills from analytics, engineering, and communication, aiming at generating value from the data itself*. Data technology may be regarded as the *method of collecting data and deducing empirical and statistical models and making decisions thereon* with the help of algorithms. We focus on two aspects: Personal data, where data science affects the privacy of the connected persons, and model deduction, which may influence our way to do science and to make decisions. We perceive data science as an applied science, as a technology, and as a decision mechanism. Thus, technology assessment seems a reasonable thing to do. Risks connected with personal data and social "transformatory" risks, that is, resulting in changes of society, seem the foremost.

C. H. Cap (✉)
Universität Rostock, Rostock, Germany
e-mail: clemens.cap@uni-rostock.de

M. Braschler et al. (eds.), *Applied Data Science*,
https://doi.org/10.1007/978-3-030-11821-1_6

This task is difficult and urgent. Data technology is not yet broadly deployed in society, therefore the possible risks have not yet shown up in a larger number of concrete cases. They are not yet well understood and thus seem frightening. In nuclear technology or aviation, over 60 years of research and a tradition of detailed accident analysis produce a different situation and allow more clear statements. Nevertheless, assessment of data science is important and must not be delayed: Data technology is a pervasive technology and risks of a ubiquitous infrastructure are difficult to avoid should there be significant negative side effects. Data science infrastructure will soon have been integrated tightly into many products, workflows, and systems. While energy production can be, and has occasionally been, converted from nuclear to non-nuclear technology, data science applications are difficult to stop once they have been deployed, due to their universal integration into technical systems and social processes. The debate is emotionalized, risks affect everybody, leakage and data theft scandals heat up the discussion, cyber-wars and surveillance allude to fear, and much money can be made by industry.

Finally, it is not clear how to conduct systematic data science risk analysis. Side effects are not of a biological nature and cannot be studied in a lab; polls provide reliable answers only after (irreversible?) social change has been completed; philosophical debates on human values may be appropriate but employ methods which, unfortunately, rarely are taken seriously by the core target groups: data scientists, company owners, and policy makers.

2 Main Risks and Side Effects

Security and privacy issues are the foremost category of problems commonly associated with data science. As they are well discussed we shall only provide a brief overview. The Whitehouse Report on Big Data[1] provides a wealth of case studies in the areas of access to credit and employment, higher education, and criminal justice and connects them with possible though avoidable flaws in the big data process. Other authors raise more fundamental questions[2] or focus on the possibility of classification (Dwork and Mulligan 2013) and a resulting loss of different perspectives, individual autonomy, and democratic functioning of society.

The first trouble is *abuse of personal data.* It frequently leads to decisions which are not in the interest of the person involved. A well-known example is travel booking. Depending on the web browsing history, cookie pricing attempts to make the most profit from a customer. Every available information on social and financial stratum, past spending habits, and even computer brands is translated to offers and booking modalities which are optimal for the selling agency. Customers who do not complete the booking are followed through social media, search engines, and even email. Service providers employ all kinds of sensors, from fitness trackers

[1] See Executive Office of the President (2016).

[2] For example, Kree and Earle (2013).

to car black-boxes, to gather information on how to best make money from the observed person. Even smart phones owned by the customer are employed as mobile selling platforms of their manufacturers. To obtain the required platform control, software technologies are used to set up walled gardens preventing every escape of their owner. See Cap (2016) for a wealth of further examples.

Often *data theft* increases the risks of data abuse. We also face the problem of *incorrect data* and *data stickiness*, that is, situations where wrong or no longer valid data affect the life of an individual, which is unable to dissociate itself from data collections or even data errors in the past.

Given the wealth of deductions which can be made on persons from their data, some critics even question whether the dignity and autonomy of man would allow for a concept of a third party "*owning*" personal data of somebody else (Hofstetter 2016).

The *digital panopticon* is a further aspect, which originated in the surveillance debate. It leverages a thesis originally from Bentham (Warriar et al. 2002): A person feeling watched by an anonymous observer is likely to adhere to the ethical standards fantasized for the observer. This concept is further reflected in the mindset of Google whose former CEO Eric Schmidt suggested that "[if] you [had] something that you [didn't] want anyone to know, [. . .] you shouldn't be doing it in the first place." According to critics this statement demonstrates a complete lack of understanding of the concept of privacy (Esguerra 2009).

The **infrastructure risk** points out an important modality of data technology. Applications require a wide deployment of data sensors. Internet of Things experts speculate on more than 50 billion networked devices by 2020 (Hosain 2016). Data technology penetrates workflows, decision processes, and business plans. It promises convenience and optimized decisions. Ultimately, society finds itself in circumstances so nicely described by the sorcerer's apprentice.[3] When the technology has been deployed, it is extremely difficult to stop it or even reduce it—for technical, social, and economic reasons. Even if data science might guarantee the best of all possible worlds, we should be careful as a society when setting up an infrastructure where this final result is granted without a possibility of a later intervention.

