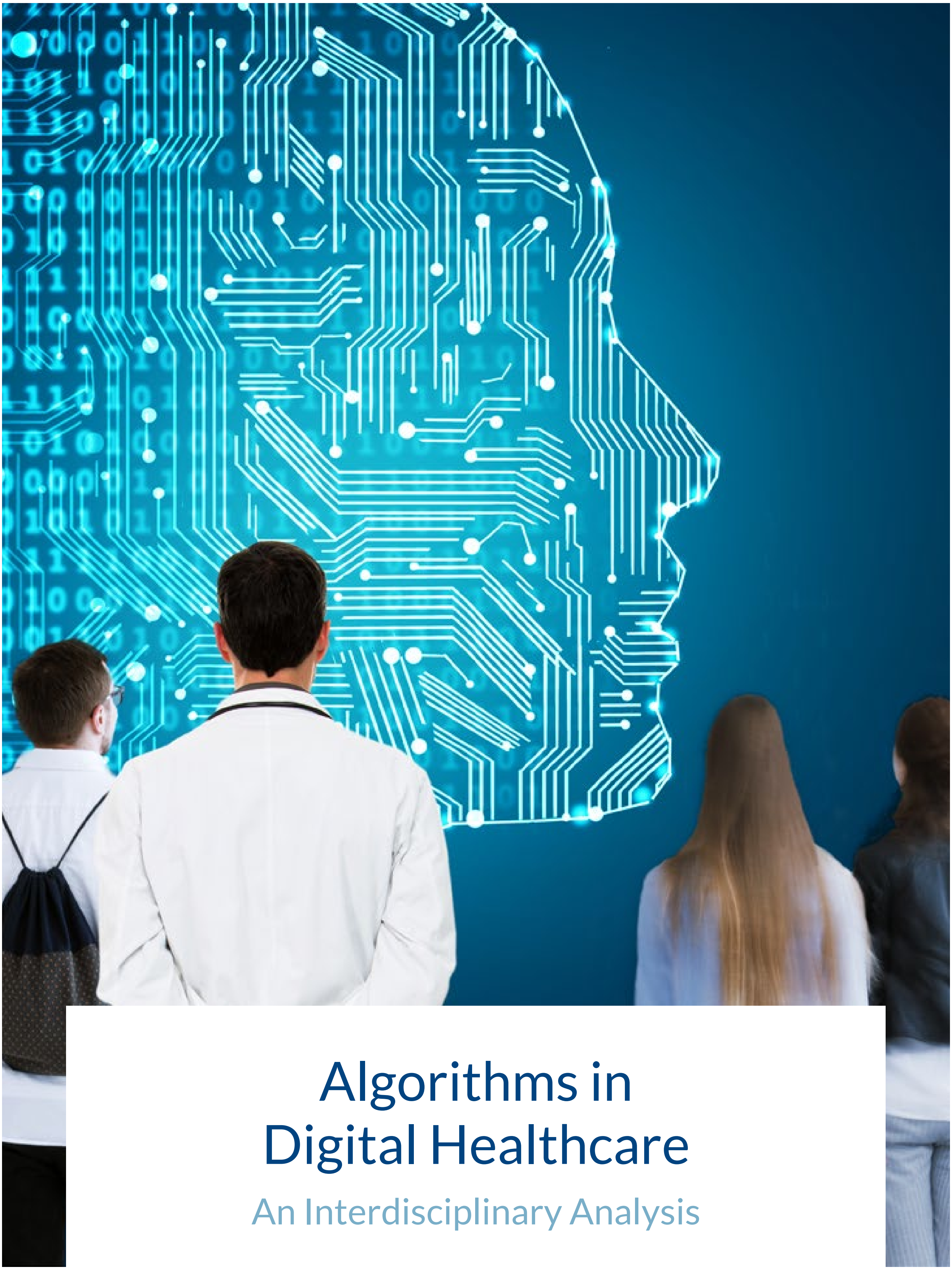


Algorithms in Digital Healthcare

An Interdisciplinary Analysis

ceres
cologne center for
ethics, rights, economics, and social sciences
of health

BertelsmannStiftung

Algorithms in Digital Healthcare

An Interdisciplinary Analysis

Authors
Marc Jannes, Dr. Minou Friele,
Christiane Jannes and Prof. Dr. Christiane Woopen,
Cologne Center for Ethics, Rights, Economics,
and Social Sciences of Health (ceres), University of Cologne,
on behalf of the Bertelsmann Stiftung

September 2019

Content

Foreword

Algorithms have become a part of our everyday lives. They recommend advertisements to us on online-shopping sites, provide estimates of our creditworthiness, and determine what information we will see in our social-media applications. In the medical field too, the opportunities for utilizing algorithms are manifold. New algorithm-based applications are constantly appearing on the market. Indeed, the pace of development is rapidly increasing. The large U.S.-based technology companies, for example, want nothing less than to revolutionize medicine.

Without question, algorithms and the use of artificial intelligence technologies can significantly improve treatment processes, thus contributing to a more efficient use of resources. This fundamental opportunity-oriented attitude should serve as the point of departure for dealing with the topic. However, the use of algorithms also presents us with new challenges – along with the question of just what kind of digital progress we as a society want. Not everything that is technically feasible is necessarily societally desirable. Any examination of healthcare-related issues touches quickly on fundamental ethical questions.

Thus, what opportunities and risks are posed by the use of algorithms in healthcare? The present study is intended to sort through and classify these issues. In it, the experts from University of Cologne's *ceres* program offer an overview of the areas of medicine in which algorithms are already being used today, and in which they are likely to be employed in the future. In addition, the authors outline the potential improvements associated with the use of such technology, along with areas of concern.

To us, the study's orientation toward practical use within the healthcare field is of critical importance. The authors examine specific use cases and derive key ethical research questions from them – for individuals, for state institutions and for society as a whole. The examples analyzed illustrate the breadth of potential uses for algorithms, from predicting mental illnesses among social-media users to providing data-driven support for physicians' therapeutic decisions, and even to helping paralyzed people regain their mobility.

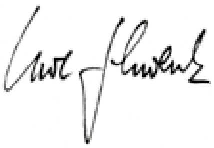
The study makes it clear just how significantly algorithms could contribute to improving care. But it also highlights the questions raised by their use, regarding issues of distributional justice and protection from discrimination, regarding liability for algorithm-based decisions, and regarding the changes looming in the relationship between doctors and patients. Not least, it casts a spotlight on the issue of trust in the healthcare system itself. In our view, we as a society therefore need to come to a widespread understanding of which developments we should collectively endorse and pursue, as well as a sense of where our "red lines" must be drawn.

The public discussion on the issues raised in this way should in turn serve as the basis for a forward-looking, creative healthcare policy – a policy that establishes conditions allowing beneficial applications to be integrated more quickly into everyday care settings. A policy containing instruments simultaneously able to protect vital individual and societal interests. With regard to the dimension and pace of digital change, our foresight must no longer stretch only until the next election. Policymaking must be guided by long-term visions of the future – and it must reach beyond purely national ambitions.

In our project “The Digital Patient,” we want to take a closer look at the ethical and societal challenges associated with digitalization in the healthcare sector. The present study marks the commencement of this effort, forming the basis for future discussions of algorithms in healthcare.

We hope you find it interesting.

Dr. Brigitte Mohn
Member of the
Bertelsmann Stiftung board



Uwe Schwenk
Director Program
“Improving Healthcare –
Informing Patients”
Bertelsmann Stiftung.

1 Algorithms: Classification and definition of terms

As discussions about big data have highlighted the associated societal transformations and challenges, algorithms have increasingly attracted the public's attention. In the healthcare sector too, the use of algorithms is deemed to hold great potential. Intelligent systems are expected to help to improve patient care and make the functioning of the healthcare system significantly more efficient, for example by supporting diagnostic and therapeutic decisions (Rasche 2017: 8; Dörn 2018: 349).

1.1 Simple algorithms

As yet, the academic literature has not settled on a single exact, universally applicable definition of the term "algorithm" (Broy and Spaniol 1999: 12; Nahrstedt 2018: 3). One reason for this is that a precise definition would arbitrarily limit its own meaning (Solymosi and Grude 2017: 1). The term is thus understood differently in different contexts, and is used at times in a non-specific, instrumentalized way (Burkhardt 2017).

It is indisputable that when discussing an algorithm, we are talking about a problem-solving procedure that employs finite sequences of clearly defined and actually realizable sub-actions (Fischer and Hofer 2008: 32; Schubert and Schwill 2011: 4). It thus represents a precisely defined working method for the solution of a (mathematical) problem (von Rimscha 2017: 3; Nahrstedt 2018: 1). However, for the calculation of this solution, the algorithm needs specific information. Today, algorithms are increasingly modeled using computers. However, for a computer to be able to carry out this kind of problem-solving procedures for different problems, the abstract algorithm must be reformulated into a concrete set of instructions (Schubert and Schwill 2011: 4 f.). This means coding it in a sequence of instructions that is comprehensible to a computer (Fischer and Hofer 2008: 32). This exact description and coding is called a program (Schubert and Schwill 2011: 4 f.).

