Ulrich Frank

Language, Change, and Possible Worlds

Philosophical Considerations of the Digital Transformation

> *The transformation we are concerned with is not a technical one, but a continuing evolution of how we understand our surrounding and ourselves – of how we continue becoming the beings that we are.*
> – Terry Winograd and Fernando Flores: *Understanding Computers and Cognition: A New Foundation for Design*

Abstract: This essay aims at the identification and discussion of specific methodological problems related to the study and the support of the digital transformation from the perspective of a discipline that is directly responsible, both as an observer and a driver: Business Informatics. After a short account of the history of information technology and its present constitution, a brief analysis of key aspects of the digital transformation will lead to the pivotal role of language and conceptual models in particular. To prepare systems for change, models need to incorporate abstractions that cover possible future worlds. From an academic perspective, this requirement is fascinating and challenging at the same time. It implies the problem of how to adequately justify the construction of possible worlds. Furthermore, it leads to the question of how we can think possible future worlds that are beyond the limits of the language we speak. Against this background, the paper proposes facets of a methodology of change.

Keywords: Artificial intelligence, contingency, induction, justification, language, narrative, research method

1 Introduction

An ever-growing amount of our surroundings is being represented digitally. It does not seem exaggerated to even state: "The world is being rebuilt in code" (Widdicombe 2014: 56). Digitization of this kind does not just mean to create images of a given reality, or to automate tasks previously performed by humans. Instead, it means to open new perspectives on how to conceptualize our surroundings, to reframe familiar patterns of work, of collaboration and of communication. It allows us to overcome traditions and concepts we have been used to for a long time, and to see new options unknown to us in the past. These new options enabled the emergence of new companies that became global giants in less than a decade. At the same time, industries that prospered only a few

https://doi.org/10.1515/9783110702255-009

years ago, lie shattered, swept away by a process of change that is unprecedented in the history of humankind. Therefore, it seems appropriate to think of it as a fundamental transformation or even as a paradigm shift. While it is not clear what the essential characteristics of this change are and what it will bring about, there is a name for it that has been a dominating theme both in academic and public discourses for some time: digital transformation. Even though the omnipresent narrative of the digital transformation suffers from multiple simplifications and dubious contributions, it represents a phenomenon of substantial relevance that will likely change the way we work, live, and think. It is fascinating and ambivalent. It is about opportunities and threats, about creation and decline, about construction and deconstruction, about liberation and domination.

The digital transformation represents a manifold, complex phenomenon that bulks against simple explanations. Nevertheless, it seems not too daring to attribute it to two main driving forces. At first, the Internet provides a common infrastructure that does not only promote the accessibility of digital resources independent from their physical location, but also allows to get access to physical resources, organizations, and people that were beyond our reach in the times before the raise of the Internet. The economic effects of the Internet are tremendous. It enables new services, new business models, promotes bundling of resources, and boosts economies of scale. It is the foundation of new forms of social interaction and promotes the dissemination of knowledge at a previously unknown level. Finally, the Internet paved the way for the second driving force, the growing availability of mass data on almost any aspect of business transactions, of social interaction, and of human life in general. The availability of mass data together with growing computing power enables an extensive use of methods of inductive statistics at relatively low costs. In addition to analyzing data for patterns of correlation which are of use for decision making purposes, approaches to so-called “machine learning” have received particular attention in recent times and contributed to the creation of a remarkable new hype about *artificial intelligence* (AI). Not only has machine learning already produced various impressive software systems, it also holds out the prospect of pushing the limits of automation. Since the functionality of these systems is hard to explain, it is not surprising that they become subject of mystification, which goes along with the emergence of utopian and dystopian scenarios of the future, with promises and warnings.

This is a worrying situation. Decision makers in politics and in organizations are under pressure to invest into research and technology, even though many lack the required appreciation of the subject. This is a clear threat to the idea of rationality and enlightenment. It reaffirms Habermas’ critique of technology and science as “ideology” (Habermas 1976) and goes beyond it at the same time: those who (re-)produce this ideology may well become its victims if the assumptions it is based on turn out to be wrong. The remarkable public attention created by the digital transformation and AI does not leave academia unaffected. This is especially the case for disciplines that are directly related to technological and institutional aspects of the digital transformation, like Computer Science and Business Informatics. They benefit from the availability of growing research funds, especially for topics related to AI. At the same time, they also suffer from the growing demand for graduates in the industry because that aggravates filling research positions. In addition, researchers sometimes face a subtle conflict. On the one hand, researchers may benefit from the hype, since research funds are often motivated by daring promises created through the hype. On the other hand, academia should maintain a critical attitude which may lead to unmasking the hype as such.

Apart from that, the digital transformation provides fantastic opportunities, and that also means: formidable challenges, for a wide range of academic studies. This is mainly for two epistemological reasons. First, understanding a system in general is hardly possible without studying how it behaves during change. Often, change will be limited to gradual modifications. The digital transformation, however, goes along with forces that may require radical change in a short period of time. Studying this kind of extensive change provides the opportunity to better understand fundamental system properties such as resilience, adaptability, and consistency. Since the digital transformation concerns interwoven technical, social, and psychological systems, it also allows for studying commonalities, differences, and interdependence of these different kinds of systems, thus contributing to the development of trans-disciplinary knowledge. Second, studying the digital transformation opens a path to an intriguing intellectual adventure. It is like a journey into unknown territory, a territory that is yet to be constructed with measures we do not know of, and that will confront us with questions that were never asked before.

Already today, the digital transformation leads to a plethora of challenging questions. Some of these questions are subject of this essay:

- What are typical drivers of the digital transformation and of automation in particular?
- What are essential differences between traditional programming and machine learning?
- Are traditional research methods still appropriate to study the digital transformation and to develop meaningful and attractive orientations for change?
- What is the role of academia in this process?
- What are convincing options to prepare for future change beyond prediction and prescription?

