



Perspective

From the Digital Data Revolution toward a Digital Society: Pervasiveness of Artificial Intelligence

Frank Emmert-Streib ^{1,2}

¹ Predictive Society and Data Analytics Lab, Faculty of Information Technology and Communication Sciences, Tampere University, 33720 Tampere, Finland; v@bio-complexity.com; Tel.: +358-50-301-5353

² Institute of Biosciences and Medical Technology, 33520 Tampere, Finland

Abstract: Technological progress has led to powerful computers and communication technologies that penetrate nowadays all areas of science, industry and our private lives. As a consequence, all these areas are generating digital traces of data amounting to big data resources. This opens unprecedented opportunities but also challenges toward the analysis, management, interpretation and responsible usage of such data. In this paper, we discuss these developments and the fields that have been particularly effected by the digital revolution. Our discussion is AI-centered showing domain-specific prospects but also intricacies for the method development in artificial intelligence. For instance, we discuss recent breakthroughs in deep learning algorithms and artificial intelligence as well as advances in text mining and natural language processing, e.g., word-embedding methods that enable the processing of large amounts of text data from diverse sources such as governmental reports, blog entries in social media or clinical health records of patients. Furthermore, we discuss the necessity of further improving general artificial intelligence approaches and for utilizing advanced learning paradigms. This leads to arguments for the establishment of statistical artificial intelligence. Finally, we provide an outlook on important aspects of future challenges that are of crucial importance for the development of all fields, including ethical AI and the influence of bias on AI systems. As potential end-point of this development, we define digital society as the asymptotic limiting state of digital economy that emerges from fully connected information and communication technologies enabling the pervasiveness of AI. Overall, our discussion provides a perspective on the elaborate relatedness of digital data and AI systems.

Keywords: artificial intelligence; machine learning; data science; social data; natural language processing; industry 4.0; ethical AI



Citation: Emmert-Streib, F. From the Digital Data Revolution toward a Digital Society: Pervasiveness of Artificial Intelligence. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 284–298. <https://doi.org/10.3390/make3010014>

Academic Editor: Isaac Triguero

Received: 22 January 2021

Accepted: 3 March 2021

Published: 4 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the last few decades, technological progress has changed nearly all areas of science [1–3]. This comprises many fields, including biology, computer science, economy, engineering, humanities, journalism, politics, public health, management, medicine, social sciences, sports and even arts. While the generation of data have long been a privilege of basic research, the computerization of society and the establishment of the Internet have enabled the availability and the distribution of information and data on almost all aspects of our daily lives. As a consequence, a quantitative analysis of such digital data can be conducted by means of artificial intelligence (AI) and machine learning with results that might have a profound effect on all levels of society [4–6].

A field that was among the first transitioning into a technology-driven area was biology [7,8]. Interestingly, the Human Genome Project [9] helped in enhancing molecular high-throughput measurements, e.g. next-generation sequencing (NGS) technologies [10], which allows the interrogation of all molecular levels, including mRNAs, proteins and DNA sequences [11,12]. In recent years, this technology has also infiltrated the biomedical and clinical sciences which allowed a quantification of those fields as well. Further

areas that have been significantly transformed by the digital revolution are economy and business. Importantly, most of the trading on the stock markets worldwide is nowadays conducted electronically, i.e., orders can be placed online and are directly sent to a broker circumventing traditional floor trading. Lastly, also the social sciences have been heavily influenced by big data [5,13]. For instance, various kinds of social media platforms, e.g., facebook, twitter or linkedin, provide a sort of virtual laboratories for conducting social and psychological experiments leading to novel insights of human behavior [14–17].

Considering the fact that technologies like the Internet, NGS or the iPhone have only been available since 1991, 2004 and 2007, respectively, it seems clear that within the next thirty years the pace of new inventions will likely further increase. Hence, novel technologies building upon existing ones, e.g., 5G, will further transform science, industry and our private lives in profound and potentially hard to foreseeable ways. In this paper, we accept this challenge and take a look ahead by discussing some of the potential changes for these fields, beneficially or detrimentally. Specifically, in this paper, we discuss fields that have been particularly effected by the digital revolution, e.g., medicine and economy. Our discussion is AI-centered showing domain-specific opportunities but also challenges for the method development in artificial intelligence. For instance, we discuss recent breakthroughs in deep learning algorithms and artificial intelligence as well as advances in text mining and natural language processing, e.g., word-embedding methods that enable the processing of large amounts of text data from diverse sources such as governmental reports, blog entries in social media or clinical health records of patients. Thereafter we discuss the necessity of further improving general artificial intelligence approaches and for utilizing advanced learning paradigms. This leads to arguments for the establishment of statistical artificial intelligence. Finally, we provide an outlook on important aspects of future challenges that are of crucial importance for the development of all fields. Importantly, we will define digital society as an asymptotic limiting state of digital economy characterized by the pervasiveness of artificial intelligence. Overall, in this paper, we provide a perspective on the elaborate relatedness of digital data and AI systems and on potential future developments.

This paper is organized as follows. In the next sections, we discuss several fields that have been significantly reshaped by the digital revolution. We discuss opportunities for the method development in artificial intelligence and potential domain-specific challenges. Then we discuss general instances of artificial intelligence approaches and learning paradigms that might be especially beneficial to all fields effected by the digitalization. Thereafter, we discuss important aspects of future challenges that are of crucial importance for the development of the respective fields. For this discussion, and the previous presentation, we assume an AI-centered perspective. Aside from mostly positive effects of a digitalization, serious problems thereof are addressed, e.g., about data privacy and fundamental issues of artificial intelligence governance. As a key problem, the asymptotic state of digital economy, we call digital society, is discussed. The paper finishes with concluding remarks.

2. Digital Medicine and Digital Health

As already mentioned, biology experienced a transition toward a technology-driven area in the 1990s. This was accomplished by introducing the DNA microarray technology allowing the measurement of genome-scale information of the concentration of messenger ribonucleic acids (mRNAs). Further technologies that followed were SELDI (Surface-enhanced laser desorption/ionization), protein-chips and various forms of NGS assays (next-generation sequencing), e.g., DNA or RNA sequencing or DNA methylation [18–20]. Importantly, many of these technologies also propagated to medical, clinical and public health studies which made also these fields essentially data-driven as a consequence of such technologies.

There are many subfields of the above subjects that utilize modern information and communication technologies in biology, medicine and public health. However, the terms

digital medicine or digital health are commonly used to indicate the general integration of such digital technologies, e.g., with smartphone or sensor technologies, with advanced analysis methods to enhance the subject related goals [21–27]. Interestingly, in [28] it is noted that “Despite a flurry of recent discussion about the role and meaning of AI in medicine, in 2017 nearly 100% of U.S. healthcare will be delivered with 0% AI involvement”. This statement underlines the difficult road ahead for translating results from basic research to the application in hospitals or healthcare systems but shows also the potential for methods from AI.