Algorithms can be classified on the basis of different levels of complexity, and have characteristic qualities such as determinacy or finiteness (Fischer and Hofer 2008: 32; Broy and Spaniol 1999: 12; Solymosi and Grude 2017: 5). Determinacy means that given the same input values, the same output values will always be produced, across multiple repetitions (Fischer and Hofer 2008: 221). Finiteness, on the other hand, describes the limited length of an algorithm; that is, the fact that it is composed of a bounded number of instructions which are themselves of limited length (Fischer and Hofer 2008: 306).

Overall, "algorithm" is one of the most important concepts in mathematics, and algorithms are in practice used in all known sciences and economic sectors (Nahrstedt 2018: 1; on the historical development of algorithms, see ibid.: 1 f.). Applications that implement an algorithm often integrate very different areas of knowledge in doing so, for instance in the description of the rules used in linguistics, or by representing specific patterns of behavior from the social sciences (ibid.: 3). Because of this, algorithms from one field can also be transferred to other fields (ibid.).

The increasing interconnectedness of things, particularly due to the internet, is leading to an exponential growth in data. This in turn allows algorithms to draw on a rising number of sources and an ever-greater quantity of data in the solution of problems (Mainzer 2016: 157; Wu et al. 2014). In the healthcare sector, this includes research and development data, laboratory and pharmaceutical data, clinical administrative data such as electronic health records or insurance data, and patient-generated data like personal health and activity data (Lipworth et al. 2017). In addition, other data relevant to health may also be accessed, depending on the specific context. Examples here might include credit card and online-activity data, census or criminality data, social-media data, or other online resources (ebd.).

1.2 Classification

Algorithms can be distinguished on the basis of different classes. However, these are not always clearly separate from one another, and indeed occasionally show some overlap (Nahrstedt 2018).

- **Deterministic algorithms** consistently deliver the same result given a constant input.
- **Randomized (non-deterministic) algorithms** incorporate elements of a randomness event.
- **Iterative algorithms** define a starting value and then, based on the previously known quantities, produce an interim result with each computational step. This in turn again serves as the basis for the performance of the next calculation (a further differentiation is made between algorithms with a known or an unknown number of iteration steps) (von Rimscha 2017: 6).
- **Recursive algorithms** are a special form of iterative algorithms in which at least one of the computational steps consists in the algorithm calling itself.

Other distinctions, for example between decision-support and optimization algorithms, relate to the way the problem is posed (Nahrstedt 2018):

- **Decision support algorithms** are often used in complex expert systems. These systems store large amounts of data relevant to a specific field of knowledge. Decision support algorithms are programmed to draw automated conclusions from the data analyzed, for example by producing diagnoses for specific situations or concrete solution proposals for problems (Mainzer 2016: 12). However, the "decisions" generated in this way must be clearly distinguished from those of human experts, for instance in the case of medical decisions. For example, decision support algorithms possess no general background knowledge, have no feelings or personal motivational principles, and cannot make judgements based on ethical values (ibid.).
- **Optimization algorithms:** Algorithms can be further distinguished on the basis of whether they offer an optimal solution for a problem, or instead generate one or multiple potential solutions that cannot or may not necessarily be viewed as optimal. Currently, the former are used in the context of cost optimizations, and for waiting line and transport problems (Nahrstedt 2018: 5).

Types of algorithms

Deterministic algorithms	always produce the same result, when given the same input.
Randomized (non-deterministic) algorithms	contain elements of randomness in their results.
Iterative algorithms	define a starting value and then, based on the previously known quantities, produce an interim result with each computational step. This in turn serves as the basis for the performance of the next calculation.
Recursive algorithms	are a special form of iterative algorithm in which at least one of the computational steps consists in the algorithm calling itself.
Decision support algorithms	are often used in complex expert systems. These systems store large amounts of data that are relevant to a specific field of knowledge. Decision support algorithms are programmed to draw automated conclusions from the data analyzed, for example by producing diagnoses in specific situations or proposing concrete solutions for problems
Optimization algorithms	produce an optimal solution for a problem. Currently, these are used in the context of cost optimization, and for waiting-line and transport problems, for example (Nahrstedt 2018: 5)

1.3 Artificial intelligence

According to Lämmel and Cleve (2012), artificial intelligence (AI) is a "sub-field of computer science that attempts to emulate human approaches to problem-solving on computers in order to arrive at new or more efficient solutions" (ibid.: 13). So-called **machine learning algorithms** also fall into the field of artificial intelligence. These are used in machine learning systems supported by new and increasingly powerful hardware and software platforms, and can help detect complex relationships within vast amounts of data, without each individual computational step having to be explicitly programmed (Hecker et al. 2017). Machine learning is based on the idea of gaining knowledge from experience (Mitchell et al. 2009). To this end, a computer is given concrete sample data from which it is supposed to derive a general rule (von Rimscha 2017: 132).

In this context, machine-learning processes are often implemented in artificial, multilayer neural networks. These networks consist of neurons that are modeled on the synapses of people or animals (Nürnberger and Bugiel 2016). The interactions between the neurons produce an artificial neural network. This is provided with data at its input layer, and the results are delivered at the output layer (von Rimscha 2017: 158). Additional hidden layers can be located between these two (ibid.). Neural networks are "capable of learning", and can, for example, distinguish between various objects in an image (e.g., lampposts and trees) on the basis of rules selected by the network itself (Nürnberger and Bugiel 2016). To make this possible, thousands of images of an object simply have to be fed into the neural network, on the basis of which the object will be reduced to the similarities in the images (ibid.). This enables the algorithm to recognize the object in future images with a high level of accuracy (ibid.).

This development is referred with as **"deep learning"**: The recognition of an object or pattern, such as an image, is accomplished through a process that goes gradually "deeper". In the beginning, only individual components are recognized, followed in the course of the proceeding by entire clusters, and finally by the whole (Mainzer 2016: 110).

A variety of different kinds of "learning" within neural networks can be distinguished, including **supervised learning, unsupervised learning and reinforcement learning.** In the case of supervised learning, the network is trained by providing it with predetermined examples for which the desired result is already known (von Rimscha 2017: 159). The prototype to be learned (e.g., the recognition of a pattern) is thus known, and the synaptic weightings within the neural network are adjusted by the machine learning algorithm until the activity pattern of the results deviate from this prototype to the smallest degree possible (Mainzer 2016: 115). As the process unfolds, each erroneous deviation can be measured against this prototype (ibid.). Outputs that deviate from the correct function value result in a correction by a "teacher" (ibid.: 119). For example, this could be the prototype of a pattern as trained (ibid.).

In the case of a unsupervised machine learning system, by contrast, the learning algorithm independently detects new patterns and correlations without making reference to predetermined prototypes (Mainzer 2016: 116). Thus, no target results are predefined, and there is no overarching authority such as a prototype or "teacher". Rather, classification takes place spontaneously on the basis of detected characteristics (ibid.). Inside the neural network, this takes place on the basis of competition and selection processes between and among the various layers' neurons. A neuron learns when it "wins the competition with the rest of the neurons within a cluster. In this way, similarities between correlations and relationships are strengthened" (ibid.).

Reinforcement learning falls between these two learning procedures. Here, as with supervised learning, a goal is predefined. However, as with unsupervised learning, the algorithm must independently find its way to the goal's realization. After each partial step toward realization of the goal, information about the environment is fed back into the algorithm, describing its progress – good or bad – in realizing the predefined goal (trial and error). This feedback is then used by the algorithm to optimize the result (Mainzer 2016: 119).

So-called **evolutionary algorithms** play a special role among machine learning algorithms. Here, the point of departure is the insight that living beings store their design plans in their genes, and that the Darwinian law of the "survival of the fittest" comes into play as organisms multiply (von Rimscha 2017: 58). This schema is translated into an evolutionary algorithm. Evolutionary algorithms use natural evolution as a model. By employing a simulated evolutionary process, they develop suitable approximation solutions for a predetermined problem (Weicker 2015: 1). The program is not written in detail by a programmer, but is instead generated in an evolutionary process (Mainzer 2016:93). In this regard, it is a stochastic, metaheuristic optimization procedure (Kruse and Borgelt 2015: 157). However, the application of such algorithms does not necessarily lead to success, because the contingencies of evolutionary algorithms, just as in nature, can lead to mistakes and unsuccessful attempts (Mainzer 2016: 94).

Many of these algorithms are already used today in the provision of health care. For example, they are employed in medical image processing for the early detection and diagnosis of tumors (Dörn 2018: 352). Another example is the use of algorithms to process vital signs and other data in order to identify abnormalities and take immediate countermeasures as necessary. The underlying data is increasingly provided by so-called health monitoring systems such as sensor technologies, which continuously collect vital data (Dörn 2018: 353).

The present report is intended to offer an illustrative overview of various areas in which algorithms are being applied within the healthcare sector. In so doing, we will describe both current and expected future developments. In addition, we will critically discuss the opportunities and challenges associated with these developments. Finally, we will point out areas in which additional action and future research are needed.

2 Central research question

This report provides an overview of the healthcare-related areas in which algorithms are today being employed, and in which they are likely to be employed in the future. It additionally examines the functions fulfilled by technology of this kind. It also examines which areas of healthcare provision can expect to see improvements thanks to the use of algorithms, and which are likely to experience increased problems.