- In this respect, what is the role of models, theories, and of language?
- Is it possible to develop scientifically grounded orientations for change without transforming academia itself?

To discuss these questions, I will build on my primary academic education in Business Informatics, which mainly relates to the creation and use of information technology in organizations. However, such a specialist perspective is not sufficient to cope with essential facets of the digital transformation. While promoted predominantly by innovations in information technology, the transformation is, at its essence, related to problems that have been the subject of philosophical discourse for a long time. Among others, those problems comprise the ambivalent role of language for recognition as well as for mastering change, the limitations of truth as the pivotal criterion to justify scientific knowledge offerings, or the role of science within processes of social and political change. Therefore, I will, even though not a philosopher by education, dare to enrich the analysis of the digital transformation with a philosophical perspective. I beg the indulgence of those readers who are professional philosophers if my arguments appear too superficial.

2 Two principal approaches to automation

The last six decades are characterized by an ever-growing amount of manual work being replaced by software. To better understand this continuing process of automation, it is advisable to analyze general preconditions and objectives of software development. The class of problems to be solved by a software system should be formalized to a certain degree. That involves the data, the software operates on as well as the operations itself. Formalization requires the definition of syntactical and semantic rules that constrain the range of valid representations. General design objectives include integrity, reuse, integration, and adaptability. Integrity means that software should prevent system states that are not consistent with the specification, in other words: states that violate syntactic or semantic constraints. As we shall see, there are remarkable differences between traditional software development and machine learning in that respect. Reuse is, on the one hand, motivated by the demand for integrity. Based on the assumption that professionally developed software artefacts are available, it is valid to conclude that their reuse within the development of a particular software system will contribute to that system's integrity. On the other hand, reuse is the pivotal measure to reduce costs. First, it allows reducing development costs, since developers are not forced to always start from scratch. Second, it enables the reduction of costs per single copy if the software system is being used by many. Integration will often be an important design objective, too. To enable the integration of two systems, they need to have a common semantic reference system (cf. Frank 2011), that is, common data structures or common functions and events. Finally, software is expected to be adaptable since the requirements it is supposed to satisfy will often be subject of change. The more conveniently and safely the relevant modifications on software can be performed, the higher is its adaptability. The following two sections serve the purpose to reconstruct two principle approaches and compare them with respect to the prerequisites of automation and the objectives of software construction.

2.1 Reduction of contingency or adaptation to the limitations of machines

Software development requires unambiguous descriptions of problem classes that can be expected to be widely invariant across all present and future use cases. Hence, if propositions characterizing a problem class are true in one context and wrong in another, the automation of reliable problem-solving processes is not feasible. In addition, automation requires representations of data that are readable by machines in an unambiguous and unified way. The history of data processing is essentially characterized by a process of stepwise reduction of contingency. Note that is not relevant here, whether contingency is seen as an ontological property or rather the reflection of an epistemological limitation. The use of punch cards and similar media in the early days of data processing enabled machines to read symbols. In addition, it was required to make sure that the data represented on these media satisfied certain syntactical rules. Furthermore, there was need to reduce ambiguity by defining the (formal) semantics of data. The introduction of data types like String, Integer, etc. addresses this requirement. However, providing for machine readable media was not enough. In addition, the entire task, e. g., payroll accounting, has to be reorganized. Furthermore, the syntactical and semantic diversity of representations and problem-solving approaches needs to be targeted, too. Standardization is a pivotal instrument to reduce this kind of contingency and to promote economies of scale. Standardization also serves the integration of software systems: if they comply with certain standards regarding, e. g., data structures, they are enabled to communicate. Furthermore, standardization fosters economies of scale and protection of investment. In any case, reduction of contingency through the introduction of standards requires adapting problems and problem-solving procedures. In the past, this did not only comprise the reorganization of business processes, but also the reconstruction of entire new business models that are tuned to exploit the potential of software and IT in general.

While automation was possible only because reduction of contingency paved the way, this approach is not without downside. In general, it restricts the freedom of implementing individual solutions. In particular, preparing representations for machine readability may compromise their expressive power and, as a consequence, the quality of communication. A typical example is the use of software to handle customer requests. Standardization may be a threat to differentiation, and, hence, to competitiveness. It may also prevent services that are tailored to specific needs of particular customers. Standardization means to freeze a certain convention. Therefore, it may well be an obstacle to progress because the costs to deviate from a standard will often be prohibitively high.

The Internet is a special case. Its tremendous success seems to be based on both, the reduction of contingency through standardization, and the waiver of rules. On the one hand, the Internet is characterized by an enormous reduction of contingency. Among other things, this reduction is realized through the construction of a unified global address space for resources, organizations, and people that is enabled by a standardized technical infrastructure. As a consequence, the Internet enabled a tremendous wave of automation. On the other hand, the World Wide Web allows for almost total individual freedom regarding the representation of data. Furthermore, the set-up costs for representing data on web pages are relatively low. The lack of constraints promoted the global dissemination of the web. While the web enables the digital representation of the world to an unprecedented extent, it comes with clear downsides as far as automation is concerned. The data represented on web pages will often be of extremely contingent nature. Not only that it lacks semantics, furthermore, it will often be unclear whether it is consistent and up to date. The next approach to promote automation is supposed to deal with contingency, that is to leave contingent representations as they are and let machines cope with them.