Challenges and obstacles: A necessity for AI to make beneficial contributions to medicine and health, but also to other fields discussed below, is the availability of (large amounts of) data. However, present genomics technologies, clinical and pathological imaging technologies, biosensors, and the internet of things (IoT) devices are essentially capable of fueling AI methods with sufficient data. There is just the requirement to gather patient-specific data over a longer period of time to establish data repositories similar to the ImageNet database for images. Then personalized or patient-tailored methods can be developed and benchmarked to enhance the current state-of-the-art in computational diagnostics and evidence based medicine. An example for this could be the concept of a digital twin [29,30].

Promising pilot studies exist that demonstrate the utility of AI methods, especially deep learning, for digital health. For instance, such studies were conducted for diabetic retinopathy [31], skin cancer [32] and medication adherence [33], to name just a few. Interestingly, most of such studies are mainly based on image analysis. This is another indicator of the early stage of digital medicine because medicine and health offer many more data types, as mentioned above, beyond imaging data.

There are three major concerns frequently raised against AI in medicine and health. The first is the fear that jobs will be lost due to the introduction of automatic analytics systems, the second criticizes the potential disruption of the personal doctor-patient relationship for similar reasons and the third issue relates to the lack of explainability of general AI methods [28,34]. The latter point means that usually AI models can be considered as black-box prediction models that are capable of achieving high prediction performance but lack intuitive explanations that describe, e.g., in standard medical terms, how the performance was actually obtained.

It is important to highlight that all three concerns do not relate to methodological issues of AI itself but to job safety, trust and communication. This means in order to pave the way for AI in medicine and health there is also educational work of the public and control bodies necessary to overcome negative and possibly ill-informed sentiments. Furthermore, it is important to mention that digital health and digital medicine require a multidisciplinary approach for their successful deployment [23]. Given the experience of similar but potentially smaller-scale endeavors from bioinformatics or systems biology there is already demonstrated success one can build on in forming the cross- and interdisciplinary teams needs.

Finally, for all the above approaches there are privacy and ethical issues that need to be dealt with properly [35,36]. On one hand, this needs to ensure that a patient is in control of its own data but also that sufficient data are collected and available for the development of data-driven AI methods. This is certainly a noal balancing act to fulfill all needs. A potential circumvention of this could be the anonymization of patient data in a way that the data are modified in a manner that individuals are no longer identifiable yet the modified data are not effecting prediction results from AI methods. For some promising pilot studies, see [37,38].

3. Digital Economy and Business

The digitalization of our world is not limited to medicine and health but effects also the way we conduct business and organize our entire economy. There have been many attempts to define ‘digital economy’ and a nice review of a large number of such definitions

can be found in [39]. Overall, the common agreement about the nature of digital economy is succinctly summarized by “...an economy based on digital technologies (sometimes called the internet economy)” provided by the Expert Group on Taxation of the Digital Economy of the European Commission. We would like to note that sometimes digital economy is all called web economy or new economy. A similar definition for digital business has been provided by [40] as “organizational strategy formulated and executed by leveraging digital resources to create differential value”. Both definitions are rather general but the diversity of these fields requires such a wide characterization to encompass all relevant aspects thereof.

Specific main sectors that are included in the above definitions are:

- e-Business;
- e-Commerce;
- Industry 4.0;
- Sharing economy;
- Crowdsourcing.

Here electronic business (e-business) “is any process that a business organization conducts over computer-mediated networks” and electronic commerce (e-commerce) “is the value of goods and services sold over computer-mediated networks” [41]. Industry 4.0 (or smart factory) stands for the fourth industrial revolution which is transforming traditional manufacturing and industrial processes into a technology mediated field [42,43]. This includes machine-to-machine communication (M2M), Internet of Things (IoT) and cyber-physical systems (CPS). According to Kagermann et al. [44], Industry 4.0 is “a new level of value chain organization and management across the lifecycle of products”. That means not only the production and manufacturing is effected but also decision making across all relevant levels including the management. Furthermore, Industry 4.0 includes not only customization but also a personalization of products [45] (see Table 1 for an overview).

Table 1. Key technologies for Industry 4.0. For a discussion of succinct differences and commonalities between these technologies see [46,47].

Technology	Definition	Reference
Machine-to-machine communication (M2M)	“Machine-to-Machine (M2M) paradigm enables machines (sensors, actuators, robots, and smart meter readers) to communicate with each other with little or no human intervention. M2M is a key enabling technology for the cyber-physical systems (CPSs)”.	[48]
Wireless sensor networks (WSN)	“WSN is designed particularly for delivering sensor-related data”.	[49]
Internet of Things (IoT)	“An open and comprehensive network of intelligent objects that have the capacity to auto-organize, share information, data and resources, reacting and acting in face of situations and changes in the environment”.	[50]
Cyber-physical systems (CPS)	“CPS are systems of collaborating computational entities which are in intensive connection with the surrounding physical world and its on-going processes, providing and using, at the same time, data-accessing and data-processing services available on the internet”.	[51]

Another important part of digital economy is sharing economy (SE). Sometimes SE is also called access economy, peer-to-peer (P2P) economy or collaborative economy [52,53]. Furthermore, SE is a wide term that has been defined as: “the sharing economy is an IT-facilitated peer-to-peer model for commercial or non-commercial sharing of underutilized goods and service capacity through an intermediary without a transfer of ownership” [54]. Hence, its underlying idea is to directly connect individual consumers and individual

providers of goods or services facilitated by a community-based on-line platform. Examples of such business models are:

- Freelancing platforms (labor market consisting of short-term contracts);
- Coworking platforms (individuals working independently or collaboratively in shared office space);
- P2P lending platforms;
- Fashion platforms.

Finally, crowdsourcing (CS) shares some similarity to sharing economy and it is defined as: “Crowdsourcing is the IT-mediated engagement of crowds for the purposes of problem-solving, task completion, idea generation and production” [55]. Important examples of CS are information sharing systems, e.g., Wikipedia or del.icio.us, voting systems, e.g., Amazon’s Mechanical Turk (MTurk) and gamification, e.g., reCAPTCHA (image recognition) or Foldit (protein folding).

Overall, we would like to note that these developments can effect any industry sector. In Table 2, we show some examples for such sectors according to the taxonomy of the Global Industry Classification Standard (GICS) [56]. Currently, the GICS categorizes industry into 11 Sectors, 24 Industry groups, 69 Industries and 158 sub-industries resulting in conceptual branches. There are alternative industry taxonomies, e.g., Standard Industry Classification (SIC), Bloomberg Industry Classification Standard (BICS), Statistical Classification of Economic Activities in the European Community (NACE) or North American Industry Classification System (NAICS) but essentially any of these is sufficient to get a high-level summary of our economy.

Table 2. Major sectors according to the taxonomy of the Global Industry Classification Standard (GICS) that categorizes the industry into 11 Sectors, 24 Industry groups, 69 Industries and 158 sub-industries; see [56].