In addition, it identifies and analyzes ethical questions associated with the introduction and use of algorithms in healthcare both today and in the future, at the individual, institutional and societal levels. Sociopolitical problems and questions constitute a further focus.

3 Methodology

With the aim of providing an overview of the diverse areas in which algorithms are employed in the healthcare sector, as well as to inform the analysis of the associated (ethical) opportunities and challenges, the authors of this study conducted an exploratory literature review. This research was conducted between 18 January 2018 and 5 April 2018. It entailed a search of two multidisciplinary and two discipline-specific databases, selected on the basis of thematic relevance. These included Web of Science and Google Scholar (the multidisciplinary databases), along with PubMed (a specialist database for the medical field) and PhilPapers (a specialist database for the fields of philosophy and ethics). Search terms in both the English and German language were used, such as: "algorithm*" AND "health" OR "medic*" OR "clinic*" AND "ethic*". Given the dynamic pace of technological development, the exploratory research was limited to publications which appeared between 2008 and 2018. Because the overall thematic field is not yet well-structured, we did not specify strict inclusion or exclusion criteria for the selection of literature, as this may otherwise have led to the exclusion of important sources. "Gray literature" and reports were also evaluated.

Due to the high number of hits in the individual databases, the results were sorted based on "relevance" or "best match." For each database, the first 150 articles were incorporated into the analysis. In addition, only publications in Englisch or German were included.

The selection of literature initially took place after viewing only the title or article abstract. Duplicates were manually discarded. When the subsequent examination of the full text determined that an article lacked relevance, it was also discarded. If hits deemed to be relevant, bibliography lists and other publications by the associated authors were also reviewed in search of potentially relevant articles (a "snowballing" process).

In addition to the literature research in the scholarly databases, the authors carried out further exploratory research using various internet search engines such as Google. The aim was to identify additional potentially relevant company activities and research projects in the area of healthcare-related algorithms. The search terms used for this purpose included "algorithm based medical products," "algorithm health product," "health machine learning" and "health deep learning," among others. The company and research-project websites identified in this manner were then further examined for additional relevant links and pertinent literature.

A total of, 77 publications, including gray literature, were included in the analysis of application areas. An additional 40 publications, including gray literature, were considered as a part of the normative analysis.

Following the literature review, various algorithm-based systems in different application areas within the healthcare sector were placed into categories. In addition, the opportunities and ethical challenges associated with the use of algorithms in healthcare were identified. The authors refered to the general and current ethical discourse, and relied on their own critical examination of the topic. In our discussion of individual case examples, we additionally classified the opportunities and challenges associated with individual application areas on the basis of three ethical levels: the individual, the institutional and the sociopolitical.

4 Current research fields and application areas

The literature review showed that algorithms are already being used in numerous areas within the healthcare sector. In addition, the use of algorithms is being researched and further developed in a wide range of application areas. The applications can be grouped into the following areas: **public health, healthcare provision processes (health services research), medical research, prevention, prediction and risk profiling, diagnostic procedures, therapeutic procedures, prognoses, rehabilitation, and nursing care**.

Areas of algorithm application

- **Public health:** This relates to the science and practice of preventing diseases, extending life, and promoting physical and mental health at the population level, taking into account the equitable allocation and efficient use of available resources (The German Public Health Association 2018.).

- **Healthcare provision processes (health services research):** When speaking of healthcare provision processes, this report is referring to logistic and / or organizational operations within the healthcare sector. One focus here is on health services research. This refers to the scientific study of the care provided to individuals and the population using health-relevant products and services under everyday conditions (Working Group Health Services Research at the Scientific Advisory Board of the German Medical Association 2004).

- **Medical research:** Encompasses all actions that are conducted in a methodologically guided manner and are aimed at obtaining insights in the field of the medical sciences (The German Reference Centre for Ethics in the Life Sciences 2018).

- **Prevention:** Refers to the entirety of all medical measures that prevent, make less probable or delay the occurrence of health-related harm (Pschyrembel Online 2018a).

- **Prediction / Risk profiling:** Refers to the probability-based forecasting of the likelihood of occurrence of specific events or states, such as a disease (Pschyrembel Online 2017).

- **Diagnostics:** Diagnostics are medical procedures with a diagnosis serving as their endpoint. Such procedures include the collection of the patient's medical history, the physical examination, instrument-based methods, laboratory and microbiological diagnostic procedures, and differential diagnostic considerations (Braun 2018a).

- **Therapeutic procedures** This refers to the treatment of diseases, disabilities and injuries. The goal of therapeutic procedures is to cure, eliminate or relieve symptoms, and to restore physical and mental functions (Pschyrembel Online 2018b).
- **Prognosis:** In the medical field, a prognosis is the prediction of the progression, duration and outcome of a disease based on scientific knowledge and experience. A prognosis can refer to the probability of survival, or sub-areas such as the ability to work (Braun 2018b).
- **Rehabilitation:** This refers to the (re-)integration of a sick, physically disabled or mentally disabled person into professional and social life (Duden).
- **Nursing care:** This term describes activities carried out to support or to promote human well-being and survival. Nursing care includes both preventative and immediate care for sick, disabled or dying people, along with medical protective custody and caregiving, as well as the more general promotion of health and the prevention of diseases (Pschyrembel Online 2018c).

Table 1 in the annex offers an overview of current and future areas in which algorithms are being or may be applied in healthcare. In the following sections, we present a number of examples of current research fields related to the use of algorithms in healthcare. Subsequently, we offer a short overview of algorithm-based products already in the market within selected application areas.

4.1 Research fields – Examples

Diagnostic procedures

In Germany, researchers from the Technical University of Darmstadt (TU Darmstadt) are working on algorithms that can accurately detect **atrial fibrillation**, and thus can improve the associated diagnostic procedures.[1] To this end, they are working with Happitech, a Dutch firm, which has developed an app that enables cardiac arrhythmias to be identified with the help of a smartphone. This uses the technique of photoplethysmography, in which blood vessels are illuminated and examined using the smartphone's light. The algorithms are able to process and classify the signals measured in such a way as to distinguish between a normal sinus rhythm and atrial fibrillation. The researchers train the algorithms using thousands of heartbeat patterns for which diagnoses recorded in hospitals were available. As a result, they are already achieving accuracy rates of nearly 100 percent.

At Stanford University in California, researchers have developed an algorithm that **analyzes chest X-rays**, thus enabling the diagnosis of up to 14 different diseases. The technology employed in this case is a machine learning algorithm that has been trained with more than 100,000 thoracic X-ray images. Tests have shown that the algorithm is better than human radiologists at detecting a pulmonary inflammation on an X-ray image. This presents great opportunities for medical practice. In the United States alone, more than 1 mil-

1 www.innovations-report.de/html/berichte/medizintechnik/signale-des-herzens-tu-darmstadt-entwickelt-algorithmen-zur-erkennung-von-vorhofflimmern.html; www.spg.tu-darmstadt.de/spg/index.en.jsp

lion people every year suffer from pneumonia, a condition that is particularly difficult to diagnose on x-rays (Rajpurkar et al. 2017).

Therapeutic procedures

Researchers from the Fraunhofer Institute for Digital Medicine (MEVIS) have developed algorithms that enable a precise **analysis of image data from liver patients**, which can then be used for surgical therapy. The algorithm-based software enables the creation of a detailed three-dimensional model of the liver and its associated vascular systems. Through use of the algorithms, the optimal surgical incisions can be calculated before an operation, and particularly critical sections can be identified. On this basis, surgeons can better plan an operation, and patient safety can be significantly increased, as even the tiniest incision errors could impair the organ's functioning (Fraunhofer-Institut 2018).

A team of researchers at the Technical University of Munich is developing and currently testing an algorithm that can be used in conjunction with so-called cochlear implants for **hearing-loss therapy**, optimizing the transfer of the acoustical signals transmitted to the brain. Those wearing such implants often have difficulties filtering out individual acoustical signals from loud environments such as a mixture of voices. The algorithm developed ensures that there is a short time lag associated with the processing of an acoustic signal. Such minimal time delays are normal with binaural hearing – that is, hearing with two ears – and will thus not be perceived by people with the cochlear implants. The initial tests have already demonstrated significant benefits associated with the new algorithm. It thus offers people affected by hearing loss the prospect of significantly improved comprehension in complex listening situations (Federal Ministry of Education and Research 2018).

4.2 Application areas – Examples

Public health

The products of MedAware offer one example from public health. These have the goal of **identifying physicians' prescription errors**, and accordingly providing the doctor with a warning.[2] To do so, the company uses machine-learning algorithms to process massive amounts of data from millions of electronic health records. This captures the prescription patterns shown by all physicians in order to determine the "normal" spectrum of treatment. On this basis, MedAware creates a mathematical model that describes this real treatment pattern. The assumption is that a physician prescription that deviates from this standard treatment spectrum has a high probability of being erroneous.