2.2 "Intelligent" machines that cope with contingency or the pivotal role of data

When the term "artificial intelligence" was coined in the fifties of the last century, it was motivated by the idea of representing the cognitive capabilities of humans on a digital computer. The early enthusiasm created by this prospect was soon replaced by growing frustration and pragmatic adjustments of the original research objectives. No longer was it the claim to develop machines that can think like humans. Instead, the focus was directed toward the formalization of knowledge (cf. Frank 1988). This second wave of AI research was based on three central assumptions. First, qualified knowledge, that is expert knowledge, is pivotal for problem solving. Second, the application of knowledge requires basically logical operations. Third, an expert's knowledge can be formalized. The conception of so-called "knowledge-based" or "expert" systems followed these assumptions. Even though some of these systems achieved remarkable problem-solving capabilities, they did not fulfill the expectations they had created. That was mainly for two reasons. First, it turned out the human problem-solving competence can often not sufficiently be reconstructed with formalized knowledge. Second, these systems did not extend the limits of automation. Like any other software system, they require the specification of solution spaces or, in other words, they allow the automation of tasks only if the problem descriptions they are directed at are free from contingency. However, typically, not necessarily, knowledge-based systems feature a declarative representation of knowledge, which allows for monotonic extensions. Therefore, these systems are in general easier to maintain than procedural program code.

The third wave of AI, even though not well-defined as such, is characterized by a different approach. It does not necessarily require reducing the contingency of a problem representation. Related to that, software that falls into this category, does not have to be specified or programmed manually. Instead, software can be generated through a "learning" process that is based on "training" the structure of a network to gradually produce results from input data that correspond to what is expected from a satisfactory solution. The algorithms that are generated through this kind of "machine learning" (Murphy 2012) are based on induction. Therefore, the availability of mass data is a crucial prerequisite. With a huge number of people leaving their footprints in the net, and a growing armada of sensors, more and more facets of the world are represented by data – which, of course, does not tell much about the quality of this representation. Once a system has been trained sufficiently to produce satisfactory results for a given set of certain input data, it is assumed that it will work for other data of sets of that kind. Apart from the promise that machine learning will reduce software development costs dramatically, it also enables the automation of tasks without a precise specification of the problem or of satisfactory solutions. In other words: it promises to enable software that can cope with contingency, which would substantially extend the limits of automation. Examples of successful and in part very impressive uses of machine learning include natural language translation, face recognition, analysis of medical imaging systems, or so-called autonomous driving.

The success of these systems has fostered remarkable enthusiasm. One of the leading researchers in the field of machine learning claims: "All knowledge – past, present, and future – can be derived from data by a single, universal learning algorithm" (Domingos 2017: 25), "[...] and as a result of all this, our lives will be longer, happier, and more productive" (Domingos 2017: 43). A growing number of startup companies aim at implementing business models based on the promises made by machine learning. Managers regard machine learning as a "game changer" that companies need to take advantage of in order to survive (see, e. g., the survey in Sinclair/Brashear/Shacklady 2018). Politicians are eager to promote research and development of machine learning because it is regarded as a key factor for a national economy's competitiveness (see, e. g., Bundesregierung 2018). Authors of popular literature do not hesitate to make predictions of the economic and ethical impact of machine learning and to propose ideas how to cope with them (e. g., Boddington 2017; Precht 2018). While it is disturbing that many actors, including academics, make affirmative statements about the prospects of a technology most of them do not sufficiently understand, the role of machine learning within the digital transformation must not be underestimated. This is not only because machine learning opens new alleys for automation, but also because it inspires the imagination of millions and their reflection about the future. We will take a closer look at prospects and challenges of machine learning in the next section.

3 Methodological challenges

Given the tremendous relevance of the digital transformation, it seems reasonable to ask how science could support it. If we assume that the transformation will leave us with options, predictions that could be developed by studying the past are hardly possible. In other words: focusing on the factual is not sufficient, instead there is need to analyze the possible. Not surprisingly, such an endeavor leads to serious methodological problems. They are also caused by the specific nature of software, since future worlds are more and more penetrated by software, or even (re-)constructed by software. Therefore, I will at first analyze what software essentially is, and how it is developed and used.

3.1 The pivotal role of language

Software is an immaterial artefact. From a formal perspective, it can be regarded as an abstract machine that is defined by a set of operations to transform input data into output data. From an engineering perspective, software is realized through the execution of operations of a processor that has read and write access to some kind of data memory. While these views of software are important for formal analysis and the construction of computers, it is not sufficient to adequately describe software as a tool for humans. Software can represent a domain of interest and provide support for users only if it is supplemented with a linguistic representation that is accessible by humans. For that purpose, it is required to map formal structures and operations to concepts that correspond to those known by its prospective users. This kind of mapping is done through designators of software artefacts that refer to concepts known in the targeted domain. In other words, for software to be usable, it is mandatory that is has a conceptual foundation. Otherwise we cannot make sense of it. This is the case for the development of software as well as for its use. Conceptual models are of pivotal relevance when it comes to support the design of a coherent and consistent conceptual foundation. They aim at clear presentation of the concepts that are needed to understand and use the software. To reduce complexity, conceptual models will often focus exclusively on static, functional, or static aspects of software. Conceptual models are created with specific modeling languages the concepts of which should allow a clear mapping to constructs of implementation languages. Figure 1 shows a simple conceptual model and illustrates how it is mapped to code.