Sector	Industry	Sub-Industry
Energy	Oil, Gas and Consumable Fuels	Coal and Consumable Fuels
Materials	Chemicals	Fertilizers and Agricultural Chemicals
Industrials	Machinery and Agricultural	Farm Machinery
Consumer Discretionary	Hotels, Restaurants and Leisure	Restaurants
Consumer Staples	Food, Beverage and Tobacco	Tobacco
Health Care	Pharmaceuticals, Biotechnology and Life Sciences	Biotechnology
Financials	Banks	Regional Banks
Information Technology	Software and Services	Internet Services and Infrastructure
Communication Services	Media and Entertainment	Publishing
Utilities	Utilities	Independent Power and Renewable Electricity Producers
Real Estate	Real Estate	Real Estate Development

In order to show the economic importance of the above fields on economy itself several studies have been conducted. For instance, sharing economy is in 2020 valued at US\$15 billion globally with a potential to raise its global market value to US\$335 billion by 2025 [57]. Measuring the value of digital economy by a study of the United Nations estimated that digital economy contributes 4.5% to 15.5% of the world GDP [58]. These numbers show the enormous impact of digital economy and its comprising subfields on our world economy and its potential to further increase.

Challenges and obstacles: For AI systems there is a large number of directions open for contribution. In general, artificial intelligence should be central for any data-driven approach in digital economy including Industry 4.0. For instance, AI can make valuable contributions to predictive maintenance (PdM) [59,60]. PdM is dealing with maintenance issues of production devices or general machines and helps in reducing down times or operational costs. By utilizing sensor information of either production or operation lines,

AI-based prediction models can be trained and used for optimizing maintenance schedules. Furthermore, AI should be helpful for any kind of IoT or CPS application because such technologies were designed for the gathering of data but not for their analysis. Finally, AI can help to further improve robotics and automation for manufacturing, production or service applications. Currently, deep reinforcement learning is showing promising results for such a novel AI approach in this context [61,62].

On a more fundamental point, it is important to note that the application of AI requires also an adjustment regarding general data analysis principles. An early standard of this has been called CRISP-DM (cross-industry standard process for data mining) [63] emphasizing the feedback between consecutive analysis steps. Recently, this has been extended considering also industry-specific needs and domain-specific knowledge [64].

There are three major concerns frequently raised against AI methods in business and economy. The first is the fear that some jobs will be lost due to the introduction of automatic analytics systems [65], the second issue relates to the lack of explainability of general AI methods and the third is concerned with the increasing gap between developed and developing countries and the general change of wealth distribution [66]. Interestingly, the first two points are essentially the same as for digital medicine and health systems; discussed above. The latter issue is addressed by artificial intelligence governance which needs to be developed accordingly.

4. Pervasiveness of Artificial Intelligence

All of the above problems can only be studied by using methods to learn from data [67]. It is important to realize that the methods for studying such problems are not always the same but they need to be adopted or even newly developed to fit a given data set optimally. Hence, there is a constant need to further enhance and extend the existing pool of machine learning and artificial intelligence methods because the technology underlying the problems which enables the generation of data is constantly changing.

For reasons of clarity, we would like to emphasize that there are many scientific areas dealing with the development of novel methods for the analysis of data. For instance, machine learning, statistics, pattern recognition or artificial intelligence are all different fields with their own history and preferences for methodological approaches and conceptual frameworks [68–70]. However, in this paper, we simplify the discussion by summarizing these fields by the term artificial intelligence because, especially, in industry this term has become the commonly accepted standard when speaking about data analysis methods and approaches. Nevertheless, one should be aware that on the academic side this is seen differently.

Due to the diversity of fields generating general data, one can expect that the methods for their analysis are similarly diverse. Currently, methods for image and audio processing [71,72] seem to be much more developed than methods for other types of data, e.g., text data, genomics data or sensor data. Hence, such data types offer a great potential for improvements. For instance, for text data a fundamental problem is a conversion of textual information into numbers in a way that conventional AI methods can process such data. Recently, word-embedding methods made great progress, above all word2vec or BERT [73–75]. However, there is still room for improvement especially for mapping onto larger units, e.g., paragraphs or documents [76].

Another area of great potential is information fusion [77]. The general idea is to combine data from multi-sensors, multi-sources or multi-processes in a way that the resulting data set contains more information than its separate, individual sources. This problem becomes apparently more difficult the more different the individual data sources or sensors are, especially, if these correspond to different data types, e.g., image data and text data.

Furthermore, transfer learning [78] should be mentioned as a field of great potential. Transfer learning means that one starts training a model for one task (called a source task) and then switches the data for another one (called target task). For instance, in

image processing a model that has been trained for classifying non-medical images could be transferred to learning to discriminate tumors in medical images [79]. This example demonstrates also that transfer learning is particularly useful when one has only a very limited amount of data for a certain task (for instance from medical images) but a much larger data set for a similar task (for non-medical images). Other learning paradigms that are severely underutilized are semi-supervised learning or multi-task learning [80,81].

Finally, a field that should be established is statistical artificial intelligence (SAI). SAI would extend the ideas of statistics to artificial intelligence, e.g., by investigating the influence of the sample size on the resulting prediction performance. This is important because a method in isolation is neither good nor bad but only in combination with data of certain characteristics a method can be evaluated. For instance, in [82] a deep learning classifier (a Long Short-Term Memory (LSTM) model) has been studied for classifying handwritten digital characters provided by the EMNIST (Extended MNIST) data [83]. As a result, it has been found that by using over 200,000 training samples the classification error is far below 5%, while for 5000 training samples the error increases to over 30%. Considering that the underlying deep learning classifier was the same for both approaches, this demonstrates the importance of quantifying the influence of the sample size on the resulting performance. Formally, such a characterization of a model is obtained by so called learning curves [84]. In general, AI methods in digital business and digital health appear to be studied in a less stringent way as compared to, e.g., methods from biostatistics. This is understandable given the fact that the latter methods find regular application in medical and clinical patient data. Nevertheless, also in those fields a steady control is required for ensuring quality standards [85] because there are examples that violating such standards can jeopardize the lives of patients [86].

5. Discussion

The above presentation discussed the individual fields separately and focused mainly on their core components. Despite this, one could already recognize that there are many commonalities among the different fields and approaches. For this reason, in the following, we focus on common aspects shared by these fields presented in an AI-centric way.

5.1. Smart Cities and Smart Government

One may wonder if there are other fields or areas beyond medicine, health, business and economy that could benefit from a digitalization and utilization of AI in a similar way as, e.g., digital economy? In fact, there are already some developments in this direction. For instance, smart city and smart government are attempts to improve the organization of cities or governments, respectively. In order to accomplish this, smart cities utilize many sensors throughout the city, e.g., via IoT technologies, to improve traffic management, road safety or energy efficiency [87–89]. Similarly, smart government utilizes mainly automation for administrative tasks (e-government) and for data-driven decision making [90]. However, the developmental state of these areas is considerably behind other fields, e.g., digital health. One reason for this may be the fact that, traditionally, neither cities nor governments are based on electronic communication technologies. Hence, there is, first, a need to introduce information processing and computing technologies for generating and gathering data and then AI-based systems can be designed for the solution of particular tasks.

For the near future, for smart cities it is expected that more sensors are needed throughout the cities for gathering information about traffic, environmental conditions and human behavior. In contrast, for smart government, text mining based AI approaches seem very promising because, essentially, all administrative tasks involve text data. Similar to digital economy, also smart cities and smart government will gradually develop toward higher states of digital cities and digital government. However, for the latter it remains to be seen how much involvement of AI is desirable or acceptable because, ultimately, even political decisions could be made based on such automated methods.