Based upon this, the firm offers various decision-making aids and risk-management tools. One such example is the MedAware Alerting System (MedAS). Whenever a physician enters a prescription into the system, this system conducts a real-time evaluation of the medication involved with reference to the current patient profile. If MedAS detects a deviation from the normal treatment spectrum for patients with similar profiles, the physician receives a warning notice. The system is also continuously updated with new patient data such as blood-test values and diagnoses. The physician additionally receives a mes-

2 www.medaware.com/our-products/

sage if new information about a medication comes to light, such as previously unknown side effects or drug interactions.

Healthcare provision processes (health services research)

MedAware also produces the MedAware Risk Management (MedRIM), a decision-support tool aimed at improving healthcare provision processes.[3] This product is intended to **optimize risk-management and quality-control processes**, while additionally providing physicians with feedback on potentially erroneous prescriptions. The system, which can be started on demand or at specified intervals, collects patient and prescription data and uses algorithms to compare this with historical data, for instance within a given hospital. If MedRIM discovers outliers, these are flagged for further analysis by an expert. This can help to identify the hospital departments with exceptionally high rates of prescription errors, the medications that are most frequently prescribed erroneously, and even the individual physicians prone to a high error rate. On the basis of these results, inadequate processes can be improved, and patent safety can be increased overall.

Prevention

Philips Healthcare has developed its IntelliVue Guardian System, **an inpatient patient-monitoring system, for preventive purposes.** [4] The system uses artificial intelligence technology to identify crisis situations that may endanger a patient's life, thus facilitating an early intervention. In the product description, Philips Healthcare explains that its early-warning system "combines software, algorithms for the support of clinical decisions, and mobile connectivity." Wearable devices play a key role in this regard. For example, a clinician can place a wireless device equipped with sensors on a patient's wrist to track vital signs such as blood pressure. The IntelliVue Guardian Solution software then uses machine-learning techniques to identify significant changes in the patients' vital parameters. The algorithms employed here have been trained with large datasets containing similar patient data. If the algorithm detects a significant deviation, data is transferred to IntelliVue Monitors or mobile devices, so that the nursing staff can be automatically notified.

Prediction / Risk profiling

AliveCor's KardiaBand™ and SmartRhythm™ products fall into the predictive category.[5] It replaces the original bracelet of the Apple Watch. Here, an FDA-approved electrocardiogram (ECG) device is combined with an analysis algorithm that draws on artificial-intelligence technology to monitor heart-rate and activity data using the Apple Watch. According to AliveCor, the system uses an artificial neural network that continually compares the detected heartrate with expected future heartrate patterns. If the system detects a heartrate and activity pattern that deviates from the expected form, the person is asked to perform an ECG. The KardiaBand sensors mounted on the wristband can then immediately perform a 30-second ECG, which can for example detect existing cardiac arrhythmias or the signs of a threatening heart disease.

3 www.medaware.com/our-products/
4 www.usa.philips.com/healthcare/clinical-solutions/early-warning-scoring/intellivue-guardian-ews
5 www.mindtecstore.com/Products-and-Infos-about-AliveCor

Diagnostic procedures

In the field of diagnostic procedures, NVIDIA and GE Healthcare have jointly developed a program aimed at improving the **speed and accuracy of computed tomography (CT) scans.**[6] With this tool, algorithms are used to identify small patterns of organ damage that are often overlooked when physicians examine a scan. By accurately capturing these small details, the technology enables faster diagnoses while simultaneously reducing error rates. In addition, the time needed to carry out the scan itself is diminished, reducing the patient's radiation exposure.

Israel's Zebra Medical Vision, a vendor specializing in machine learning, has developed its Zebra's Radiology Assistant product for similar diagnostic purposes.[7] The system can **automatically analyze computed tomography scans for various diagnostic findings**, and forward the results in real time to physicians or hospital systems, as needed. The product is intended to help radiologists manage their rising workloads without sacrificing quality. According to the company's statements, Zebra Medical Vision trains its algorithms using millions of clinical-image data points, with the goal of enabling faster diagnoses and facilitating the parallel capture of multiple diagnostic findings. This process entails the use of machine-learning techniques.

According to the manufacturer, the system can be used for a wide range of purposes, such as identifying patients with an elevated risk of cardiovascular, lung or bone diseases, as well as other diseases. In March 2018, Zebra Medical Vision additionally announced that it had received European Union approval for its algorithm that detects intracranial bleeding, a part of its Deep Learning Imaging Analytics platform. This algorithm is meant to detect cerebral hemorrhages accurately and with minimal delay, with the aim of offering additional support to physicians in acute care.[8]

The ImmunoXpert™ test produced by MeMed represents another development in the area of diagnostic procedures.[9] According to the company, this is an innovative **in-vitro diagnostic test designed to distinguish bacteria from viruses.** The ImmunoXpert™ test measures immune-system biomarkers within blood serum, and uses pattern-recognition algorithms to distinguish accurately between these types of infections. This offers the significant advantage of possible diagnoses even when the site of infection is inaccessible or unknown. The test could significantly reduce the quantity of antibiotics prescribed, because – according to the company – one primary reason for the improper administration of antibiotics is the clinical difficulty in distinguishing bacterial from viral infections. This leads to ineffective treatments and to the appearance of antibiotic-resistant strains of bacteria, and results in annual costs of an estimated several billion dollars worldwide. ImmunoXpert™ is already approved for use within the European Union (CE-IVD certified).

A company called HeartFlow has developed a **non-invasive diagnostic software system** for patients with coronary heart-disease symptoms.[10] For these people, a computed tomography (CT) scan is initially carried out; then, if there are signs of arteriosclerosis, for example an analysis with HeartFlow is conducted. For this purpose, the system uses the CT scan

6 www.techemergence.com/ai-medical-devices-three-emerging-industry-app-li-cat-ions; https://blogs.nvidia.com/blog/2017/11/26/ai-medical-imaging/
7 https://us.zebra-med.com/
8 https://us.zebra-med.com/
9 www.me-med.com/html5/?_id=11282&did=2466&G=11049&SM=112
10 www.heartflow.com/

to create a 3D model of the patient's coronary vessels. Then, algorithms are used to create and analyze an exact 3D model of the blood flow, enabling a calculation of how significantly this flow is restricted. The results can be used to determine the subsequent steps in a treatment plan. According to the company, the system is based on the findings of more than 200 scientific studies on the issue of coronary heart disease, as well as decades of research and development.

Therapeutic procedures

One example of a therapeutic product already on the market is the Sugar.IQ app, produced by Medtronic and IBM, which acts as a **personal assistant for patients with diabetes**.[11] By means of a small subcutaneous sensor, patients' blood-sugar levels are sent via Bluetooth to their smartphones every five minutes. The system triggers an alarm as soon as the value rises above or falls below the individually determined maximum or minimum threshold. In addition, the algorithm identifies and analyzes specific patterns of blood-sugar level changes, for instance during certain physical activities. This enables improved control of diabetes symptoms, as well as targeted adjustments to prescribed treatments. The system is based on IBM's Watson artificial-intelligence technology.

11 www.medtronicdiabetes.com/products/%20sugar.iq-diabetes-assistant

5 Ethics of algorithms in healthcare

Algorithms in themselves are neither ethically good nor problematic. Rather, any assessment of their moral or ethical value [12]must take place in the context of their application, and must account for their functions across different social contexts (Mittelstadt et al. 2016; Wagner et al. 2017: 12). In addition, the degree to which relevant principles and values are either observed or breached in the course of their development and application must be clarified. In healthcare, these principles include above all the orientation toward patient well-being; the preservation and promotion of the capacity to make decisions and act on them; (informational) self-determination; protection from harm; equality of treatment and protection from discrimination; and the fair use of limited resources. In the course of developing and using algorithms in the healthcare field, various opportunities and challenges arise at the individual, institutional and societal levels.

5.1 Responsibility at the individual, institutional and societal levels

Responsibility at the individual level

With regard to the individual perspective, one key set of questions runs as follows: What are individuals allowed to do on ethical grounds (ethical permission), what should or must they do (ethical imperative), and what is neither required nor allowed (ethical prohibition)? However, the question of what rights and obligations others (individuals, institutions and society as a whole) have toward the individual must also be examined. These questions are particularly relevant when it comes to dealing directly with conflicting interests and values. At the individual level, for example, patients may see health benefits deriving from improved diagnostic abilities. However, they may at the same time see themselves as being forced to disclose a significant portion of their private data. The value of well-being from a health perspective may thus come into conflict with the value of personal privacy and individual control over how personal data is used.

Physicians and other healthcare actors, as well as medical researchers, are required in such cases to weigh conflicting interests conscientiously, and to safeguard the rights of affected individuals. They bear a responsibility for observing ethical principles and values both due to their role or function (as physician, healthcare employee, etc.), *and* to their position as an individual.

12 In academic discussions, the concept of ethics is commonly distinguished from the concept of morality. Ethics is usually understood as the critical examination of the phenomena of morality, here used in the sense of moral values and conventions.

Ethically relevant levels

Individual level
The individual actor – the individual affected person

Institutional level
Design and organization of framework conditions

Societal level
Political and legal decisions,
social and cultural impact

Responsibility at the institutional level

Like individuals, institutions too are actors in the development and use of algorithms. State institutions such as data-protection oversight agencies have to ensure that sensitive information is protected against unauthorized access. They therefore have different rights than do private persons, for example.