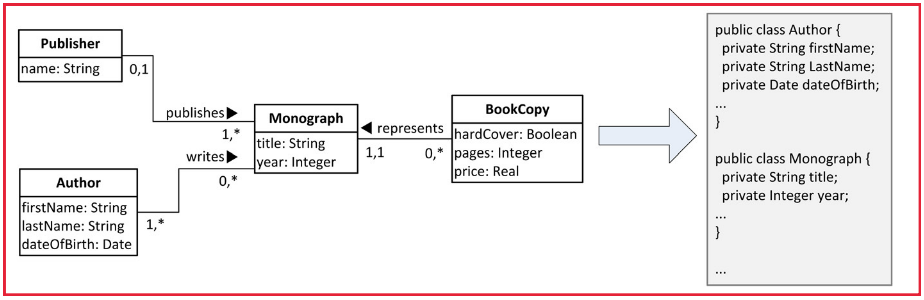

Fig. 1: Example of a conceptual model and mapping to code. (Credit: Ulrich Frank)

Note that it is not mandatory to develop an explicit conceptual model as it is shown in Figure 1. It is conceivable that programmers work with implicit conceptual models and represent them rudimentarily in code through designators. It is not only mandatory to conceptualized software when it is developed. If software is not supplemented with a representation that corresponds to concepts known in the relevant domain, it is not possible to use it.

It is not trivial to develop comprehensive and consistent conceptual models. Exceptional requirements may be overlooked. The possible variety of a subject may be inappropriately assessed. For example, in the simple model in Figure 1 a monograph is published by one publisher at the most. However, there are rare cases, where more than one publisher takes care of the publication of a monograph. If a conceptual model does not appropriately represent a domain, it is likely to result in software that does not fulfill the requirements. With respect to the economics of software, its reuse in many particular cases is of pivotal relevance in order to achieve economies of scale. However, that creates the challenge that the variety of requirements across the set of intended use cases needs to be accounted for. To that end, it is required to find, or construct, abstractions, that is, concepts, that are appropriate for the entire range of intended applications, and that allow for convenient and safe adaptation to more specific requirements. Especially in those application areas that are characterized by remarkable variety of requirements the quest for a wide range of reuse is a considerable challenge. A common strategy to cope with contingent requirements is to reduce contingency. This can be done by reducing or reconstructing a problem to fit it to the capabilities of computers: "Many of the problems that are popularly attributed to 'computerization' are the result of forcing our interactions into the narrow mold provided by a limited formalized domain." (Winograd/Flores 1986: 75) Similarly, the variety within a range of intended use cases can be reduced by creating incentives for users to adapt their requirements to a given software. This can be achieved through attractive terms and conditions which are enabled by economies of scale.

From an academic perspective, software construction is even more demanding, because it is not sufficient to ask whether requirements are satisfied. Furthermore, it is required to reflect upon the impact that the technical language of a particular domain has on the identification of problems, the organization of work, or for developing a satisfactory understanding of the domain.

> An awareness of one's own vocabulary is the first step to questioning it with a design attitude and exploring how different vocabularies yield more creative problem representations and enable the development of better designs. (Boland/Collopy 2004: 15)

That leads to the question whether the language we use to conceptualize a domain and to design software systems is adequate. The concepts that are identified during an analysis of a domain might be approved by domain experts. But that does not mean they are an appropriate foundation for structuring the domain, for identifying problems, for organizing work or for developing a satisfactory appreciation of the domain. The example in Figure 1 illustrates these aspects. Without resolving the ambiguity of a term such as "book," by distinguishing between monograph and printed book, it would hardly be possible to develop a consistent information system. At the same time, it is obvious that the concepts presented in the model are not sufficient to cover the variety of publications, since they do not account for edited books, journals, articles, etc.

Therefore, it is not only hard to tell whether requirements are complete, but furthermore whether they are appropriate since they are a reflection of concepts that might have been defined differently. The pivotal and delicate role that language plays for the analysis of requirements, as well as for the construction and use of software, becomes even more apparent in the light of the digital transformation.

3.2 In search of new languages

If we assume that the world is more and more constructed through software, it is essential to account for the role of software in times of change. If we further assume that the digital transformation creates new opportunities for the design of products, for the organization of work, or for the arrangement of social interactions, in other words: for new possible worlds, we are confronted with extraordinary methodological challenges. On the one hand, they relate to the evaluation and justification of our constructions. In many disciplines, justification is preferably based on a neo-positivist, "evidence-based" approach that reflects the correspondence theory of truth. However, developing possible future worlds that may serve as an orientation for change can hardly be tested against "reality," because they intentionally deviate from the "factual." In addition, the design of possible future worlds will involve value judgments. On the other hand, the challenges relate to the limits of recognition. Our primary tool to conceive of a possible future is the language we speak. However, at the same time, language limits our imagination, that is, the world we can conceive of (TLP 5.6). This limitation is necessary to cope with complexity and to establish sense. At the same time, we need to be aware of it if we do not want to give up the quest for a critical attitude:

> Our view is limited to what can be expressed in the terms we have adopted. This is not a flaw to be avoided in thinking – on the contrary, it is necessary and inescapable. Reflective thought is impossible without the kind of abstraction that produces blindness. Nevertheless we must be aware of the limitations that are imposed. (Winograd/Flores 1986: 97)

Being aware of this epistemological limitation is a necessary, but not sufficient prerequisite of relaxing it. Against this background, designing a possible future world is confronted with multiple contingencies and overwhelming complexity. Since there are many possible futures, we need to develop a space of possibilities that comprises those options that appear, for convincing reasons (!), to be the most desirable ones. While that requires a new language, it is not sufficient to follow Wittgenstein's advice: "We want to establish an order in our knowledge of the use of language: an order with a particular end in view; one out of many possible orders; not *the* order." (PI § 132) Instead, the choice or construction of a proper language is essential, or, as Rorty put it: "Philosophers have long wanted to understand concepts, but the point is to change them so as to make them serve our purposes better." (Rorty 2000: 25)

But how can we tell how this new, better language should look like, even if we agree with Rorty that it should be suited to foster "democratic politics" (Rorty 2000: 25)? How could we decide for a language that enables us to conceive a possible future world if both the language and the possibilities of the future are beyond our imagination? In any case, this constitutes an epistemologically extremely risky, if not hopeless situation that Derrida characterizes as an "absolute danger":