It is certainly intriguing to think about such possibilities even if only applied retrospectively. An example demonstrating the problems with such an approach for automatic decision making is the Brexit. Assuming we would have an AI system that would allow us to answer the following political question:

- Question: Should the UK leave the European Union?

Phrased like this it is a binary classification task which could be solved by a supervised learning method because the answer is either to leave or to stay in the European Union. A problem is that contrary to general artificial intelligence or machine learning approaches, there are no samples of similar 'events' available one could use for the training of this supervised learning task. Leaving technical difficulties aside, assuming we would have access to an AI system that could provide a faithful answer to this question (for instance based on PU-learning), what would this entail? Would this answer be convincing to the people who's opinion was the opposite? Given the fact that a general classification result does not come with an explanation, such an answer could be misunderstood or even bewildering to the public in large.

This example demonstrates potential limitations even of error-free AI systems capable of making the correct political decisions. Hence, in a such a context, AI by itself could not provide the final solution but needs to be complemented with additional features not unlike to what is currently discussed for explainable AI (XAI) [34,91–93]. Overall, this example shows that AI-driven decision making on higher levels, e.g., on a management or governmental stage, possess new challenges that need to be addressed.

5.2. Human–Machine Interaction

A related topic, but coming from a different perspective, is human–machine interaction (HMI) [94]. In HMI, also called human–computer interaction (HCI) [95], one assumes that a machine or computer is not capable of performing the complete task by its own. Instead, some form of human-involvement is needed for a succesful implementation and execution of the task. That means there is an interface between people and machines or computers.

Prominent application examples are doctor-in-the-loop for supporting medical decision making by health practitioners [96] or augmented reality [97] for merging physical and virtual perception of a user. The former finds application in digital health whereas the latter is used in digital business, e.g., for virtual viewings of properties. From an abstract point of view, also data science falls within the category of human–computer interaction because a complex data analysis process involves many individual steps which may not be automatically connectable but requires human intervention, e.g., via an explanatory analysis [98].

A general question, also related to the example from the previous section about political decision making, is if one needs always a form of a human-computer interface for solving AI tasks in higher organizational or societal layers, as represented by health-related problems, the economy or cities, or if this is only needed in special cases? Classically, the ideal case seems to be an human-free AI system because this eliminates a potential bias of the subjectivity of humans. However, as the example about political decision making or the doctor-in-the-loop shows this is not clear for such problems.

5.3. Data Privacy, Cybersecurity and Bias in AI

Another aspect that is also shared by the fields discussed in this paper is data privacy (information privacy) and cybersecurity. For instance, due to the increasing usage of connected technologies, e.g., IoT or CPS, Industry 4.0 systems are vulnerable to cyber attacks [99]. Similarly, patient data from hospitals or retirement homes contain sensitive personal information that needs to be protected from third party usage. These problems are also well known for social media data, e.g., for using facebook or twitter [100,101].

A practical example where the usage of personal data by a third party may lead to unwanted consequences is the insurance or financial industry. For instance, private data could be used for the evaluation of an insurance premium or its coverage. Similarly,

banks may base their decision to provide credits on similar data about an individual. This problem is certain to become more severe the more the single layers of our society are interconnected with each other, see Figure 1, in a way that the corresponding data from those layers become simultaneously available.

Hierarchical organization of society

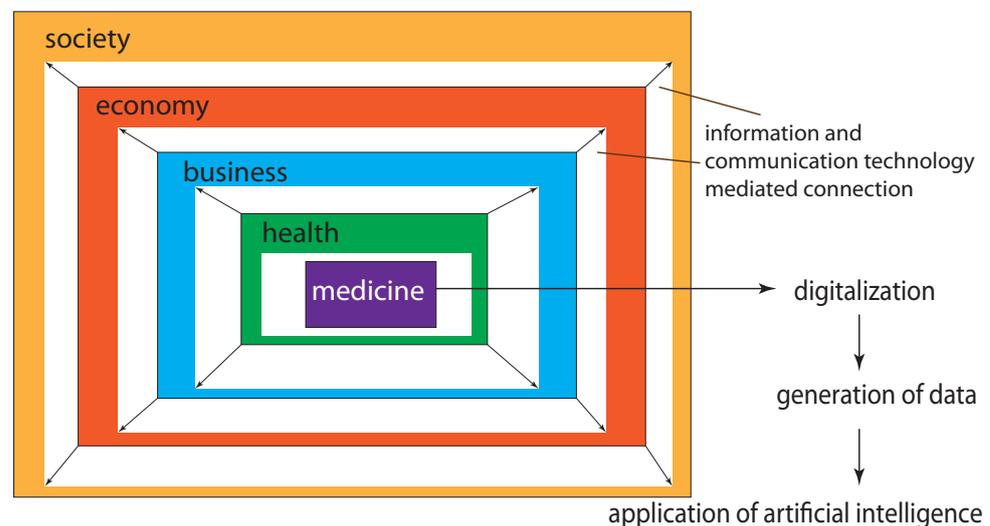


Figure 1. A simplified view of the hierarchical organization of society. The digitalization of the shown fields progresses from the inside toward the outside. This leads eventually to a full penetration of society with artificial intelligence.

A related issue to data privacy is artificial intelligence governance [102]. AI governance is concerned with ethical standards, safety, transparency and public fear. This addresses the needs of individual users but also regulation bodies. Again there are cross links to our political decision making example above, especially relating the last point. For companies, on the other hand, there is the concern that AI governance might be too strict and rigid preventing the implementation of viable business ideas. Overall, this will be a balancing act to find the right regulations satisfying the needs of all parties.

Finally, we would like to mention another issue of great importance which is the bias in AI systems [103]. Sources for bias in AI can be either due to biased data aggregation or algorithmic unfairness. For automatic decision making systems, this could for instance lead to gender-bias or racial-bias both of which demonstrate the severity of this problem. For this reason it is important to avoid any form of bias in an AI system. However, this requires the understand of bias, mitigating bias and accounting for bias [104], which is usually noal. In our opinion, so far this topic has not found widespread appreciation and appears underrepresented. Only very recently this seems to change [105–107]. A possible reason for this delayed interest in this topic could be the fact that AI is still mostly concerned with low-level decision making, e.g., via classification or regression, whereas high-level decision making, e.g., for management, administration or legal problems, is still in its infancy. However, the more AI tends toward the latter areas the more the problems with bias in AI systems will become apparent.