Institutional actors within the healthcare sector, such as insurers, physician associations or professional chambers, have the responsibility to create conditions under which healthcare-relevant data is properly processed and used for the purposes of immediate and preventative healthcare. In doing so, they often face ethical challenges that an individual actor would be unable to overcome. Thus, appropriate institutional frameworks are necessary to enable individual actors within the healthcare system to deal commensurately with ethical challenges.

For example, it is foreseeable that the increased use of algorithms will over time lead to fundamental changes in the job descriptions and activity profiles of medical practitioners (Amarasingham et al. 2016). At any point at which algorithms achieve better results than humans, tasks could accordingly be delegated to such systems in the future. For example, if an algorithm achieves better results than human experts in analyzing images for the purposes of early skin-cancer detection, it would seem to make little sense to continue educating, employing and training staff in the same way as is done today. In this regard, education and training institutions also bear a measure of responsibility. As they (further) develop their curricula, they must in the future teach the sector's skilled employees how to use algorithm-based systems, and how to interpret and review the automatically generated results (Wang et al. 2016).

The demarcation between individual and institutional responsibility may seem artificial at first glance, as it is always people who are acting in the end. However, the distinction is relevant. Institutional framework conditions and guidelines structure and guide people's actions in many relevant respects (Göbel 2017: 49). For example, an institution's guidelines might explicitly refer to ethical principles, and structures can be constituted so as to promote corresponding actions on the part of employees. However, such frameworks can also render appropriate action more difficult due to problematic or contradictory instructions.

Institutional structures can be designed to respond well to ethical challenges, for instance because there is a high degree of mutual trust and competency. By contrast, they can also be constituted so that individuals have little hope of support from within the institutional structures.

This distinction becomes relevant, for example, with regard to issues of responsibility. In the case of a flawed outcome, for example, is the medical professional that used the algorithm responsible for the harm caused? Or is it the department or facility that originally decided to use the algorithm? Are the programmers that developed and trained the algorithm responsible? These questions, which may also have legal relevance, cannot be answered at the individual level. Rather, it is always necessary here to examine the institutional structures as well. These should be constituted so as to respond to these and similar challenges in a reasonable way, making an ethically acceptable solution achievable.

Core ethical principles and values

- Informational self-determination
- Capacity to make and act on decisions (freedom)
- Protection from harm
- Protection against discrimination
- Fair treatment when interests conflict
- Fair handling of scarce resources
- Beneficence

Responsibility at the societal level

Finally, individuals and institutions are a part of society. Thus, the issues identified here must also be considered in the context of societal challenges. For example, algorithms are used to further several public-health goals, including that of generally improving the provision of healthcare, and of detecting epidemics as early as possible (Wilder et al. 2018). To this end, they are increasingly able to connect and analyze data from entirely different areas of life.

Such analyses could produce various advantages and disadvantages for different societal groups. The detection of specific data correlations, for example, can lead to discrimination against individual groups. This should be a concern, for example, if algorithms are used to examine the influence of lifestyle on the development of specific diseases, with corresponding benefits or penalties for insurance premiums. People that lead a lifestyle associated with a higher risk of illness could even be identified by the algorithm and excluded from certain medical services (Lippert-Rasmussen 2016).

Here, it is critical to be aware of that norms and standards have already been incorporated into the algorithm's programming with the choice of what data is to be processed, and with the design of the rules governing that process (Kraemer et al. 2011; Mittelstadt et al. 2016).

As noted above, algorithms are trained to process very specific types of data. In addition, a set of baseline data is employed that is used both as an initial point of departure and as a reference norm. This reference may itself already contain some bias, that ultimately determines the overall performance of the algorithm.

The broader public has seen an example of this phenomenon in the facial-recognition failure experienced by a Google photo app. The algorithm used in this application was trained using image data that primarily included photos of people with light-colored skin. Due to this limited dataset, the program was not sufficiently trained to recognize individuals with darker skin colors as people. The automatic tagging function instead referred to them as gorillas (Kasperkevic 2015). A false classification of this kind represents an example of discrimination that is in no way ethically acceptable. If it were to occur in a medical application, it would wrong people and, in addition, be hazardous to their health.

When it comes to problems touching on mutual respect and safety, individuals and institutions are often unable to provide a solution. Societal discourse and political solutions, for instance in the form of legal regulations, are needed here. Above all, a sociopolitical consensus must be developed regarding the objectives and purposes that should guide the development and use of these new technologies. For example, should algorithms be used with the primary goal of reducing healthcare costs? Should it be permitted that algorithms developed for healthcare purposes are also used for commercial ends? These and other questions are of a sociopolitical nature, and demand corresponding discussions and solutions.

Other questions arise when looking at the technology's possible impact on future sociocultural developments. Given the new technical possibilities for helping people live healthier lives, will new health-related obligations emerge as well? For example, could an obligation arise to collect one's own individual vital data, thus facilitating the early detection of potential disease risks and reducing treatment costs? These questions too require a broad public discussion aimed at creating awareness of possible developments and finding common answers to current and future challenges. Such answers must satisfy ethical claims by observing the core principles cited above – that is, they must promote the capacity to make and act on decisions, support (informational) self-determination, protect against potential harm and discrimination, and help allocate scarce resources in a fair manner.

5.2 Opportunities and challenges

As has been made clear in the previous sections, diverse hopes and expectations are bound up with the use of algorithms in healthcare. Some additional specific applications will be described and discussed in chapter 6. Numerous current reports on projects engaged in the development and use of algorithms in healthcare give the impression that the realization of preventive, predictive, diagnostic, prognostic, therapeutic and rehabilitative goals are all imminent. Such predictions should be taken with a healthy dose of skepticism.

Many ideas and projects still have a long road ahead of them before being realized, and before they can be practically used in health research and care provision on a quality-assured basis. As in other domains, it is highly unlikely that all projects will succeed, or that all expectations will be met. The following presentation of the opportunities and challenges associated with algorithms in medicine and healthcare must in this respect be understood as a reflection of the field's expectations, wishes and hopes. It also highlights the challenges that could be associated with the various applications. Our goal is thus neither to

Opportunities and challenges – at a glance

Opportunities	Challenges
Improved early detection of diseases	Cannot substitute for human judgment Lack of differentiation between correlation and causation
Faster and more accurate diagnoses	Lack of control if processes proceed automatically (black-box effect) Safety risks due to complexity and lack of transparency
Improved safety standards	Complications with regard to allocating responsibility
Therapies tailored to individual patients	Promotion of automatism & threat to right to self-determination
Increased efficiency and cost-effectiveness and reduction of burden on med. staffers	Bias risks related to how threshold levels are set Bias risks from insufficiently large underlying dataset
Less susceptible to error than human actors (increased patient safety)	Re-identification produces threat to informational self-determination rights Risk of data theft and data misuse
Discovery of correlations in massive quantities of data for the purposes of generating hypotheses, with ultimate goal of identifying causalities	New professional demands on healthcare actors

deliver a forecast for the future nor to assess the beliefs and assumptions associated with the opportunities and challenges we identify.

Opportunities

The use of algorithms in healthcare is bound up with fertile expectations, and in some cases very great expectations indeed. These include: a significant increase in the speed with which health-relevant findings are obtained by researchers and introduced into practical care; a significant expansion of the existing knowledge base and the range of medical services that depend on it; and an increase in the accuracy of diagnoses and treatment recommendations, along with associated increases in the safety of medical procedures (Dörn 2018: 352; Wired 2017; De Witte 2017). With the automatic processing of a variety of health-related personal data comes the hope of developing individualized medical treatments – a so-called precision medicine – while simultaneously reducing healthcare-system costs (IBC 2017: 7; De Witte 2017).

Big-data algorithms

The expectations cited for digitalized medical research and healthcare are primarily associated with the potential of using algorithms to process large amounts of data from many different sources in a short period of time. These sources include patient records, medical research publications, insurance records, real-time vital-sign data from wearable devices and other sensors, and other diverse data generated through the use of digital services such as online surveys and social media. All this data can be linked together, evaluated by algorithms seeking specific information, and thus rendered useful for healthcare purposes (Deutscher Ethikrat 2018: 63).

However, the mere availability of a significant amount of data does not in any way guarantee that it will be evaluated in a meaningful or useful manner. With respect to big data, experts say that current practice often fails to observe customary scientific-research principles, and violates the principles of evidence-based medicine (Antes 2016). Critics particularly charge that too little attention is given to the development of a theoretical framework to guide the evaluation of the data (Mayer-Schönberger et al. 2013: 70).

To enable meaningful analysis of the data being made available by advancing digitalization in various medical fields, this data must be edited and curated. This task can only be performed by human experts. However, algorithms can provide valuable support. Algorithms can be used to facilitate the data analysis; for example, they may be trained and used with a high degree of accuracy to process only the data that is necessary to realize a particular objective, such as producing a prognosis for a complex disease.

The use of algorithms thus promises to simplify the handling of ever greater and more diverse quantities of data, generated in very different medical and extra-medical contexts.