> The future can only be anticipated in the form of an absolute danger. It is that which breaks absolutely with constituted normality and can only be proclaimed, presented, as a sort of monstrosity. For that future world and for that within it which will have put into question the values of sign, word, and writing, for what which guides our future anterior, there is as yet no exergue. (Derrida 1976: 5)

3.3 Prospects and limitations of induction

The enthusiastic promises that accompany the third wave of AI are based on the power of induction and the availability of mass data. The inductive analysis of huge amounts of data may not only reveal patterns shared by many objects, and, thus, increase our knowledge about these objects. In addition, induction may also serve the automated construction of algorithms that enable the transformation of input data toward an intended result. Since induction is also relevant for human learning, especially in early age, it promotes the idea of machines that gradually learn from data to train their "neural" networks to an ever-growing level of "intelligence" that will eventually match or surpass human capabilities. Consequently, AI might be suited to automate scientific research, thus contributing to the growth of human recognition. Pentland, who already predicts the replacement of traditional social sciences by data driven "social physics," outlines an age of groundbreaking scientific achievements:

> For the first time, we will have the data required to really know ourselves and understand how our society evolves. By better understanding ourselves, we can potentially build a world without war or financial crashes, in which infectious disease is quickly detected and stopped [...] and in which governments are part of the solution rather than part of the problem. (Pentland 2014: 18 – 19)

Based on a similar assessment, Anderson predicts the "end of theory" and of "the scientific method" (Anderson 2008). There is no doubt that induction is suited to uncover facets of the world unknown to us. Induction has been applied to the analysis of, e. g., customer or voter behavior, to the assessment of applicants, or to the translation of natural languages – frequently with impressive results. However, induction as well as its limitations have been known for long. In machine learning, induction is a mechanical process of discovering common patterns. From an epistemological point of view, this is not satisfactory, because induction as pattern detection does not offer a convincing explanation. Rescher who appreciates induction as an instrument of inquiry, therefore proposes to regard induction as an act of "responsible estimation": "it is not just an estimate of the true answer that we want, but an estimate that is sensible and defensible: tenable, in short" (Rescher 1980: 9). Therefore, it does not seem appropriate to speak of inductive inference. Possible conclusions are "not derived from the observed facts, but invented in order to account for them" (Hempel 1966: 13).

Nevertheless, the promises of machine learning may appear intriguing. With large amounts of the world being digitized, and the availability of tremendous computing power, it seems conceivable that all regularities, both static and dynamic, that exist can be discovered by machines. That would leave us with the challenge to somehow justify these estimations. This problem could be addressed in a pragmatic way, that is, by redefining the concept of rational justification. That seems to happen already, when decision makers justify their decisions with patterns produced by inductive analysis of mass data. From an academic perspective, that would hardly be convincing. There is already growing awareness of the problem, which is, among others, expressed in the emergence of a new field of study named "explainable Artificial Intelligence" – a term that reflects a massive criticism of AI research, because it indicates that it lacks an essential part of any scientific knowledge offering, namely justification.

It goes beyond the scope of this essay to investigate the potential of automated induction. In addition to epistemological aspects that would also require accounting for economic and ethical issues. To give one example only: it may seem acceptable that the prediction of consumer behavior fails in some cases, however, this is likely to be different with software to enable "autonomous" driving. Apart from that, the potential of machine learning to replace human inquiry is questionable for two important reasons. First, only those phenomena can be subject of inductive analysis that allow for an appropriate digital representation and are suited for being investigated by machines. That, of course, relates to the old methodological conflict between neo-positivist and hermeneutic approaches. Second, with respect to develop images of possible future worlds in order to provide an orientation for change, induction suffers from a principal restriction. It depends on the analysis of data about the factual world. If we assume that the future is not entirely determined by historical path dependencies, relying on induction would clearly limit our options. Domingos does not deny that, even though he does not regard it as a problem: "We're only interested in knowledge about our world, not about worlds that don't exist" (Domingos 2017: 25).

3.4 Recognition and decision making without concepts?

Machine learning is not restricted to the inductive analysis of data that are supplemented with a conceptual definition. There are various approaches to machine learning that lack a conceptual foundation of this kind. Often, these approaches are referred to as "sub-symbolic," (see, e.g., Wichert 2013) which is misleading, because any representation on digital machines has to be symbolic. Solving problems without concepts is in clear contrast not only to the prevalent conception of science, it is furthermore hardly conceivable. Typical examples of this kind of machine learning include image and, in particular, face recognition, but also the detection of tumors. The simplified example in Figure 2 illustrates the principal idea. A traditional approach to software engineering would recommend a conceptualization of the problem, e.g., by introducing concepts like "eye," "nose," etc. Different from that, machine learning is based on digital representations of the images that do not include any conceptual information. After redundant data is removed, the resulting data sets serve as input to a "neural" network, which is trained until its weights are adjusted to produce the same characteristic number for all pictures of that person. If the network is to decide whether a further picture represents the same person, it compares the resulting number to the one that was produced during the training phase. If the number is close enough, it would be concluded that it is the same person. If the result is not satisfactory, the network could be refined through further training data.