5.4. From Big Data and Cloud Computing toward Advanced Analytics

It is important to realize that despite the fact that data are the driving force (or fuel) of general AI methods, they are not sufficient but necessary. This insight triggered the big data era where essentially all fields started to store all sorts of sector-specific data and cloud storage became popular. The next step that built upon big data was cloud computing because one does not only need data but also the capabilities to process these efficiently. Unfortunately, neither of these two steps generates value by itself but they provide only the potential for deriving value thereof. In large, this can only be accomplished by the

application of AI and machine learning methods. Hence, we need to move away from data storage and data processing as a main focus toward advanced analytics [108,109]. This includes visualization and dimension reduction methods for an explanatory data analysis [110,111] as well as unsupervised (clustering or hypothesis testing [112,113]) and supervised learning (classification or regression [69,114,115]) methods. Furthermore, there are some modern learning paradigms that deserve more attention especially in combination with deep learning architectures [116–119]:

- Semi-supervised learning;
- Reinforcement learning;
- Transfer learning;
- Adversarial learning.

In addition, technology-mediated forms of AI are needed for optimizing the underlying technologies for digital health and digital economy, e.g., IoT, CPS or M2M.

On a statistical note, we would like to remark that, in general, the characteristics of a data set has a strong influence on the performance of any model. For this reason, also methods for investigating the statistical robustness of methods are of importance. This includes also resampling methods like cross validation (CV). In our discussion above, we called the field that provides a systematic analysis of such issues statistical artificial intelligence.

5.5. From Digital Economy to Digital Society

We want to finish our discussion with an outlook on the potential end-point of all these developments. What does this mean? Given the above explored information and communication technologies that find already application ranging from digital health to digital economy, a natural question arising is, if it is possible to foresee where all this might lead us?

In our discussion about digital economy, we saw that digital economy is characterized by different sectors, of which Industry 4.0 is one of them. We think that this is a good way to indicate the continuous transformation or evolution of fields and would also suggest to utilize such an enumeration for digital economy itself, e.g., in the form of digital economy 4.0 [120]. Thinking ahead, one may wonder what is the potential end-point of such a developmental process. Let us call this end-point digital economy X whereas ' X ' corresponds to an unknown number. We do not know what the value of ' X ' might be but it seems natural to assume that the limiting state or end-point of digital economy will provide a connected network of the individual layers manifesting society mediated by information and communication technologies (see Figure 1). Hence, the resulting state has no borders between economy, business and society. That means every part of our home, work, education and recreation will be fully penetrated by AI and the society will be also part of a business sector (see Figure 1 for a simplified visualization of a hierarchical society). We call this end-point resulting from this development digital society.

In a society that is fully connected by information and communication technologies enabling the pervasiveness of AI the asymptotic limiting state of digital economy X is called digital society.

From this scenario it seems clear that such a development would be undesirable because it would eradicate any kind of privacy. Hence, from an ethical point of view the question emerges what is the largest number ' X ' we would be willing to tolerate?

It is interesting to note that there are similarities of our argument to Granovetter's theory of the embeddedness of economic actions in society via social networks, which he called new economic sociology [121]. This, in retrospective very plausible argument, however, penetrates also businesses and firms alike and in a fully connected world by information and communication technologies leads to a digital society.

In general, detailed predictions about the future developments involving AI in our society are difficult [122]. Or to say it with the words of Stephan Hawking: "The rise of

powerful AI will be either the best or the worst thing ever to happen to humanity. We do not yet know which". From reading this, it may not be surprising to find advocates along this broad spectrum of possible scenarios [123,124]. For instance, Kurzweil represents an optimistic view that AI, along with nanotechnology and genetics, will improve our lives for the better, whereas Joy assumes a pessimistic role who even sees humanity threatened by such technologies to the extent of being extinct [125]. We do not want to miss mentioning that there are also less extremal views on the future of AI. Such advocates may be seen as pragmatists because they believe in a beneficial application of AI, by at the same time maintaining control over all crucial aspects of safety and security [123]. The only thing that seems to be clear at this moment is that most results achieved so far are largely overhyped and we are still (far) away from full pervasiveness of artificial intelligence [126].

The best, and possibly most pragmatic, way to go forward seems to conduct more interdisciplinary research and education to enlighten our way ahead and to hopefully avoid devastating end-points [127,128].

6. Conclusions

We hope that our perspective on the development of digital medicine and digital economy toward a digital society, leading to a pervasiveness of artificial intelligence in all layers of society, demonstrates the need for a concerted effort in this area. A final point we would like to mention that seems to be underdiscussed in the current literature is 'responsibility'. Specifically, should AI scientists be responsible for their inventions and consequences these possibly have on society [129]? This can affect 'optimists' and 'pessimists' alike because from both negative consequences can arise in either over- or underutilizing opportunities.

In summary, artificial intelligence in combination with digitalization offers a multitude of avenues to go forward that could change our lives in many profound ways. Interestingly, all of these ways seem to be inclusive with respect to different scientific fields and application domains because AI raises questions in technology, mathematics, business and ethics alike. Hence, whatever the future will look like it will be multidisciplinary.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Chang, R.M.; Kauffman, R.J.; Kwon, Y. Understanding the paradigm shift to computational social science in the presence of big data. *Decis. Support Syst.* **2014**, *63*, 67–80. [[CrossRef](#)]
2. Chen, H.; Chiang, R.H.; Storey, V.C. Business intelligence and analytics: From big data to big impact. *MIS Q.* **2012**, *36*, 1165–1188. [[CrossRef](#)]
3. Helbing, D. *Thinking Ahead—Essays on Big Data, Digital Revolution, and Participatory Market Society*; Springer: Berlin, Germany, 2015; Volume 10.
4. Ma'ayan, A.; Rouillard, A.; Clark, N.; Wang, Z.; Duan, Q.; Kou, Y. Lean Big Data integration in systems biology and systems pharmacology. *Trends Pharmacol. Sci.* **2014**, *35*, 450–60. [[CrossRef](#)]
5. Olshannikova, E.; Olsson, T.; Huhtamäki, J.; Kärkkäinen, H. Conceptualizing big social data. *J. Big Data* **2017**, *4*, 3. [[CrossRef](#)]
6. Kitchin, R. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*; Sage: Thousand Oaks, CA, USA, 2014.
7. Schena, M.; Shalon, D.; Davis, R.W.; Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **1995**, *270*, 467–470. [[CrossRef](#)]
8. Marx, V. Biology: The big challenges of big data. *Nature* **2013**, *498*, 255–260. [[CrossRef](#)] [[PubMed](#)]
9. Quackenbush, J. *The Human Genome: The Book of Essential Knowledge*; Curiosity Guides, Imagine Publishing: New York, NY, USA, 2011.
10. Shendure, J.; Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **2008**, *26*, 1135–1145. [[CrossRef](#)] [[PubMed](#)]
11. Dehmer, M.; Emmert-Streib, F.; Graber, A.; Salvador, A. (Eds.) *Applied Statistics for Network Biology: Methods for Systems Biology*; Wiley-Blackwell: Weinheim, Germany, 2011.