Example: On a recent plane travel, the author was asked to have his boarding pass scanned as precondition for buying bottled water. Leaving aside duty-free and tax requirements, which could have been satisfied by manual ticket inspection or a higher price, the infrastructure problem was that the sales assistant could not even open the drawer of the cash register without a scan of personal data. Recently the Amazon Echo voice device behaved as the literal sorcerer's apprentice: A 6-year-old Texan girl was chatting with the device of her parents and asked it to order a dollhouse—which promptly was delivered. When the story was reported by a TV station, the reporter said: "I love the little girl, saying 'Alexa order me a dollhouse'." Devices in many a viewer's living room heard this command and placed orders for dollhouses (Nichols 2017).

[3]Famous German poem "Der Zauberlehrling" by Johann Wolfgang von Goethe, in which the control on an initially helpful spirit is hopelessly lost.

A disruptive change of science: In traditional science, theoreticians develop mental models which then are put to the test by observation. The models are cognitive constructions of the mind and do not constitute "reality"—although the act of confusing models with reality often helps a scientist to improve models: The earth is not flat but until the development of better astronomic instruments this model was helpful; the earth is not a sphere either, but only advanced questions really required the model of a rotational ellipsoid. The earth is not an ellipsoid either. All these mental models are "wrong" but helpful in the sense that they provide the physicist with constructions for "understanding" the world. Ultimately, in quantum physics the attempts to model observations by "machinery" are believed to fail. We recall Richard Feynman: "If you think you understand quantum mechanics, you don't understand it." Still the physicists' minds heavily and successfully use imagined machinery[4] since these cognitive tools fit the human mind.

Data science replaces this machinery by empirically validated models. In the optimal scenario, it drops theory and delivers the "best" numeric description of billions of experiments. This approach is hard to beat empirically. Why bother for an explanation if black-box predictions match myriads of experiments? This may be particularly attractive in complex system science such as medicine. Why bother to develop an explanatory description of a disease if a computer can diagnose and treat it much better?

Example: Anderson (1989) describes how a neural network can learn to control an inverted pendulum without prior knowledge of dynamics. The algorithm produced sets of real numbers as connection weights which interpolate complex functions with sufficient precision—it does not produce any "understanding" of the "learned" problem. These weights heavily depend on the structure of the network and on the randomization throughout the learning process.

Data science produces a solution for a problem (e.g., treating a disease), an abstract mathematical model for a complex object (e.g., an inverted pendulum), and in a few cases even additional insight into correlations and statistical mechanisms—it usually does not provide the mental models a human will use for "understanding" a system. We can argue that our mental models are incorrect, since modelling employs complexity reduction. However, the human mode of understanding our world and communicating about it is exactly in those "wrong" but vivid and demonstrative mental models which are close enough to the human mind. A disruptive transformation of science which replaces the human researcher by an algorithm is not desirable. The offer is, of course, tempting, but maybe we should reject it.

A GIS research project at the University of Zurich demonstrates a more refined approach (Schönholzer 2017): Studies indicate that repeated use of navigation systems weakens the sense of orientation of the user. Thus, the group now studies

[4]Feynman also acknowledges this. "I see vague pictures of Bessel functions [. . .] and dark brown x's flying around. And I wonder what the hell it must look to the students." See Root-Bernstein and Root-Bernstein (1999).

how the interaction of a user with a navigation system should be restructured to avoid superfluous use of the system and prevent further degradation of human capabilities. An interesting paradox arises: The more we develop helpful tools the more our natural skills degenerate. This is well known from other areas: An increasingly immobile life-style, for example, calls for regular compensation on the treadmill. With an increasing number of smart assistants and with data science applications taking over our decisions, we will face such paradoxes more and more often. Which mechanisms will prevent our cognitive decline? How reasonable are technical tools which, when used, destroy our abilities? Can this effect be counteracted successfully, as the Zurich GIS research project intends to do? Would it be more reasonable to use these tools less often? Do we have the necessary self-discipline? What protects us on a larger scale from first using such tools, then showing signs of degeneration, thus requiring such tools and finally becoming dependent on machines and associated business cases to manage our lives?

A **replacement of humans by algorithms** is closer than we might be aware of. We witness the trend in autonomous cars; financial trading already is dominated by algorithms and not only the allocation (Park et al. 2014) but also the selection (Miller-Merell 2012) of human resources will soon be taken over by computers. The latter means computers will decide which humans get a job, where and why. The pattern of replacing humans by machines was seen in the first industrial revolution, where it was for physical tasks. While a replacement of humans for repetitive, routine, dangerous, and boring tasks seems fine, we might cross a fundamental boundary when human *deciders* are systematically replaced by algorithms. Even if all involved parties benefit from better decisions, the issue at stake is the loss of core values of meaning of human life. While not everybody might agree that meaningful work is one of the core purposes of our existence, possible alternative worlds are not very attractive. Scenarios comprise dystopias such as “Brave New World” by Aldous Huxley, where the purpose of life is reduced to consumption and instantaneous satisfaction of needs, or more recently and drastically by the science fiction movie series Matrix, where humans are the mere appendix of a world governed by machines. Even if the outcome were a true paradise: Which effects would drive human evolution and prevent degeneration? What would we enjoy in our lives when it is no longer reward or success that sweetens labor? Can mankind exist for long being served by robots, regularly aroused by stimulating drugs, a kind of “soma” as described by Huxley?