The advantage of this approach is obvious. It does not only allow for a substantial reduction of development costs, it may even enable automation in cases where human developers would not be able to cope with the complexity of a

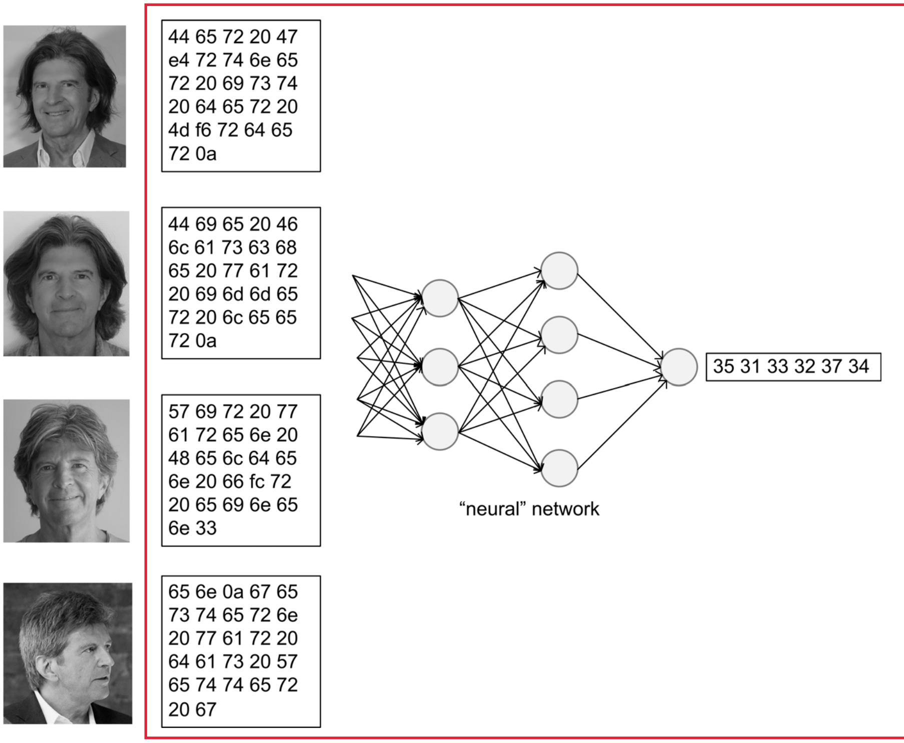

Fig. 2: Illustration of face recognition by a neural network, inspired by Murphy (2012). (Credit: Ulrich Frank)

problem. However, this advantage comes at a price. It cannot be guaranteed that the results will always be correct. Depending on the requirements of specific use cases, that may be acceptable or not. It is, in a way, similar to human perception: we can hardly explain it, and it may be deceptive. Nevertheless, it is extremely useful. However, from an academic perspective, it creates a serious problem. Since machine learning is part of university curricula, it should be based on justified recognition. If that is not the case, the approach of choice to evaluate neural networks are benchmarks, that is, testing the performance of a system against given data sets. If these systems perform sufficiently well, it is likely that they are used, and that the decisions they suggest are followed. Such a scenario is critical, because it is suited to shatter the idea of rationality and of enlightenment by giving up the quest for justification.

4 In a nutshell: A methodology of change

As we have seen, the methodological challenges that come with the digital transformation are enormous. One could only avoid them by deciding that research is not responsible for designing images of the future. However, taken the huge importance of the digital transformation for society as a whole, and also for the conception of science, that is hardly a convincing option. That leads us to the question how research methods could look like that are suited to develop an orientation for the digital transformation, or, if there is even need for a methodology of change. The following two sections give a brief outlook of possible features of such a methodology (a more elaborate, but still preliminary description can be found in Frank 2017).

4.1 Construction of possible future worlds

If the construction of possible future worlds is regarded as a scientifically grounded offering, prevalent research methods are of limited use only. They usually rely on some conception of truth and certain procedures to check the truth of propositions. In its purest form, scientific knowledge is offered as theories. However, as truth is not sufficient to justify possible future worlds, common concepts of theory are not suited to capture our imagination of the future. Giving up on truth or, at least, relaxing its pivotal role in the justification of knowledge, suggests looking for a new kind of methodology. I am reluctant to follow the radical turn that Rorty suggests:

> To say that one should replace knowledge by hope is to say much the same thing: that one should stop worrying about whether what one believes is well grounded and start worrying about whether one has been imaginative enough to think up interesting alternatives to one's present beliefs. (Rorty 1999: 34)

Nevertheless, I feel tempted by the idea to supplement knowledge with hope, which means to supplement the analysis of the factual with the search for the possible. As a consequence, there is need for an extended conception of theory. It should not only reflect descriptions/explanations of the "real," but also emphasize the need for an outlook beyond the factual. In any case, to develop and communicate our imagination of the future, we need models.

> The cognitive artifacts we create are models: representations to ourselves of what we do, of what we want, and of what we hope for. The model is not, therefore, simply a reflection or a

> copy of some state of affairs, but beyond this, a putative mode of action, a representation of prospective action, or of acquired modes of action (Wartofsky 1979: xv).

The essential role of models as object and objectivation of our imagination is also emphasized by Wood who uses the term "map" instead. "And this, essentially is what maps give us, *reality*, a reality that exceeds our vision, our reach, the span of our days, a reality we achieve no other way" (Wood 1992: 4–5).

Focusing on models of possible worlds as a subject and outcome of scientific inquiry does not mean to give up on theories as the pivotal representation of scientific knowledge. If we regard theories in the sense of the original meaning as an outlook beyond the ostensible, they would also comprise possible worlds. The concept of "possible worlds" is used in logic to overcome the limitation that a sentence must be either true or false (*tertium non datur*). Modal logic allows for assigning a proposition to many possible worlds. While in each of these worlds the *tertium non datur* postulate is satisfied, the overall picture allows for a contingent truth value. The "many-worlds interpretation" of quantum physics that assumes the co-existence of multiple parallel worlds, has been around for long (cf. Carroll 2019; DeWitt 1970; Everett 1957). An extended concept of theory that would also comprise models of the possible would mean to supplement "committed cognitive claim to truth" (Wartofsky 1979: 2) with the pragmatic claim to usefulness.