12. Emmert-Streib, F.; Altay, G. Local network-based measures to assess the inferability of different regulatory networks. *IET Syst. Biol.* **2010**, *4*, 277–288. [[CrossRef](#)] [[PubMed](#)]
13. Shah, D.V.; Cappella, J.N.; Neuman, W.R. Big data, digital media, and computational social science: Possibilities and perils. *Ann. Am. Acad. Political Soc. Sci.* **2015**, *659*, 6–13. [[CrossRef](#)]
14. Conte, R.; Gilbert, N.; Bonelli, G.; Cioffi-Revilla, C.; Deffuant, G.; Kertesz, J.; Loreto, V.; Moat, S.; Nadal, J.P.; Sanchez, A.; et al. Manifesto of computational social science. *Eur. Phys. J. Spec. Top.* **2012**, *214*, 325. [[CrossRef](#)]
15. Emmert-Streib, F.; Yli-Harja, O.; Dehmer, M. Data analytics applications for streaming data from social media: What to predict? *Front. Big Data* **2018**, *1*, 1. [[CrossRef](#)]
16. Matz, S.C.; Kosinski, M.; Nave, G.; Stillwell, D.J. Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12714–12719. [[CrossRef](#)] [[PubMed](#)]
17. Emmert-Streib, F.; Dehmer, M. Data-driven computational social network science: Predictive and inferential models for Web-enabled scientific discoveries. *Front. Big Data* **2021**, *4*, 10.
18. Petricoin, E.F.; Liotta, L.A. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr. Opin. Biotechnol.* **2004**, *15*, 24–30. [[CrossRef](#)]
19. Marzese, D.M.; Hoon, D.S. Emerging technologies for studying DNA methylation for the molecular diagnosis of cancer. *Expert Rev. Mol. Diagn.* **2015**, *15*, 647–664. [[CrossRef](#)]
20. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63. [[CrossRef](#)] [[PubMed](#)]
21. Chen, C.E.; Harrington, R.A.; Desai, S.A.; Mahaffey, K.W.; Turakhia, M.P. Characteristics of digital health studies registered in ClinicalTrials.gov. *JAMA Intern. Med.* **2019**, *179*, 838–840. [[CrossRef](#)]
22. Steinhubl, S.R.; Topol, E.J. Digital medicine, on its way to being just plain medicine. *NPJ Digit. Med.* **2018**, *1*, 20175. [[CrossRef](#)] [[PubMed](#)]
23. Kostkova, P. Grand challenges in digital health. *Front. Public Health* **2015**, *3*, 134. [[CrossRef](#)]
24. Ali, F.; El-Sappagh, S.; Islam, S.R.; Ali, A.; Attique, M.; Imran, M.; Kwak, K.S. An intelligent healthcare monitoring framework using wearable sensors and social networking data. *Future Gener. Comput. Syst.* **2021**, *114*, 23–43. [[CrossRef](#)]
25. Ali, F.; El-Sappagh, S.; Islam, S.R.; Kwak, D.; Ali, A.; Imran, M.; Kwak, K.S. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Inf. Fusion* **2020**, *63*, 208–222. [[CrossRef](#)]
26. Kumar, V.; Jangirala, S.; Ahmad, M. An efficient mutual authentication framework for healthcare system in cloud computing. *J. Med. Syst.* **2018**, *42*, 1–25. [[CrossRef](#)]
27. Srinivas, J.; Das, A.K.; Kumar, N.; Rodrigues, J.J. Cloud centric authentication for wearable healthcare monitoring system. *IEEE Trans. Dependable Secur. Comput.* **2018**, *17*, 942–956. [[CrossRef](#)]
28. Fogel, A.L.; Kvedar, J.C. Artificial intelligence powers digital medicine. *NPJ Digit. Med.* **2018**, *1*, 1–4. [[CrossRef](#)]
29. Liu, Y.; Zhang, L.; Yang, Y.; Zhou, L.; Ren, L.; Wang, F.; Liu, R.; Pang, Z.; Deen, M.J. A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access* **2019**, *7*, 49088–49101. [[CrossRef](#)]
30. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access* **2019**, *7*, 167653–167671. [[CrossRef](#)]
31. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **2016**, *316*, 2402–2410. [[CrossRef](#)] [[PubMed](#)]
32. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)]
33. Labovitz, D.L.; Shafner, L.; Reyes Gil, M.; Virmani, D.; Hanina, A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. *Stroke* **2017**, *48*, 1416–1419. [[CrossRef](#)]
34. Emmert-Streib, F.; Yli-Harja, O.; Dehmer, M. Explainable Artificial Intelligence and Machine Learning: A reality rooted perspective. *WIREs Data Min. Knowl. Discov.* **2020**, *10*, e1368. [[CrossRef](#)]
35. Vayena, E.; Ienca, M. Digital medicine and ethics: Rooting for evidence. *Am. J. Bioeth.* **2018**, *18*, 49–51. [[CrossRef](#)]
36. Milosevic, Z. Ethics in Digital Health: A deontic accountability framework. In Proceedings of the 2019 IEEE 23rd International Enterprise Distributed Object Computing Conference (EDOC), Paris, France, 28–31 October 2019; pp. 105–111.
37. Lee, H.; Kim, S.; Kim, J.W.; Chung, Y.D. Utility-preserving anonymization for health data publishing. *BMC Med. Inform. Decis. Mak.* **2017**, *17*, 104. [[CrossRef](#)]
38. El Emam, K.; Rodgers, S.; Malin, B. Anonymising and sharing individual patient data. *BMJ* **2015**, *350*, h1139. [[CrossRef](#)] [[PubMed](#)]
39. Bukht, R.; Heeks, R. Defining, conceptualising and measuring the digital economy. Development Informatics working paper. *SSRN Electron. J.* **2017**. [[CrossRef](#)]
40. Bharadwaj, A.; El Sawy, O.A.; Pavlou, P.A.; Venkatraman, N. Digital business strategy: Toward a next generation of insights. *MIS Q.* **2013**, *37*, 471–482. [[CrossRef](#)]
41. Mesenbourg, T.L. *Measuring the Digital Economy*; US Bureau of the Census: Washington, DC, USA, 2001; pp. 5–6.
42. Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [[CrossRef](#)]
43. Xu, L.D.; Xu, E.L.; Li, L. Industry 4.0: State of the art and future trends. *Int. J. Prod. Res.* **2018**, *56*, 2941–2962. [[CrossRef](#)]