Reverse and meta risk assessment: A thorough technology assessment also raises the following questions: What could possibly go wrong if risks are not analyzed correctly? What if they are communicated incorrectly to deciders or to the general public? What if they are perceived as larger than they “really” are?

A negative image of data science in public primarily affects the data science profession; there may be consequences for the acceptance of data technology products at large; ultimately, regulatory and legislative processes may damage the industry as such. It is important to recognize the differences between a “true” risk evaluation and a debate on the perceived risks (namely, negative *image*).

There certainly is the risk of rejection of data science due to a perceived abuse. It affects society at large through the loss of possible benefits. The trend to data avoidance or to a more sparing use of data[5] is promoted by privacy activists and may lead to insufficient data bases and incorrect models. The discussion therefore must also focus on the opposite question: May an individual claim the right to withhold data for research, thereby damaging the right of society to deduce possibly important results on these data? The Charta of Digital Rights[6] gives an affirmative answer to this question in Article 11(2).

Chosen ignorance is also an aspect to be discussed as part of a reverse risk assessment. Although being very skeptical toward possible side effects and naïve data science enthusiasm, the author neither considers nor suggests complete abstinence from data science. The famous quote of nuclear bomb physicist Edward Teller can guide us (Shattuck 1997): "There is no case where ignorance should be preferred to knowledge—especially if the knowledge is terrible." The allusion to nuclear weapon technology is not an over exaggeration of the author: Hannes Grassegger and Mikael Krogerus (Grassegger and Krogerus 2016) use this metaphor to speculate on the impact of data science on Brexit and 2016 US presidential elections. They cite[7] a psychologist on his application of data science to psychometric data: "I did not build the bomb. I only pointed out its existence." This is similar to the apologetic position taken by some physicists toward the discovery of nuclear chain reactions as basis for atomic bombs.

Data science applied incorrectly: In this article we presume that data science is done correctly and will not consider the risks of bad data science. They are said to comprise wrong results, bad analytics, bad data, and associated costs (Marr 2015).

3 Important Aspects

Data science and its core applications can be described as technology of optimization. Its ultimate vision is irresistible: *Observe everything, determine the best model automatically and provide us with the optimal answer to our decision problems.*

Deep Blue[8] is an impressive success of artificial intelligence. The system consistently wins chess against human opponents. While it provides for the pride of its programmer, ultimately it destroys the magic of the game of chess.

IBM's Watson is known as winner of the game-show Jeopardy! The impression this success made on the public amounts to a framing error, since of course, but

[5]This can be expressed more precisely in German with the hard to translate terms of "Datensparsamkeit" and "Datenvermeidung".

[6]Charter of Digital Fundamental Rights of the European Union. See https://digitalcharta.eu/wp-content/uploads/2016/12/Digital-Charta-EN.pdf

[7]English translation by the author.

[8]See Wikipedia article on Deep Blue: https://en.wikipedia.org/wiki/Deep_Blue_(chess_computer)

contrary to wide-spread belief, success in Jeopardy! is not so much connected with "knowledge" or even cognitive reasoning but rather with mere storage of many facts and clever engineering. Game-shows where single line answers must be given in a short period of time and decide on winning or losing dangerously misrepresent the ability to store facts as "knowledge" or "intelligence." The weakness of Watson's statistical reasoning is demonstrated impressively by its famous mistake that Toronto was a city in the USA (O'Connor 2011), although formally it is not completely wrong as there are several Torontos in the USA.

Watson, of course, is capable of "reading" medical research papers at speeds much higher than a human doctor. Direct comparisons with human specialists provide promising headlines for patients, especially when it comes to rare diseases or to overlooking symptoms (Galeon 2016). Claims, however, that Watson soon might be the best doctor in the world (Friedman 2014) are dangerously misleading. We are reminded of the famous surprise of Weizenbaum (1976) when psychologists started to discuss therapeutic benefits of his ELIZA program, which was intended as a study in pattern matching and later was attributed human-like feelings by observers.

We argue that the debate is not about better data science, smarter algorithms, and faster processors but that it is affected by a fundamental framing error. From a doctor, a patient also requires support in sufferings, pain reduction, moral support, sharing in the desperate feelings, or other forms of human help, which cannot be outsourced to a machine.

Contemporary medicine, despite impressive success, already gets this wrong often by reducing patients to columns of data, to coded diagnoses, and to amounts of medicine to be taken or operations to be done. The human part of the helping profession gets lost over its scientific success. As the medical system happily makes money with this framing error, it remains uncorrected. The process of dying becomes less and less visible in a world of single-person households. The gap widens between the original task of medicine (to help patients in their sufferings) and sad effects of optimization (to provide legal proof that everything possible has been done and was properly billed; to prolong the life of the patient irrespective of personal wishes and life quality). While the topic is more delicate than our short and one-sided perspective, we must acknowledge the problem and the ongoing public debate on the issue.[9] We witness a serious side effect caused by the social and economic reactions to a one-sided scientific approach, which does not solve the original problem as well as its scientific proponents claim.