With respect to guiding change, models of possible worlds need to satisfy two main postulates. First, a possible future world must be feasible. Second, research should focus on those possible worlds that seem to be particularly attractive, which will usually imply that they are better than the actual world. While both postulates, especially the second, create a substantial challenge with respect to justification, the creation of possible worlds is not supposed to prescribe "scientifically grounded" blueprints of a bright future. Instead, they are intended as knowledge offerings that might serve those who will actually create the future as an inspiration and as a guidance. Regarding the digital transformation, conceptual models are of pivotal relevance, since they form the foundation for software systems, which in turn will chiefly contribute to the construction of the future. But how could a method guide the design of such models if we account for the challenge to somehow overcome the limitations of the language we speak? There is no clear recipe to meet this challenge. There are, however, two approaches that seem useful: stepwise destruction and construction of concepts through abstraction. These approaches are core elements of a new paradigm of conceptual modeling called multilevel modeling (cf. Atkinson/Kühne 2001; Frank 2014). Conventional conceptual modeling is done with a given modeling language that defines the scope of possible models. Multilevel modeling allows to modify the language itself during the act of modeling. Furthermore, an arbitrary number of language levels is possible, where each additional level represents an abstraction of existing terms. The abstraction comprises both, classification and generalization. Adding a further level of abstraction also means to increase the range of possible models, that is, of possible worlds that are covered by that language. In case, multilevel models are executable, one aspect of the feasibility postulate is satisfied. Unfortunately, a more detailed description of multilevel modeling is beyond the scope of this essay. The model in Figure 1 may serve to illustrate the principal idea. The concept "BookCopy" could be abstracted onto a concept like "PaperRepresentation," which could be further abstracted onto "Representation," etc. "Representation" would create the question, what kind of representations are conceivable, which may, among others, lead to the concept of digital representation. From a functional perspective, one could ask for the purpose of a model. At first, it might be "publication," which could then be abstracted to something like "communication" and "documentation." Communication could then be further differentiated to new concepts that enable distinguishing different kinds of communication. This kind of critical deconstruction and reconstruction would not only help to conceive of a future that will be constituted by a language yet not known, it would also broaden our perspective on the current world:

> All formation of new concepts, all change in concepts, involves discovery of the world – that is, the development of a new way of looking at the world […] which may be more or less borne out as time goes on. Every theory of formation of new concepts is also about discovering the way the world is. (Schön 1963: 34)

However, not every (re-)construction of possible future worlds will be convincing. Therefore, the justification of models of possible future worlds is of crucial importance. Truth in general, and especially the correspondence theory of truth, are of little use in that respect – even though aspects of feasibility will usually include propositions that can be assigned a truth value. The only way to achieve satisfactory justification of what we regard as useful and what we find worth hoping for is through discourse and agreement. But that, of course, comes with the notorious challenge to evaluate whether those who participate in a discourse are sufficiently qualified. With respect to the pivotal role of machine learning, it will be important to emphasize the need for a justification of results achieved through induction, both from an epistemological and an economic point of view.

4.2 The role of narratives

The design of conceptual models that describe substantial parts of the organization of future worlds (mainly through software) may provide valuable guidelines for change but is not sufficient. First, conceptual models intentionally fade out all aspects that bulk against formalization and automation. Second, the future is not created through an act of engineering that is guided by models developed by scientists. In the end, it needs to be constructed by those who live in the future. Hence, they have to be involved. That will be possible only if the vision of a possible world makes sense to them. In other words: a possible future world can serve as an orientation for change only if people can imagine how it would be to live in such a world. Only then they can participate in a discourse about its evaluation and contribute to its evolution. But how could sense be mediated? Probably, the most effective approach to mediating sense is story telling. "Narratives are about people acting in a setting, and the happenings that befall them must be relevant to their intentional states [...]" (Bruner 1991: 7). In order to provide for sense-making, a story needs to connect to the practices that form people's lives. However, at the same time, it should support their imagination to overcome the limitations of practices they take for granted. This is a serious challenge. It is needless to say that narratives that serve as a supplement to models and theories must not be confused with science fiction. Instead, narratives of possible worlds should aim at a differentiated picture of possible future worlds that include comprehensible descriptions of prospects, conflicts, and threats. In addition, it should of course be made clear that narratives of this kind address a contingent matter: the future they describe should be possible, but that neither means that it will become reality, nor that there is a deterministic way to achieve it.

Since narratives are not an accepted medium to communicate scientific knowledge, there are no corresponding examples of scientific publications – at least none that I would know of. However, there are a few authors that create images of a possible future who use a narrative like style to reach a broader audience. Lanier, for example, outlines the vision of a future, democratic version of the Internet, not only by describing technical details, but also by using narratives (Lanier 2013). Van Reybrouck performs a reconstruction of the concept of democracy to propose possible future models of democratic politics (cf. Van Reybrouck 2016). In addition to a critical analysis of current implementations of democracy, he also illustrates his ideas with narratives.

5 Instead of conclusions

We are living in challenging times, in times of change and contradiction. We cannot predict the future, but we may be able to support a transformation to the better. To that end, it is not advisable to follow the traditional path of scientific inquiry that is restricted to the past and the present. It seems more appropriate to extend the scope of our interest to the investigation of possible future worlds. As I outlined only briefly, such a research program is hardly compatible with established research methods. Furthermore, it is confronted with tremendous methodological problems. The approaches I proposed as elements of a method or even a methodology of change are no solutions but create further challenges. Against this background, one may feel tempted to avoid the risk, that is, to continue focusing on the "factual" and fade out the possible. That, however, would mean to miss a fantastic opportunity. Is it not most inspiring to not only ask why things are, but how they could be? That includes the troubling question, how our cognitive capabilities, how our imagination, how *we* would change if we managed to acquire new languages that include concepts based on abstractions we are not yet aware of. And, of course, we would compromise our credibility if we did not also ask how academia could be changed to serve its purpose better. I agree with Lewis who argues that the idea of possible worlds creates a "philosopher's paradise" (here I would like to add: not only a paradise for philosophers): "We have only to believe in the vast realm of *possibilia*, and there we find what we need to advance our endeavours" (Lewis 1986: 4).