Swifter, more comprehensive data matching

Improvements are expected particularly from the ability of algorithm-supported systems to automatically compare vast amounts of data in very short periods of time. The capability of machines here significantly exceeds the corresponding capabilities of human actors. On the basis of this kind of data-matching process, algorithms at times achieve the same or even higher accuracy levels than human experts in tasks such as the diagnosis of specific diseases. Especially with regard to rare diseases, they prove superior to humans in rendering diagnoses (Esteva et al. 2017; Rajpurkar et al. 2017). A variety of techniques such as automatic text and image analysis are used in this process. For example, algorithm-based image-analysis procedures enable a swift, automated check to be made for potential skin diseases. Algorithms can generate diagnosis suggestions on the basis of a photo of the affected skin area and a digital questionnaire designed to capture any additional symptoms (Dörn 2018: 354).

Algorithms are also already being used to automatically identify drug interactions and side effects, on the basis of information from digital patient records and the medical research literature (Dörn 2018: 651). Additional algorithm-assisted diagnostic and therapeutic systems promising faster diagnoses and more personalized therapies are currently being developed. The number of inadequate or unnecessary treatments could also be reduced as a consequence of improved diagnostic results.

Consistent performance level and avoidance of errors

Increasingly powerful processors promise significantly accelerated data-processing speeds and a more consistent level of performance. While human performance can be impaired by fatigue, for example, and is more broadly subject to general fluctuations, computing power functions consistently and reliably, regardless of the time of day or how long the system has been in service. The only requirement in this regard is a secure energy supply (Dörn 2018: 352).

The use of algorithms can correct any possible errors made by overtaxed employees. In this respect, algorithms contribute to increased safety in the healthcare system. In addition, they can generally serve to reduce workloads in the medical and nursing-care fields. For example, algorithms can be used to monitor the vital signs of patients or persons otherwise in need of care. Aberrations in vital signs as captured by sensors in real time are automatically reported to medical or caregiving staff. In addition, intelligent algorithms can automatically detect dangerous situations and trigger an emergency call. Ideally, the time gained as a result of such technological support can be used for treating patients (Dörn 2018: 355), provided staff levels are not reduced.

Assumption of routine tasks

The use of algorithms opens up new opportunities for automated processes in other areas too. Many routine tasks, for instance in laboratory medicine, cardiology and radiology, could in the future be taken over by algorithms (Rasche 2017: 8). This may relieve the burden on skilled workers, thus providing them with time for other activities, but also offers the possibility of reduced healthcare costs due to savings on staff salaries. The accompanying threat to job security illustrates the ambivalence that can be associated with digitalization. The opportunity to reduce healthcare costs with new forms of automation must be considered in conjunction with the associated individual and labor-market policy challenges.

Individualized and increasingly preventative medical care

In the long term, the comprehensive analysis of health-relevant data could make personalized medical care possible (IBC 2017: 7). By combining a broad range of personal health-related data such as genetic information, information on lifestyle habits and preexisting conditions, therapies could be tailored to the individual person with great precision. The hope is that thanks to this individualization, patient care in the future will not only be better, but also more cost-effective, given that side effects would be reduced and the chances of recovery would be maximized.

In addition, advocates hope that today’s primarily reactive medicine can increasingly be replaced by a preventive medicine (IBC 2017: 8). To make this possible, data from a variety of sources, such as wearable devices and other sensors, would be combined and analyzed with reference to risk factors. Given the risk profile thus determined, a person’s lifestyle could be preventively adjusted. If health-relevant values were to fall above or below acceptable threshold levels, either in the environment or the person’s body, an automatic signal could be generated and sent. In such a case, the affected individual would be warned of the presence of a health risk before it became a genuine danger to health.

Accelerated transfer from research to practical care – and back again

In addition to positive expectations for medical practice, significant new opportunities are forecast within the medical-research field. The greater availability of data and the use of algorithms promises to speed the pace at which data flows between research and practice, and to make this exchange more comprehensive. The faster flow of information could take place in both directions – from research institutions into the field, and from practical care settings back into the research sector. The research process could be given a broader data foundation through the automatic analysis of data from the field; similarly, the transfer of new research results into practical medical care could be significantly accelerated. In imagining this sustainably accelerated exchange between the research and practice sectors, experts refer to the emergence of a *learning healthcare system* (Krumholz 2014).

Challenges

Potential misjudgments of algorithms' capacities

Given their high speed and ability to process vast amounts of data, the capacities of algorithmic systems could easily be overestimated. With regard to the storage and management of data, machine-based systems are systematically superior to humans; moreover, the degree of this superiority will presumably continue to increase in the future. However, when it comes to evaluating information, they are systematically inferior to humans.

Human judgement is required in many if not most areas of medical care, nursing care and research, for instance when it becomes necessary to balance various risks and opportunities. As stressed by experts such as Rasche (2017), if there are multiple diagnostic or therapeutic options, each associated with their own advantages and disadvantages, then an algorithm can have at most a supporting function. They cannot replace human judgement. Thus, when considering algorithm-generated recommendations, it is important to make a clear distinction between *recommendations* and *decisions*. Digital assistance systems may generate recommendations – advice that depending on the quality and quantity of the processed data may well have the "character of a qualified second opinion" (Rasche 2017: 8). However, they cannot make a decision. This task must always lie with a human being (ibid.).

Allocating and distributing responsibility

This also applies to the use of algorithms in systems that automatically administer drugs, trigger electrical impulses, or send messages to medical or nursing personnel. An example here are sensors implanted under the skin that monitor a diabetic person's blood values and automatically release insulin as necessary.[13] In applications of this kind, the algorithm to a certain extent seems to be independently "deciding" whether, when and how much insulin is to be administered. Yet the decision is based on and determined by the programming and the function settings. In other words, humans are making the decisions – in this case the programmers, patients, physicians, nursing staff and perhaps even other people involved. However, the system's programming, use and choice of settings could have ethically and legally problematic implications, particularly with view to the allocation of responsibility.

13 www.medtronicdiabetes.com/products/minimed-670g-insulin-pump-system

Responsibility for errors and the problem of insufficient transparency

An algorithm can obviously cause harm if damage is produced by qualitatively poor or outright faulty programming or use. However, it would be strange to suggest that the algorithm itself is responsible for the harm in the literal sense. Even very sophisticated algorithms are unable to take on responsibility. They are unable to engage in morally responsible decision-taking. Only humans are able to do this. Hence, if harm occurs due to the use of an algorithm, those who were involved in the technology's programming and application decisions are responsible. However, given the often large number of people participating in such decisions, there is some question as to who specifically is responsible for which factors and errors (Mittelstadt et al. 2016). Is it the programmer, the institution that is offering the system, the treating physician or perhaps even the patient? As yet, this question remains unresolved. Answering it is an increasingly urgent task given the growing use of algorithms.

The problem of allocating responsibility is complicated still further by technical aspects of the issue. This is particular true with regard to the transparency of data-processing rules, and the associated practical opportunities afforded to clinics and other users to assume responsibility (Mittelstadt et al. 2016). In order to be able to make decisions, individuals must first have access to sufficient relevant information and practical knowledge relating to the subject at hand. However, even for computer scientists, the ways in which algorithms function are at times difficult and at times even impossible to comprehend in full (European Group on Ethics in Science and New Technologies 2018). In the event of an error, it may be that due to the so-called black-box effect, neither the user nor the computer scientist can detect its origin precisely. This renders the analysis of possible failures significantly more difficult, or perhaps even impossible, and risks a loss of control (Jaume-Palasí and Spielkamp 2017).

In this regard, semi-supervised or unsupervised machine learning are prone to raise problems. The individual steps of such processes are often not fully comprehensible even to computer scientists and programmers. Should the algorithm function incorrectly, humans would thus be unable to identify the element that is causing the behavior. Even in the case of supervised-learning algorithms, questions arise regarding transparency and the allocation of responsibility, for instance between individual programmers and users. Algorithms of this kind are used to filter and process information. In so doing, they influence human decisions. A data-processing failure can as a consequence lead to erroneous human decisions, for instance if information relevant to the decision is classified as irrelevant. If experts rely on the capabilities of an algorithm of this kind, factors relevant to the decision might easily be overlooked. In the worst case, even the awareness of the possibility that decision-relevant information could be overlooked by an algorithm is lost (Mittelstadt et al. 2016).

Consent to automated processes

It must be borne in mind that the use of automated systems is bound by the informational and consent requirements that hold more generally with regard to medical procedures (Sect. 630e BGB). For example, a system using sensors to monitor a patient's vital signs, with an algorithmic component that automatically detects values over or under specific medical threshold values and accordingly sends a message or even triggers delivery of a drug, would require the patient's consent. Otherwise, it would violate the patient's personal rights and even (in the case of drug administration) bodily integrity.

Automatic generation of spurious correlations

In addition to the challenges noted above, limitations to algorithms' capacities due to the current state of digital development must also be considered. While these constraints may well be overcome in the course of future technological development, expectations regarding currently available systems should not be exaggerated.

As noted above, many of the current challenges stem from the fact that digital systems cannot distinguish between correlations and causalities (Fasel and Meyer 2016: 9; Wagner et al. 2017). For example, an algorithm programmed to compare data respectively on alcohol consumption and cardiovascular diseases may indicate a relationship between the two variables. However, it does not mean that the relationship is one of causality. In fact, it may be no more than a spurious correlation. In this regard, exaggerate reliance on automatic analysis of large datasets can lead to wrong conclusions and systematic fallacies.