Contemporary data science may be regarded figuratively as the endeavor to close the gap between the best treatment for cancer and Watson calling Toronto a city in the USA. We shall consider the philosophical position that this attempt is futile at its roots for other than data science reasons. The remainder of this section provides some arguments for this end.

[9]See, for example, Borasio (2016) or Thöns (2016).

4 The Battle Field: Individual Freedom Versus Institutional Optimization

In a world of global optimization, the concept of freedom becomes meaningless. The model of human existence (grow up, educate yourself, find your place in society, provide some meaningful service to your peers, learn to cope with the inevitable sufferings, and die a dignified death) is destroyed. The data scientific endpoint of human evolution is the wheel chair where a neuro interface reads out the mental states of the occupant and provides all the decisions and actions, food and drugs, movement and entertainment, guaranteeing the "best" possible world for a user, who has never met the challenges of his ancestors, who has not learned to cope with difficulties, who has never met the bitter sweet teaching of failures due to an ever-optimized life.

Of course, data scientists are not working explicitly toward this dystopian target! The mechanism is more delicate and well-illustrated by a thought of C. S. Lewis (1972): "Of all tyrannies, a tyranny sincerely exercised for the good of its victims may be the most oppressive. It would be better to live under robber barons than under omnipotent moral busybodies." However, they construct decision algorithms fostering the illusion of "the best."

The debate on digital nudging and the political process ("selling" the "best" options to the constituents without debate as being without alternative[10]) demonstrates that such a development already is taking place in the political arena. The economic sector is more advanced. The business case of the free and informed individual, negotiating the best deal on a level playing field, has been lost. The consumer faces an anonymous digital opponent, which [sic[11]] knows his or her habits, preferences, past choices, financial and mental capabilities, friends and likes, and more. A user interface and a choice of language which is optimized down to minuscule psychometric details influences the emotions and optimizes the maximum financial benefit which can be made from this user. There is no possibility for the individual to deal fairly with an opponent which "owns" billions of psychometric profiles of past shopping interactions and employs optimized persuasion technologies. For example, a project at the University of Liechtenstein aims at discovering those design modifications in an online poll which are best to produce a specific bias in the poll.[12] There is no economic incentive to change this problematic trend.

Our digital future can be described as a feudalistic society where the owners of the "land" (data and algorithms) are the landlords and the data subjects work as their slaves. In the age of enlightenment, Kant taught us to use our own minds and ultimately reject undue dominance over our thinking. In the age of data processing

[10]An astonishingly large number of proposed legislation in the German Bundestag contains the phrase: "Alternatives: None". Decisions are no longer open to democratic debate in a mindset of optimization.

[11]"Which" and not "who": The opponent is a machine, algorithm or web portal, not a human.

[12]See https://www.uni.li/de/thema/information-systems/digital-nudging/digital-nudging-1

we urgently need to regain control over our own data in order not to end up as digital slaves of algorithms and a few anonymous institutions controlling them to our "benefit" (Cap 2017).

Attempts for solutions come in different degrees of practicality. A philosophical approach will promote a new age of digital enlightenment. It will comprise a renewed understanding of the value of freedom and of the importance of liberal-minded ideologies (in the European interpretation of the word, not in the sense of US politics). While academically appealing, this is completely insufficient for practical purposes and needs further implementation in education, legislation, and political measures. Some approaches are outlined by Cap (2016). An important aspect may be to educate the consumer that it is not in his or her monetary interest to offer personal data to companies, since every information on the consumer gets translated into profit-maximization strategies. For example, in B2C commerce the mechanism of cookie pricing makes the ultimate price dependent on search histories, product interests, and past shopping activities of the consumer. If this abuse of asymmetric information relationships between buyer and merchant is made more transparent to the public, market, legal, and governance mechanisms might produce a buyer reaction which reduces this abuse. Manifestos and chartas[13] and similar activities may further raise awareness.

5 The Mistake: Choice of an Incorrect Frame

Framing is the process of selecting a mindset for the semantic interpretation of a concept. The choice of a frame is at the core of every evaluation. For example, taxes may be framed as "heavy burden" or as "valuable contribution to society." The art of "convincing" or, in an alternative frame, "manipulation" often boils down to the choice of a frame (Wehling 2016) suitable to the specific intentions.

Frames which are commonly used to define science may be the naïve "finding out a so-called truth" or, more elaborated, "falsifying hypotheses." Data science technologies and their applications may be described in the frame of an "empirically validated optimal choice." This framing provides a setting which can never, rationally, be rejected. Why would anyone dislike what is best[14] for him? A rejection seems particularly absurd when it is based on empirical evidence which, rooted in world wide data collection, cannot realistically be falsified; when a margin of statistical error can be provided, and is sufficiently small; and when the decision is made by a computer which, per wide-spread belief, cannot err. A critical mind might

[13]User Data Manifesto 2.0 https://userdatamanifesto.org/, the European Digital Charta https://digitalcharta.eu/, the Swiss manifest for digital democracy http://digital-manifest.ch/, or the digital manifesto in Helbing et al. (2015).

[14]We leave aside for a moment the question of who may choose the target function for optimization. This leads to a likewise important debate, which we do not pursue at this place.

become skeptical at a choice of a defining frame which claims immunity against every criticism.