References

Anderson, Chris (2008): "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." In: *Wired* 23 June 2008. https://www.wired.com/2008/06/pb-theory/, accessed 21 March 2020.

Atkinson, Colin/Kühne, Thomas (2001): "The Essence of Multilevel Metamodeling." In: Gorgolla, Martin/Kobryn, Cris (Eds.): *"UML 2001" – The Unified Modeling Language: Modeling Languages, Concepts, and Tools*. New York, NY: Springer, pp. 19–23.

Boddington, Paula (2017): *Towards a Code of Ethics for Artificial Intelligence*. New York, NY: Springer.

Boland, Richard/Collopy, Fred (2004): "Design Matters for Management." In: Boland, Richard/Collopy, Fred (Eds.): *Managing as Designing*. Stanford, CA: Stanford Business Books, pp. 3–18.

Bruner, Jerome (1991): "The Narrative Construction of Reality." In: *Critical Inquiry* 18(1), pp. 1–21.

Bundesregierung (2018): *Strategie Künstliche Intelligenz der Bundesregierung*. https://www.bmbf.de/files/Nationale_KI-Strategie.pdf, accessed 21 March 2020.
Carroll, Sean (2019): *Something Deeply Hidden*. Munich: Random House.
Derrida, Jacques (1976): *Of Grammatology*. Baltimore, MD: Johns Hopkins University Press.
DeWitt, Bryce S. (1970): "Quantum Mechanics and Reality: Could the Solution to the Dilemma of Indeterminism Be a Universe in Which All Possible Outcomes of an Experiment Actually Occur?" In: *Physics Today* 23(9), pp. 30 – 35.
Domingos, Pedro (2017): *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. London: Penguin.
Everett, Hugh (1957): "'Relative State' Formulation of Quantum Mechanics." In: *Reviews of Modern Physics* 29(3), pp. 454 – 462.
Frank, Ulrich (1988): *Expertensysteme: Neue Automatisierungspotentiale im Büro- und Verwaltungsbereich?* Wiesbaden: Gabler.
Frank, Ulrich (2011): *Multi-Perspective Enterprise Modelling: Background and Terminological Foundation*. ICB Research Report No. 46. https://www.icb.wiwi.uni-due.de/fileadmin/fileupload/ICB/research/research_reports/ICB-Report-No46.pdf, accessed 21 March 2020.
Frank, Ulrich (2014): "Multilevel Modeling: Toward a New Paradigm of Conceptual Modeling and Information Systems Design." In: *Business & Information Systems Engineering* 6(6), pp. 319 – 337.
Frank, Ulrich (2017): "Theories in the Light of Contingency and Change: Possible Future Worlds and Well-Grounded Hope as a Supplement to Truth." In: *Proceedings of the 50th Annual Hawaii International Conference on System Sciences*, pp. 5727 – 5736.
Habermas, Jürgen (1976): "Technik und Wissenschaft als 'Ideologie'." In: *Technik und Wissenschaft als "Ideologie."* Frankfurt a. M.: Suhrkamp, pp. 48 – 103.
Hempel, Carl G. (1966): *Philosophy of Natural Science*. Upper Saddle River, NJ: Prentice Hall.
Lanier, Jaron (2013): *Who Owns the Future?* New York, NY: Simon & Schuster.
Lewis, David (1986): *On the Plurality of Worlds*. Oxford: Blackwell.
Murphy, Kevin P. (2012): *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press.
Pentland, Alex (2014): *Social Physics: How Good Ideas Spread – The Lessons from a New Science*. Melbourne: Scribe.
Precht, Richard David (2018): *Jäger, Hirten, Kritiker: Eine Utopie für die digitale Gesellschaft*. Munich: Goldmann.
Rescher, Nicholas (1980): *Induction: An Essay on the Justification of Inductive Reasoning*. Pittsburgh, PA: University of Pittsburgh Press.
Rorty, Richard (1999): *Philosophy and Social Hope*. New York, NY: Penguin.
Rorty, Richard (2000): "Universality and Truth." In: Brandom, Robert B. (Ed.): *Rorty and His Critics*. Malden, MA: Blackwell, pp. 1 – 30.
Schön, Donald A. (1963): *Displacement of Concepts*. London: Tavistock.
Sinclair, Andrew/Brashear, Jeffrey/Shacklady, John (2018): *AI: The Momentum Mindset*. https://www.accenture.com/_acnmedia/pdf-73/accenture-strategy-ai-momentum-mindset-exec-summary-pov.pdf, accessed 21 March 2020.
Van Reybrouck, David (2016): *Against Elections: The Case for Democracy*. London: The Bodley Head.
Wartofsky, Marx W. (1979): *Models: Representation and the Scientific Understanding*. Dordrecht: Reidel.

Wichert, Andreas (2013): "Proto Logic and Neural Subsymbolic Reasoning." In: *Journal of Logic and Computation* 23(3), pp. 627 – 643.
Widdicombe, Lizzie (2014): "The Programmer's Price." In: *The New Yorker* 24 November 2014, pp. 54 – 64.
Winograd, Terry/Flores, Fernando (1986): *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, NJ: Ablex.
Wittgenstein, Ludwig (1961) [TLP]: *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul.
Wittgenstein, Ludwig (1973) [PI]: *Philosophical Investigations*. Oxford: Prentice Hall.
Wood, Denis (1992): *The Power of Maps*. New York, NY: Guilford Press.