While some data set seems to indicate that people who, for instance, regularly drink high-priced red wine have a reduced risk for the development of cardiovascular diseases. The presumption that the lower incidence of cardiovascular diseases in consumers is caused by such wine could be false. The two factors are correlated with one another – but a causal relationship is in no way certain. In this case, for example, it has been shown that people who can afford to purchase expensive wine are exposed to lesser amounts of social stress due to their income-derived status, and that this may be a causal factor for a decreased risk in terms of cardiovascular diseases (Mortensen et al. 2001).

Human researchers can recognize such relationships due to their capacity to create hypotheses and their grounding in other areas of knowledge. An algorithm cannot do so. In short, a human expert makes distinctions on the basis of hypotheses and theories that are regularly adapted to the demands made by the course of scientific progress. Algorithms are currently limited in this respect (Antes 2016).

Distortion of results due to bias-related phenomena

Additional challenges are associated with the so-called bias phenomenon. In this context, bias refers to a condition in which an algorithm's data-processing rules lead to systematic distortion or one-sidedness. For example, algorithms are used to automatically analyze and classify cell samples with regard to specific disease markers (Kraemer et al. 2011). In many cases, such a classification will be unambiguous. In other cases, however, the assignment can be uncertain. For these instances, a threshold must be set that determines whether a cell sample will be labeled as relevant to the disease or not. Setting a norm or threshold of this kind requires to decide which outcome is preferable: that a system potentially produces false positive alarms, or falsely negatively labeled samples (ibid.).

A bias may also be created by the fact that an algorithm is operating with inadequate underlying data. This may either be because the algorithm – as in the above-cited example of Google's image-recognition algorithm – was trained with insufficient or especially one-sided datasets. Alternatively, it may also be due to the incompleteness or contradictory nature of the datasets being used as the machine-learning algorithm is applied. At this point in time, health-relevant data has often been collected only incompletely. Data in patient records, for example, is often coded poorly or inconsistently, and the information itself incomplete.

Such shortcomings have an impact on the performance of algorithms, as they are unable to evaluate such data, or can do so only in a flawed manner. The fact that there may be a large amount of data available relating to certain groups of people, and comparatively little for others, may produce additional imbalances in the underlying data. Patients in hospitals that already work digitally, for example, produce more data than those in less fully digitalized facilities. An asymmetry of this kind can also lead to bias (De Laat 2017).

Bias-related errors can significantly impair the reliability of systems for practical use. The analyses generated are necessarily either incomplete or even incorrect. The opportunity noted above – that the use of algorithms will significantly improve the safety and reliability of healthcare services – is thus valid to only a limited extent today. It remains to be seen whether the problems caused by various types of bias can be overcome in the foreseeable future.

Inadequate data-collection and data-exchange standards

Digital content can ultimately be processed in an automated or semi-automated way only if a sufficient degree of interoperability between different systems is in place. It is hence important to develop and establish common data-exchange standards (Hahn and Schreiber 2018: 340). For the time being, this area shows some significant shortcomings. This is true at both the technical and the user level, for example with regard to the collection of data and the degree to which it is coded appropriately in patient records. On the technical level, fast and reliable data connections ("telematics infrastructures") must be established nationwide; without these, algorithm-supported systems will not be able to function properly. At the user level, the use (and ongoing improvement) of such systems is dependent on data being correctly entered and retrieved (Amarasingham et al. 2016).

These systems in turn create new individual and professional challenges. Healthcare professionals may have to acquire new technical knowledge and agree on common standards, for example for doctor's referrals and laboratory data.[14] Engineers and programmers face the challenge of designing systems in a user-friendly way (Rüping 2015). To do so, they need information on how the users typically use the systems, what challenges they experience, and what mistakes they typically make. Only once this knowledge is acquired can algorithm-supported systems be designed to be used securely and in a manner fulfilling their intended purpose.

In practice, this constitutes a further challenge: specifically, initiating a cooperative venture or dialogue of this kind, and then if necessary creating uniform standards for this as well. However, this cooperative effort is already being bolstered by a variety of support mechanisms. For example, in the context of a medical informatics initiative, the German Federal Ministry of Education and Research is supporting diverse interdisciplinary consortia working on projects of this nature.[15]

14 Thomas Kriedel, a board member of the National Association of Statutory Health Insurance Physicians (KBV), addressed both of these issues in a May 2018 meeting of the organization's representatives: www.aerzteblatt.de/nachrichten/94956/Einfuehrung-der-Tele%C2%ADma%C2%ADtik%C2%ADinfra%C2%ADstruk%C2%ADtur-hakt-weiterhin