Many jokes "work" by employing a sudden and unexpected shift of the interpretational frames and many human tragedies are caused by sticking to an inappropriate frame. In an ancient Greek allegory King Midas wishes that everything he touches turns into gold. He realizes that he has chosen a wrong frame when not only his furniture but also his wine and his wife turn into solid gold. The defining frame for data science applications (empirically validated optimal choice) contradicts the reasonable frames for human existence and does not go well with concepts of humanity. Which mechanisms stop our society to make the same error as the figurative King Midas, when confronted with the promises of data science? It is, in fact, an interesting paradox. While everybody wants to lead a good life, the perspective of doing so by following the decisions of a machine ultimately is dehumanizing.

Human life is about empathy, about dealing with imperfection and coping with the sometimes-painful limitations of an often-absurd existence. The fundamental strengths of a human being are the ability to cope with this situation, by giving meaning to our life. Almost all productions of human culture, from music to literature and from astronomy to physics, are witnesses of a more or less successful coping with this situation. The frames of "optimal decisions" or of "maximization of profits" may be helpful in a few particular situations, however too broad an adherence to these frames or too successful an implementation of them turns the understanding of human values upside down.

Let us give a vivid example: How would the stereotypic data scientists fall in love? Would match-making algorithms choose their partners? As the efficiency of current algorithms in this field is subject of controversial debates (Tierny 2013), let us conduct this as a thought experiment! How would they *make* themselves fall in love with the person selected by the machine? Would they *really* fall in love or fall for an illusion? What if they felt more for a person the algorithm explicitly warns them of? The western concept of love as individual spontaneous attraction, which hopefully might grow into stable and trusted relationships, conflicts with eastern cultures where partners for life are selected by parents. Which target function should drive the selection process? Who will decide on that? The cultural conflicts between the stereotypic eastern and western partnering processes are replaced by questions on algorithmic parameters. The resolutions of these conflicts form the basis for many a personal fate and, ultimately, collective cultural development. Do we want to settle these conflicts by law, by tradition, or by individual decision? Do we agree to have them settled by Moore's law (Waldrop 2016), when by the technological coincidence fast computers and "intelligent" algorithms provide us with "optimal" choices?

Maybe the frame of optimization is not appropriate for many areas of human existence. If we draw this conclusion, why should we tolerate a creeping—and creepy—development toward this?

In theory and in a free society, the individual may choose differently. In practice, such a divergence faces constant pressure on those individuals who have not led the best life available to them by ignoring to comply with suggestions of the machine.[15] Collectively, many people are not reflecting their lives in philosophical depth but rather follow the simple, nudged choices attractively offered via interfaces optimized to the effect desired by the operator. This trend is not isolated but transforms society into a form where these options of divergence are no longer offered and freedom of choice ultimately vanishes.

Again, the argument must not be parsed one-sidedly. Optimization and improvement *are* necessary; dangerous is the systematic, unreflected, and globalized acceptance of it as a singular dominant trend which is implemented in all processes, products, and work-flows.

6 The Second-Order Problem of Science and Data Science

Second-order problems[16] are problems which cannot be solved with the problem-solving behavior available to the system in which they arise. They are usually met with one or more of the following reactions: Denial that the problem exists, attempts to solve an unsolvable problem, or increasing those efforts which caused the problem in the very beginning (known as "more of the same"—paradox). The common aspect, unfortunately, is: The more a system tries to solve the problem, the bigger it grows the problem. Best intentions combined with an inability to recognize the paradox ultimately constitute the second order problem. A successful solution needs a second order approach and requires deliberation outside of the framework of the original problem. Often this relates to shifting an interpretational frame.

Unfortunately, very successful first-order problem solutions often tend to produce second-order problems. This is particularly true if large-scale systemic effects of first-order solutions are ignored or if only a single methodological approach is taken. The spiral of violence is a phenomenon well known from domestic abuse, and from cultural conflicts up to nuclear armament: If force is an accepted (first order) answer to violence between partners of similar power, this may lead to a spiral of ever-increasing violence with devastating destruction on both sides. The appropriate answer is to refuse using the first-order problem "solution" at the very beginning of the conflict.

By the same reasoning, dominance of the catholic church as single and dominant authority for explaining the world produced a crisis out of which in 1500–1800 AD numerous religious conflicts arose and modern science as a new form of survival evolved. Science and ratio as a method proved to be extremely successful and enabled impressive technological progress and welfare. The unilateral emphasis of

[15] For a more detailed elaboration, see Han (2010).

[16] See Watzlawik et al. (1979) for a systematic approach and Watzlawik (2005) or Watzlawik (2009) for a layman approach to the problem.

science and technology, however, also produced numerous problems. Climate change, destruction of our natural habitat, or the debates on post-truth politics are indicators thereof. Independently of where one stands with respect to these concrete debates, one cannot avoid acknowledging the existence of such a conflict in public debates. Even traditional scientific communities realize that and publish on various forms of inner-scientific problems, which come by the names of replication crisis or publication crisis.[17]

Of course, this certainly does not mean that the scientific approach is flawed and that we should return, for example, to the Bible as the source for scientific understanding—although social, political, and pseudo-scientific movements are growing which promote this goal. It should, however, be taken as an indication that the scientific method may have reached limits in the sense of a second-order problem. More lab experiments and more analyzed data might not be the correct mind-set which helps us out of the crisis. If one is willing to accept this hypothesis, data science points into the wrong direction. It is a first-order solution which is particularly good in aggravating the problem with its "more of the same" strategy.