44. Kagermann, H.; Helbig, J.; Hellinger, A.; Wahlster, W. *Recommendations for Implementing the Strategic Initiative Industrie 4.0: Securing the Future of German Manufacturing Industry; Final Report of the Industrie 4.0 Working Group*; Forschungsunion; National Academy of Science and Engineering: München, Germany, 2013.
45. Wang, Y.; Ma, H.S.; Yang, J.H.; Wang, K.S. Industry 4.0: A way from mass customization to mass personalization production. *Adv. Manuf.* **2017**, *5*, 311–320. [[CrossRef](#)]
46. Chen, M.; Wan, J.; Li, F. Machine-to-machine communications: Architectures, standards and applications. *Ksii Trans. Internet Inf. Syst.* **2012**, *6*. [[CrossRef](#)]
47. Wan, J.; Chen, M.; Xia, F.; Di, L.; Zhou, K. From machine-to-machine communications towards cyber-physical systems. *Comput. Sci. Inf. Syst.* **2013**, *10*, 1105–1128. [[CrossRef](#)]
48. Stojmenovic, I. Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems. *IEEE Internet Things J.* **2014**, *1*, 122–128. [[CrossRef](#)]
49. Wu, F.J.; Kao, Y.F.; Tseng, Y.C. From wireless sensor networks towards cyber physical systems. *Pervasive Mob. Comput.* **2011**, *7*, 397–413. [[CrossRef](#)]
50. Madakam, S.; Lake, V.; Lake, V.; Lake, V. Internet of Things (IoT): A literature review. *J. Comput. Commun.* **2015**, *3*, 164. [[CrossRef](#)]
51. Monostori, L. Cyber-Physical Systems. In *CIRP Encyclopedia of Production Engineering*; Chatti, S., Tolio, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 1–8. [[CrossRef](#)]
52. Cheng, M. Sharing economy: A review and agenda for future research. *Int. J. Hosp. Manag.* **2016**, *57*, 60–70. [[CrossRef](#)]
53. Botsman, R.; Rogers, R. *What's Mine Is Yours: The Rise of Collaborative Consumption*; Harper Business: New York, NY, USA, 2010.
54. Schlagwein, D.; Schoder, D.; Spindeldreher, K. Consolidated, systemic conceptualization, and definition of the “sharing economy”. *J. Assoc. Inf. Sci. Technol.* **2020**, *71*, 817–838. [[CrossRef](#)]
55. Taeihagh, A. Crowdsourcing, sharing economies and development. *J. Dev. Soc.* **2017**, *33*, 191–222. [[CrossRef](#)]
56. Hrazdil, K.; Trottier, K.; Zhang, R. A comparison of industry classification schemes: A large sample study. *Econ. Lett.* **2013**, *118*, 77–80. [[CrossRef](#)]
57. Lim, W.M. Sharing economy: A marketing perspective. *Australas. Mark. J. (AMJ)* **2020**, *28*, 4–13. [[CrossRef](#)]
58. UNCTAD. *Digital Economy Report 2019: Value Creation and Capture—Implications for Developing Countries*; UNCTAD: New York, NY, USA, 2019.
59. Grall, A.; Dieulle, L.; Bérenguer, C.; Roussignol, M. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Trans. Reliab.* **2002**, *51*, 141–150. [[CrossRef](#)]
60. Lee, W.J.; Wu, H.; Yun, H.; Kim, H.; Jun, M.B.; Sutherland, J.W. Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia CIRP* **2019**, *80*, 506–511. [[CrossRef](#)]
61. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529. [[CrossRef](#)] [[PubMed](#)]
62. Gu, S.; Holly, E.; Lillicrap, T.; Levine, S. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In *Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 29 May–3 June 2017; pp. 3389–3396.
63. Shearer, C. The CRISP-DM model: The new blueprint for data mining. *J. Data Warehous.* **2000**, *5*, 13–22.
64. Tripathi, S.; Muhr, D.; Manuel, B.; Emmert-Streib, F.; Jodlbauer, J.; Dehmer, M. Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing. *arXiv* **2020**, arXiv:2007.14791.
65. Brynjolfsson, E.; McAfee, A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*; WW Norton & Company: New York, NY, USA, 2014.
66. Bughin, J.; Seong, J.; Manyika, J.; Chui, M.; Joshi, R. *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy*; McKinsey Global Institute: San Francisco, CA, USA, 2018.
67. Abu-Mostafa, Y.S.; Magdon-Ismael, M.; Lin, H.T. *Learning from Data*; AMLBook: New York, NY, USA, 2012; Volume 4.
68. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*; Elsevier Academic Press: San Diego, CA, USA, 2003.
69. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*; Springer: New York, NY, USA, 2009.
70. Emmert-Streib, F.; Yli-Harja, O.; Dehmer, M. A clarification of misconceptions, myths and desired status of artificial intelligence. *Front. Artif. Intell.* **2020**, *3*, 91. [[CrossRef](#)]
71. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2012; pp. 1097–1105.
72. Lee, H.; Pham, P.; Largman, Y.; Ng, A.Y. Unsupervised feature learning for audio classification using convolutional deep belief networks. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2009; pp. 1096–1104.
73. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2013; pp. 3111–3119.
74. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
75. Perera, N.; Dehmer, M.; Emmert-Streib, F. Named Entity Recognition and Relation Detection for Biomedical Information Extraction. *Front. Cell Dev. Biol.* **2020**, *8*, 673. [[CrossRef](#)]