We now provide a less philosophical, more concrete, well-known example from communication technology: In a world of printed letters, an answer within weeks was acceptable and gave the recipient sufficient time for a well-tuned reply. Then email was invented, leaving only days for an answer. The efficiency of the email solution was so high that ultimately half-day answers were expected. The traditional format of a letter with salutation and complimentary closings was perceived as a burden. Currently even faster formats, such as social network lifelines, messengers, and twitter-like forms of communication, are replacing email. The mere possibility of real-time like answers, often also the expectation thereof, produced a new communicative situation which does not leave room for thinking between the actions. The effects of a twittering president of a super-power on international diplomacy and on stock exchange prices can currently be witnessed.

The original problem—too large latency in communication by postal mail—has been solved too well and has in turn produced a second-order problem. For this, society currently pursues some unsuited first-order solutions. For example, digital non-natives are accused of not coping sufficiently with the new speed at the work place. This line of reasoning may be correct but is not helpful. Despite their ability of fast tweeting, the digital native generation has acquired their own deficiencies: Numerous studies describe a frightening loss of medium- and long-term attention, of the ability to understand, read, and appropriately react on emotions of other people, and even very significantly drop in *the* core human value, that is, empathy.[18]

---

[17]See, for example, Saltelli and Funtowicz (2017). The main reasons they give—and document with a wide range of references—comprise almost all properties frequently found in second-order problems, such as denial of the problem, being a victim of one's own success, no reaction to changes in systemic boundary conditions, flawed incentives leading to misallocations of resources, and more.

[18]A meta-analysis on 72 samples of 13,737 American college students demonstrates this empirically and provides a wealth of pointers into further literature. It identifies as reasons the changes in

Using electronic means of communication less frequently is not a solution either as this amounts to a voluntary decision of being less efficient than one's competitor. A suitable second-order solution remains to be found. The financial system exhibited similar systemic issues since the Lehman crisis,[19] likewise a medical system which focuses on economic efficiency instead of patient relief, as outlined above.

The author is convinced that the essential risks and side-effects of data science have a similar second-order nature: Data science is driven by the vision of the ultimately optimized scientific model, and data technology is driven by the dream of best economic results and most efficient applications. There is no first-order argument why this should be wrong! However, in the end, there is a problem with human values. As is the nature of a second-order problem, the situation fails to be understood with the tools of data science itself.

Starting with technology assessment and ending with Lehman, or the proverbial rise and fall of human values, may seem a helpless exaggeration and a much too big arena for our analysis. However, too small and too detailed a perspective is usually at the roots of second-order problems: Systemic effects are neglected, wide range effects are dismissed as far-fetched, and arguments outside of a narrow scope are perceived as irrelevant or methodologically flawed. Instead, the discourse of analysis should be widened, the pursued goals should be questioned, and the methods and scopes criticized and readjusted. This will not be done by data science itself: Why should data scientists limit themselves in their research, why should data technology companies hold back their run for economic success? It looks like we *do* have a second-order problem at hand.

Striving for improvement and optimal solutions is at the core of human development. However, deviating behavior, individual preferences, and even deficiency is human as well. The proper balance of these two aspects has successfully guided human development for centuries. Data science can destroy this balance.

7 Conclusion

This text fails to deliver technical solutions as they are usually expected in technical papers. Its goal is to raise awareness for a difficult, if not paradoxical, situation through intentionally pointed and painful metaphors. There is a wealth of proposals for quick fixes in other publications,[20] but these are merely band-aid and hide the issues at stake. The data scientist expecting a short "recipe" on how to do things

media and communication technology as well as an increased expectation of success and human optimality. See Konrath et al. (2011).

[19]For a description why financial systems failed in the 2008 crisis not because they ignored best practice but because they *followed* established governance, and for a description of the collective blindness in recognizing self-serving governance mechanisms, see Turnbull (2016).

[20]See, for example, Hofstetter (2016) or Cap (2016).

"right" will be frustrated, as the message is a criticism on one of the mind-sets of data technology: Ubiquitous analysis for pervasive optimization.

The privacy debate produced the "right to be left alone" (Brandeis and Warren 1890–1891) as a consequence of human dignity. Of course, the right should be balanced and must not be understood as an appeal to turn a collaborative society into hermits.

The data science debate must focus on the "*right to be different*" as a "*right to deviate*" from what is considered the optimal choice—whether this choice has been obtained by a majority consensus, by scientific methods, or calculated by an algorithm. Tolerance and the protection of minorities are the values to be saved. This goal is beyond what a particular scientific discipline can achieve and it culminates in the paradoxical insight that an optimal world is bad, or that there is no such thing like a best or true frame for understanding our existence.

The particularly problematic aspect is, of course, that the normative character of objectivized statements and optimized processes is an essential feature for the scientific and technological success of the last 300 years. So how would a "right to deviate" from established norms be implemented without destroying the beneficial aspects of such norms?