76. Lau, J.H.; Baldwin, T. An empirical evaluation of doc2vec with practical insights into document embedding generation. *arXiv* **2016**, arXiv:1607.05368.
77. Mahler, R.P. *Statistical Multisource-Multitarget Information Fusion*; Artech House, Inc.: Norwood, MA, USA, 2007.
78. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [[CrossRef](#)]
79. Huynh, B.Q.; Li, H.; Giger, M.L. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J. Med. Imaging* **2016**, *3*, 034501. [[CrossRef](#)]
80. Van Engelen, J.E.; Hoos, H.H. A survey on semi-supervised learning. *Mach. Learn.* **2020**, *109*, 373–440. [[CrossRef](#)]
81. Caruana, R. Multitask learning. *Mach. Learn.* **1997**, *28*, 41–75. [[CrossRef](#)]
82. Emmert-Streib, F.; Yang, Z.; Feng, H.; Tripathi, S.; Dehmer, M. An introductory review of deep learning for prediction models with big data. *Front. Artif. Intell.* **2020**, *3*, 4. [[CrossRef](#)]
83. Cohen, G.; Afshar, S.; Tapson, J.; van Schaik, A. EMNIST: An extension of MNIST to handwritten letters. *arXiv* **2017**, arXiv:1702.05373.
84. Emmert-Streib, F.; Dehmer, M. Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 521–551. [[CrossRef](#)]
85. Emmert-Streib, F.; Dehmer, M.; Yli-Harja, O. Ensuring Quality Standards and Reproducible Research for Data Analysis Services in Oncology: A Cooperative Service Model. *Front. Cell Dev. Biol.* **2019**, *7*, 349. [[CrossRef](#)]
86. Baggerly, K.A.; Coombes, K.R. Deriving chemosensitivity from cell lines: Forensic bioinformatics and reproducible research in high-throughput biology. *Ann. Appl. Stat.* **2009**, *3*, 1309–1334. [[CrossRef](#)]
87. Djahel, S.; Doolan, R.; Muntean, G.M.; Murphy, J. A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. *IEEE Commun. Surv. Tutor.* **2014**, *17*, 125–151. [[CrossRef](#)]
88. Barba, C.T.; Mateos, M.A.; Soto, P.R.; Mezher, A.M.; Igartua, M.A. Smart city for VANETs using warning messages, traffic statistics and intelligent traffic lights. In Proceedings of the 2012 IEEE Intelligent Vehicles Symposium, Madrid, Spain, 3–7 June 2012; pp. 902–907.
89. Ejaz, W.; Naeem, M.; Shahid, A.; Anpalagan, A.; Jo, M. Efficient energy management for the internet of things in smart cities. *IEEE Commun. Mag.* **2017**, *55*, 84–91. [[CrossRef](#)]
90. Kankanhalli, A.; Charalabidis, Y.; Mellouli, S. IoT and AI for smart government: A research agenda. *Gov. Inf. Q.* **2019**, *36*, 304–309. [[CrossRef](#)]
91. Xu, F.; Uszkoreit, H.; Du, Y.; Fan, W.; Zhao, D.; Zhu, J. Explainable AI: A brief survey on history, research areas, approaches and challenges. In *CCF International Conference on Natural Language Processing and Chinese Computing*; Springer: Berlin, Germany, 2019; pp. 563–574.
92. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
93. Holzinger, A.; Kieseberg, P.; Weippl, E.; Tjoa, A.M. Current advances, trends and challenges of machine learning and knowledge extraction: From machine learning to explainable ai. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*; Springer: Berlin, Germany; 2018; pp. 1–8.
94. Gorecky, D.; Schmitt, M.; Loskyll, M.; Zühlke, D. Human–machine-interaction in the industry 4.0 era. In Proceedings of the 12th IEEE International Conference on Industrial Informatics (INDIN), Porto Alegre, Brazil, 27–30 July 2014; pp. 289–294.
95. Card, S.K. *The Psychology of Human–Computer Interaction*; CRC Press: Boca Raton, FL, USA, 2018.
96. Holzinger, A. Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Inform.* **2016**, *3*, 119–131. [[CrossRef](#)] [[PubMed](#)]
97. Carmigniani, J.; Furht, B.; Anisetti, M.; Ceravolo, P.; Damiani, E.; Ivkovic, M. Augmented reality technologies, systems and applications. *Multimed. Tools Appl.* **2011**, *51*, 341–377. [[CrossRef](#)]
98. Emmert-Streib, F.; Moutari, S.; Dehmer, M. The process of analyzing data is the emergent feature of data science. *Front. Genet.* **2016**, *7*, 12. [[CrossRef](#)]
99. Tuptuk, N.; Hailes, S. Security of smart manufacturing systems. *J. Manuf. Syst.* **2018**, *47*, 93–106. [[CrossRef](#)]
100. Isaak, J.; Hanna, M.J. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer* **2018**, *51*, 56–59. [[CrossRef](#)]
101. Buccafurri, F.; Lax, G.; Nicolazzo, S.; Nocera, A. Comparing Twitter and Facebook user behavior: Privacy and other aspects. *Comput. Hum. Behav.* **2015**, *52*, 87–95. [[CrossRef](#)]
102. Winfield, A.F.; Jirotko, M. Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2018**, *376*, 20180085. [[CrossRef](#)] [[PubMed](#)]
103. Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; Galstyan, A. A survey on bias and fairness in machine learning. *arXiv* **2019**, arXiv:1908.09635.
104. Ntoutsis, E.; Fafalios, P.; Gadiraju, U.; Iosifidis, V.; Nejdil, W.; Vidal, M.E.; Ruggieri, S.; Turini, F.; Papadopoulos, S.; Krasanakis, E.; et al. Bias in data-driven artificial intelligence systems? An introductory survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, e1356. [[CrossRef](#)]
105. Courtland, R. The bias detectives. *Nature* **2018**, *558*, 357–360. [[CrossRef](#)] [[PubMed](#)]
106. Yapo, A.; Weiss, J. Ethical implications of bias in machine learning. In Proceedings of the 51st Hawaii International Conference on System Sciences, Hilton Waikoloa Village, HI, USA, 3–6 January 2018.

107. Challen, R.; Denny, J.; Pitt, M.; Gompels, L.; Edwards, T.; Tsaneva-Atanasova, K. Artificial intelligence, bias and clinical safety. *BMJ Qual. Saf.* **2019**, *28*, 231–237. [[CrossRef](#)]
108. Franks, B. *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*; John Wiley & Sons: New York, NY, USA, 2012; Volume 49.
109. Halper, F. Advanced Analytics: Moving Toward AI, Machine Learning, and Natural Language Processing. TDWI Best Practices Report. 2017. Available online: <https://analyticsconsultores.com.mx/wp-content/uploads/2019/03/Advanced-Analyhtics-Moving-Toward-AI-Machine-Learning-and-Natural-Language-Processing-Fern-Halper-TDWI-SAS-2017.pdf> (accessed on 4 March 2021).
110. Ma, Y.; Zhu, L. A review on dimension reduction. *Int. Stat. Rev.* **2013**, *81*, 134–150. [[CrossRef](#)]
111. Tukey, J. *Exploratory Data Analysis*; Addison-Wesley: New York, NY, USA, 1977.
112. Baldi, P. Autoencoders, unsupervised learning, and deep architectures. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, Bellevue, WA, USA, 27 June 2012; pp. 37–49.
113. Emmert-Streib, F.; Dehmer, M. Understanding Statistical Hypothesis Testing: The Logic of Statistical Inference. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 945–961. [[CrossRef](#)]
114. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin, Germany, 2006.
115. Emmert-Streib, F.; Dehmer, M. High-Dimensional LASSO-Based Computational Regression Models: Regularization, Shrinkage, and Selection. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 359–383. [[CrossRef](#)]
116. Kingma, D.P.; Mohamed, S.; Rezende, D.J.; Welling, M. Semi-supervised learning with deep generative models. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2014; pp. 3581–3589.
117. Li, Y. Deep reinforcement learning: An overview. *arXiv* **2017**, arXiv:1701.07274.
118. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [[CrossRef](#)]
119. Han, T.; Liu, C.; Yang, W.; Jiang, D. A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. *Knowl. Based Syst.* **2019**, *165*, 474–487. [[CrossRef](#)]
120. Skobelev, P.; Borovik, S.Y. On the way from Industry 4.0 to Industry 5.0: From digital manufacturing to digital society. *Industry 4.0* **2017**, *2*, 307–311.
121. Granovetter, M. Economic action and social structure: The problem of embeddedness. *Am. J. Sociol.* **1985**, *91*, 481–510. [[CrossRef](#)]
122. Dufva, T.; Dufva, M. Grasping the future of the digital society. *Futures* **2019**, *107*, 17–28. [[CrossRef](#)]
123. Makridakis, S. The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures* **2017**, *90*, 46–60. [[CrossRef](#)]
124. Brockman, J. *Possible Minds: Twenty-Five Ways of Looking at AI*; Penguin Books: London, UK, 2020.
125. Joy, B. Why the future doesn't need us. *Wired Mag.* **2000**, *8*, 238–262.
126. Marcus, G.; Davis, E. *Rebooting AI: Building Artificial Intelligence We Can Trust*; Pantheon: New York, NY, USA, 2019.
127. Helbing, D. Societal, economic, ethical and legal challenges of the digital revolution: From big data to deep learning, artificial intelligence, and manipulative technologies. In *Towards Digital Enlightenment*; Springer: Berlin, Germany, 2019; pp. 47–72.
128. Lankshear, C.; Knobel, M. *Digital Literacies: Concepts, Policies and Practices*; Peter Lang: Bruxelles, Belgium, 2008.
129. Herrlich, P. The responsibility of the scientist: What can history teach us about how scientists should handle research that has the potential to create harm? *EMBO Rep.* **2013**, *14*, 759–764. [[CrossRef](#)] [[PubMed](#)]