We might realize the danger of a slippery slope with a thought experiment. It is well known that algorithms may turn racist by learning a bias from observing humans. Bornstein (2017) describes the case of an algorithm for bail decisions in criminal cases, which learned to discriminate against Afro-American people based on police behavior. But what if an algorithm developed a preference or bias from "facts" "alone"? Would we be willing to accept predictions if they violated our sense of racial fairness? How would we deal with obvious violations of anti-discrimination laws by algorithms? Would we twist the facts? Would we legislate which facts may enter the algorithm? Would we forbid the use of algorithms even though a statistical analysis could tell us the objective "damage" which this prohibition would produce? Of course, every decision system will lead to some form of decision. Only a human sense of fairness produces the distinction between unethical bias and fair prediction, but this evaluation varies in time. It is in no way objective or scientific but reflects our values and depends on cultural, political, and economic conditions.

New and due to data science is our ability to quantify this human sense of fairness. An algorithmic answer seems available to this question "If we dropped our fairness towards group X and allowed our algorithms to discriminate against them then, on a world-wide scale, this would allow an increase in productivity of Y units and in security of Z units." The next question only seems logical. Which combination of X, Y and Z is acceptable from an ethical and economical point of view? Who is going to decide which point of view shall be used? Who will implement the answers and how?

We might follow Edward Teller (1998) who suggested "[that] we must learn to live with contradictions, because they lead to deeper and more effective understanding." The fall of man in physics was the discovery of the nuclear chain reaction, the fall of man in computing may prove to be the discovery of data science methods.

This analogue may point us to a solution. Politicians and generals in the cold war had a hard time learning that in a world with multiple nuclear armies a conflict cannot be reasonably won by using atomic bombs. The ultimate solution is clear: Know the bomb but do not use it!

Similarly, we are in the process of understanding that an economic system depending on percental growth leads to exponential development and exhausts available resources. While this is often abused as a romanticized argument in ideologic discussions, it also is a simple mathematical phenomenon. Although the author is fascinated by the mathematical elegance of science and of data scientific methods, it is his conviction that an indiscriminate application of data science may produce more problems than it solves. This is particularly true if the methods are applied on data related to humans, on human behavior, with a goal of optimizing costs and processes where humans are closely involved or in a situation where our human understanding of complex world phenomena are involved. Contrary to the observation on percental growth, this conviction is not a mathematical phenomenon but an assumption based on a personal observation of human nature and greed, so it is difficult to reach an agreement on it with formal reasoning alone. The short form of a solution is somewhat like in nuclear war technology: Study data science but do not apply it!

As in nuclear technology we need a debate on the details. We have learned to use nuclear engineering to make war, to produce energy, to diagnose, and to treat health problems—and we have been working on an understanding which applications are acceptable. Data science needs a similar debate. I expect this debate to be long and tedious. I am not sure which force will prove stronger—greed, rationality, or humanitarianism.

References

Anderson, C. W. (1989). Learning to control an inverted pendulum using neural networks. *IEEE Control System Magazine, 9*(3), 1989.

Borasio, G. (2016). *Selbstbestimmt sterben*. München: DTV Verlagsgesellschaft.

Bornstein, A. M. (2017). Are algorithms building the new infrastructure of racism? *Nautilus, 55*. http://nautil.us/issue/55/trust/are-algorithms-building-the-new-infrastructure-of-racism.

Brandeis, L., & Warren, S. (1890–1891). The right to privacy. *Harvard Law Review, 4*, 193–220.

Braschler, M., Stadelmann, T., & Stockinger, K. (2019). Data science. In M. Braschler, T. Stadelmann, & K. Stockinger (Eds.), *Applied data science – Lessons learned for the data-driven business*. Berlin: Springer.

Cap, C. H. (2016). Verpflichtung der Hersteller zur Mitwirkung bei informationeller Selbstbestimmung. In M. Friedewald, J. Lamla, & A. Roßnagel (Eds.), *Informationelle Selbstbestimmung im digitalen Wandel*. Wiesbaden: Springer Vieweg, DuD-Fachbeiträge.

Cap, C. H. (2017). Vertrauen in der Krise: Vom Feudalismus 2.0 zur Digitalen Aufklärung. In M. Haller (Ed.), *Öffentliches Vertrauen in der Mediengesellschaft*. Köln: Halem Verlag.

Dwork, C., & Mulligan, D. (2013, September). It's not privacy, and it's not fair. *Stanford Law Review*. https://www.stanfordlawreview.org/online/privacy-and-big-data-its-not-privacy-and-its-not-fair/

Esguerra, R. (2009, December 10). *Google CEO Eric Schmidt dismisses the importance of privacy*. Electronic Frontier Foundation. https://www.eff.org/de/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy

Executive Office of the President. (2016, May). Big data: A report on algorithmic systems, opportunity, and civil rights. *The Whitehouse*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Friedman, L. (2014, April 22). IBM's Watson supercomputer may soon be the best doctor in the world. *Business Insider*. http://www.businessinsider.com/ibms-watson-may-soon-be-the-best-doctor-in-the-world-2014-4

Galeon, D. (2016, October 28). IBM's Watson AI recommends same treatment as doctors in 99% of cancer cases. *Futurism*. https://futurism.com/ibms-watson-ai-recommends-same-treatment-as-doctors-in-99-of-cancer-cases/

Grassegger, H., & Krogerus, M. (2016, December 3). Ich habe nur gezeigt, dass es die Bombe gibt. *Das Magazin* Nr. 48. https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/

Han, B.-C. (2010). *Müdigkeitsgesellschaft*. Berlin: Matthes & Seitz.

Helbing, D., et al. (2015). *Eine Strategie für das digitale Zeitalter*, 12.11.2015, Spektrum Verlag, http://www.spektrum.de/kolumne/eine-strategie-fuer-das-digitale-zeitalter/1376083

Hofstetter, Y. (2016). *Sie wissen alles: Wie Big Data in unser Leben eindringt und warum wir um unsere Freiheit kämpfen müssen*. München: Penguin.

Hosain, S. Z. (2016). Reality check: 50B IoT devices connected by 2020 – Beyond hype and into reality. *RCRWirelessNews*, 28th June 2016. http://www.rcrwireless.com/20160628/opinion/reality-check-50b-iot-devices-connected-2020-beyond-hype-reality-tag10

Konrath, S., O'Brien, E., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time. *Personality and Social Psychology Review, 15*(2), 180–198.

Kree, I., & Earle, J. (2013, September 2013). Prediction, preemption, presumption: How big data threatens big picture privacy. *Stanford Law Review*. https://www.stanfordlawreview.org/online/privacy-and-big-data-prediction-preemption-presumption/

Lewis, C. S. (1972). *God in the dock. Essays on theology and ethics*. Grand Rapids: Eerdmans Publishing.

Marr, B. (2015, September 28). *The biggest risks of big data*. http://www.datasciencecentral.com/profiles/blogs/the-biggest-risks-of-big-data

Miller-Merell, J. (2012, September 27). Hiring by algorithm. The new self-checkout of HR. *SmartRecruiters Blog*. https://www.smartrecruiters.com/blog/hiring-by-algorithim-the-new-self-checkout-of-hr-2/

Nichols, S. (2017, January 7). TV anchor says live on-air 'Alexa, order me a dollhouse'. *The Register*. https://m.theregister.co.uk/2017/01/07/tv_anchor_says_alexa_buy_me_a_dollhouse_and_she_does/

O'Connor, A. (2011, February 15). Watson dominates Jeopardy but stumbles over geography. *New York Times*. https://artsbeat.blogs.nytimes.com/2011/02/15/watson-dominates-jeopardy-but-stumbles-over-geography/

Park, J., Seo, D., et al. (2014). *Practical human resource allocation in software projects using genetic algorithm*. In SEKE 2014, pp. 688–694. https://pdfs.semanticscholar.org/c317/a294cb7995f5ca7799ab1f7f878e7b0a4749.pdf

Root-Bernstein, R., & Root-Bernstein, M. (1999). *Sparks of genius – The thirteen thinking tools of the world's most creative people*. Boston: Houghton Mifflin.

Saltelli, A., & Funtowicz, S. (2017). What is science's crisis really about? *Futures, 91*, 5–11.

Schönholzer, F. (2017). Digitale Pfadfinder. *UZH News*, 14.11.2017. http://www.news.uzh.ch/de/articles/2017/Digitale-Pfadfinder.html

Shattuck, R. (1997). Forbidden knowledge: From prometheus to pornography. *Harvest Book*. Here quoted from Wikiquote https://en.wikiquote.org/wiki/Edward_Teller

Teller, E. (1998). Science and morality. *Science, 280*(5367), 1200–1201.

Thöns, M. (2016). *Patient ohne Verfügung: Das Geschäft mit dem Lebensende*. München: Piper.

Tierny, J. (2013, February 11). A match made in the code. *The New York Times*. http://www.nytimes.com/2013/02/12/science/skepticism-as-eharmony-defends-its-matchmaking-algorithm.html

Turnbull, S. (2016). Defining and achieving good governance. In G. Aras & C. Ingley (Eds.), *Corporate behavior and sustainability*. New York: Gower.

Waldrop, M. (2016). The chips are down for Moore's law. *Nature, 530*, 144–147.

Warriar, L., Roberts, A., & Lewis, J. (2002). 2002- *Surveillance – An analysis of Jeremy Bentham Michel Foucault and their present day relevance.* http://studymore.org.uk/ybenfou.htm

Watzlawik, P. (2005). *Vom Schlechten des Guten oder Hekates Lösungen*. München: Piper.

Watzlawik, P. (2009). *Anleitung zum Unglücklichsein*. München: Piper.

Watzlawik, P., et al. (1979). *Lösungen. Zur Theorie und Praxis menschlichen Wandels*. Bern: Hans Huber.

Wehling, E. (2016). *Politisches Framing: Wie eine Nation sich ihr Denken einredet – und daraus Politik macht*. Köln: Halem.

Weizenbaum, J. (1976). *Computer power and human reason: From judgement to calculation*. New York: Freeman.