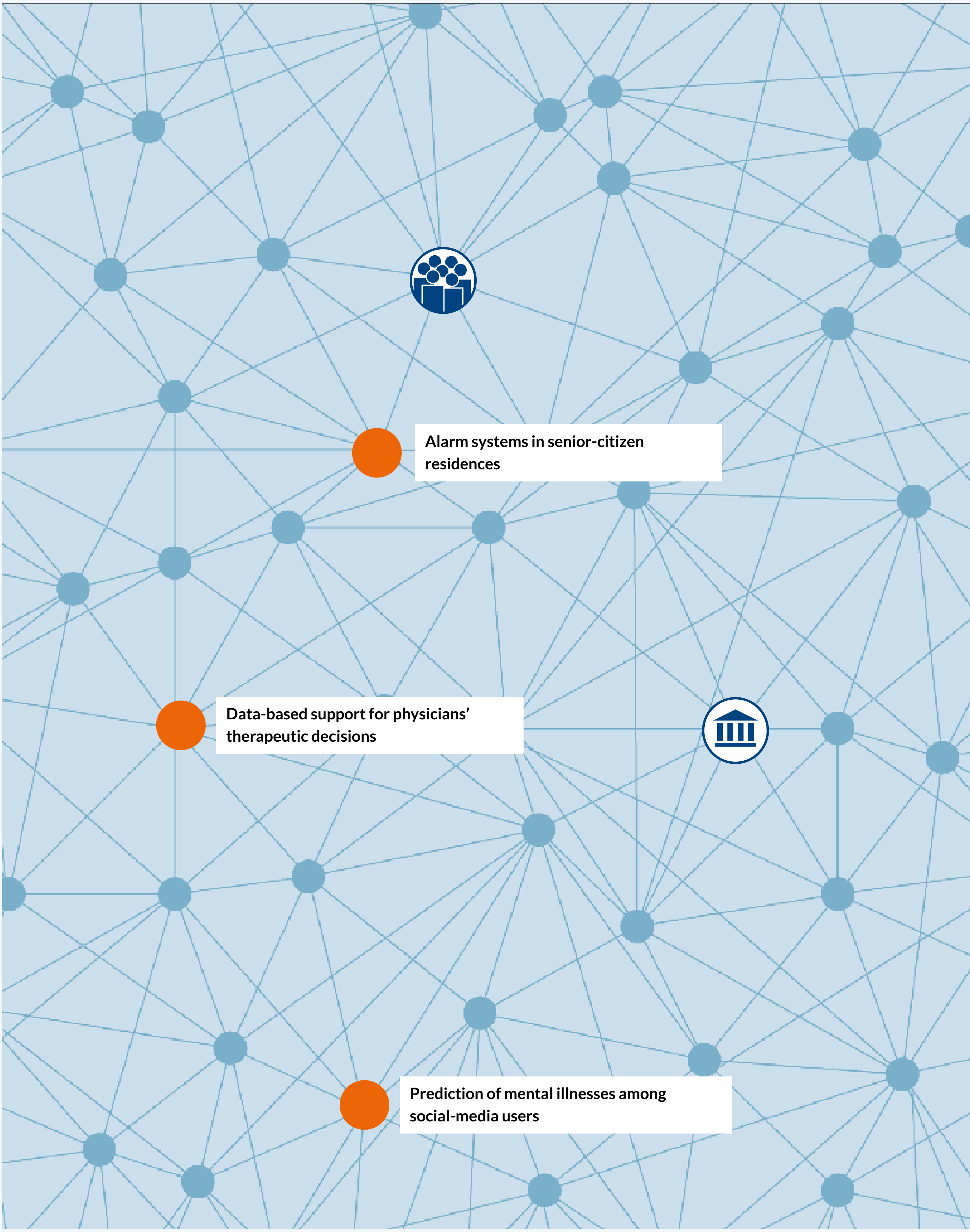
15 www.medizininformatik-initiative.de/

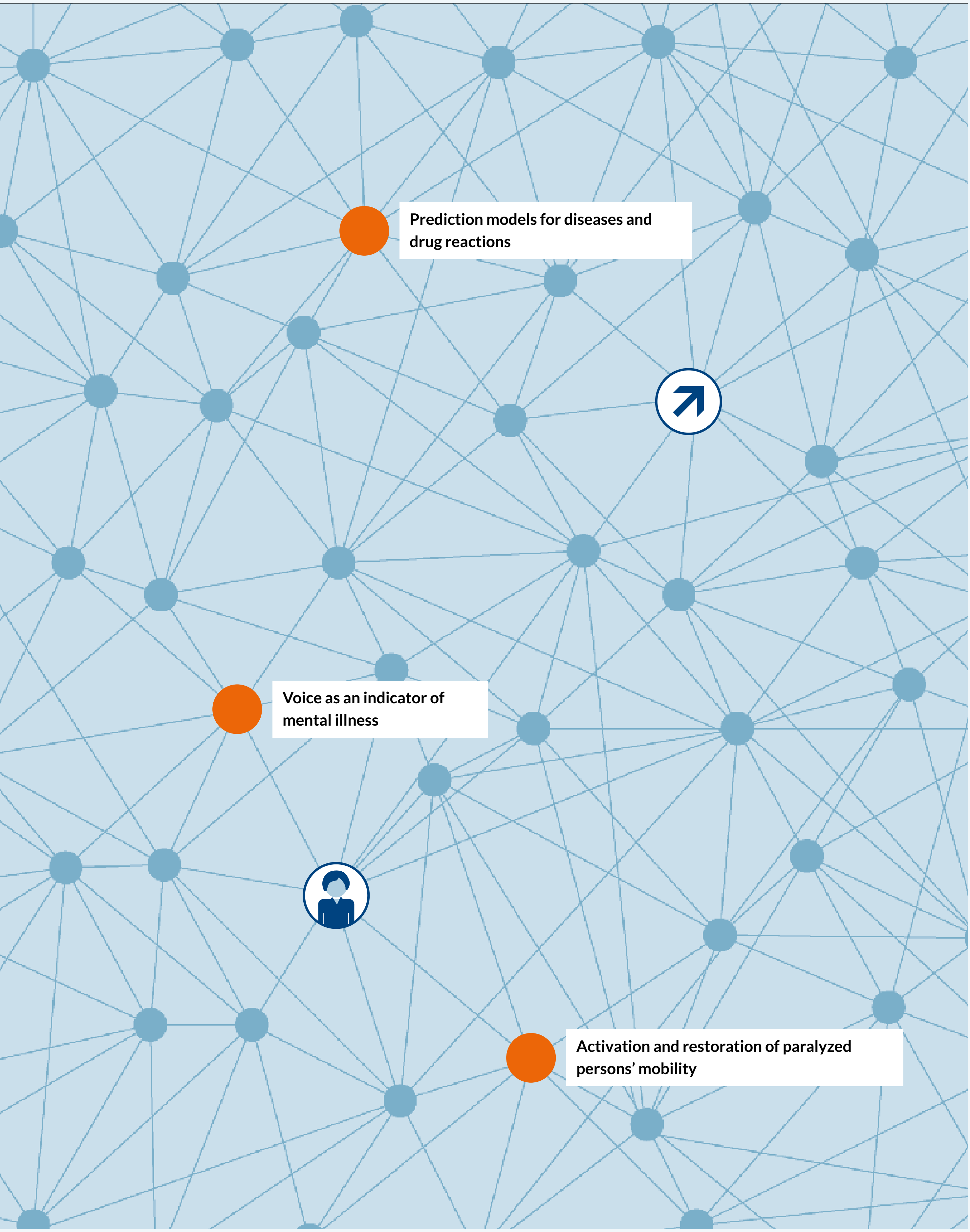
The problem of data anonymity and security

Further issues occur in the course of the exchange and processing of sensitive health related data (Abouelmehdi et al. 2017). For example, when recording data, treating physicians or nursing staff members may fail to exchange data exclusively in a properly encrypted way. In addition, there may be uncertainties regarding the circumstances in which medical institutions are allowed to pass on healthcare data to programmers or medical-sector computer professionals. Powles and Hodson have analyzed a case in which managers of a British hospital passed on a significant quantity of patient data to DeepMind, a company that used this data for the development of a clinical analytics tool (Powles and Hodson 2017). The research was neither aimed at any individual diagnostic nor did it provide any therapeutic benefit for the affected data subjects. Moreover, the patients' privacy rights were infringed (ibid.).

With regard to the sharing and processing of personal data for research purposes, algorithms can generate an additional problem: The anonymization of personal data is becoming difficult. It is becoming increasingly feasible to re-identify a person even in the case of data that has supposedly been completely and comprehensively anonymized. Many experts already assume that anonymization can no longer be guaranteed (Mittelstadt and Floridi 2016).

Personal data can also be used for purposes that do not benefit the data subject. This is obviously true in the case of data misuse that directly harms the individual in question, for example because the knowledge obtained about the person is used against him or her, perhaps through the rejection of insurance coverage. While maybe less obvious, it is just as relevant if personal data is used for legal purposes without appropriate authorization by the individuals involved. This too would constitute misuse.


Alarm systems in senior-citizen residences
Data-based support for physicians' therapeutic decisions
Prediction of mental illnesses among social-media users


Prediction models for diseases and drug reactions
Voice as an indicator of mental illness
Activation and restoration of paralyzed persons' mobility

6 Application scenarios

In the following sections, we will again focus on opportunities and challenges, this time in the context of concrete application scenarios. The selected examples are derived from current research and practice. They include such varied applications as the use of algorithms to analyze social-media image content in order to predict depressive disorders, the use of algorithms in clinical decision-support systems, and the use of algorithms in monitoring systems designed to allow elderly people and other individuals needing care to live as long as possible in their accustomed environments. Each field of application produces specific opportunities and poses its own challenges.

6.1 Prediction of mental illnesses among social-media users

Social-media sites such as Facebook or Instagram are used today by hundreds of millions of people worldwide. A huge amount of new content in the form of photos and comments is produced on these platforms on a daily basis (see Statista 2018). This content is also potentially of interest for medical research purposes. For example, studies have shown that depressive people prefer dark- and gray-toned colors in photographs, and often engage in only limited social contact (Carruthers et al. 2010; Bruce and Hoff 1994). Based on these observations, Reece and Danforth (2017) proposed a method by which depressive episodes could be recognized early on the basis of an automated analysis of photographs published on Instagram.

The authors' objective was to identify specific markers of depression in posted photos with the help of an algorithm. To this end, the researchers investigated the hypothesis that the uploaded photos and other metadata of Instagram users with depression could be reliably distinguished from the contributions made by healthy users. In addition, they also hypothesized that Instagram postings by users with depression uploaded even before their first clinical diagnosis could be reliably distinguished from those uploaded by healthy users. This hypothesis is particularly interesting, because the behavior of a user who is aware of his or her condition could be influenced by this knowledge. For example, awareness of this kind could even have an impact on the person's self-presentation on Instagram.

In order to verify their hypotheses, Reece and Danforth (2017) used machine learning techniques in conjunction with image-processing tools, extracting certain characteristics from the images. In selecting markers for their analysis, the researchers focused on image characteristics that had proved to be reliable predictors of depression in previous studies.

Social media

Early detection Medical research

Diagnosis of depression Usage behavior

Machine learning

These characteristics included, among other features, the presence and number of people in a given photo. These qualities were analyzed with the help of a facial-recognition algorithm, and served as an indicator for the participating study subject's level of social activity. The authors were also interested in the location that the photo had been taken in (e.g., inside or outside). Additional pixel-level qualities were also recorded, such as average color levels and brightness. In addition, the study captured characteristic of social-media sites, such as whether other users had commented on a photo, and how many "likes" it had received. The authors also examined elements of each subject's usage behavior, such as the frequency with which they visited and used Instagram.

The underlying data for the model incorporated 43,950 photos posted by 166 Instagram users, 71 of whom had been diagnosed as depressive. The recruitment took place through Amazon's Mechanical Turk online marketplace. The Center for Epidemiologic Studies Depression Scale (CES-D) was used to identify the extent of potential participants' depression. Healthy participants were further examined in order to ensure they showed no signs of depression. Suitable participants were subsequently asked about their use of Instagram up to that date, and were requested to log into their Instagram accounts using an app embedded in the survey so that they could share their data with the researchers.

The results of the study confirmed both hypotheses cited above. With regard to the image characteristics, for example, it was shown that the photos posted on Instagram by depressive patients were more likely to have blue, grey or dark tones, and clearly received fewer "likes". Depressive users additionally tended to filter out all colors from their photos, while at the same time showing an aversion to artificially brightening the images. Moreover, the photos uploaded by depressive users rarely depicted multiple people.

In order to test the accuracy of the predictive model, the researchers compared these results with the data contained in a comprehensive meta-analysis (Mitchell et al. 2009). In this previous study, the authors had reviewed 118 studies that analyzed the degree to which general practitioners were able to correctly diagnose depressions in their patients without the help of surveys or other measurement instruments.

Overall, Reece and Danforth's model (2017) showed notably better results with regard to the diagnosis of depression than did general practitioners carrying out in-practice diagnoses without the help of measurement instruments. The algorithm correctly detected the presence of depression in a majority of the patients solely on the basis of the photos uploaded to Instagram, while more than half of the general practitioners' diagnoses were false positives. This means that the physicians falsely diagnosed a considerable number of healthy patients as depressive. According to the researchers, the algorithm's predictive power could significantly be improved if the features of the text posted by users were also analyzed. The authors note that text analyses of this kind in previous studies have already proved successful in recognizing various health-related and psychological signals in social networks.

Opportunities

Mental illnesses are considered to be the fifth-largest factor in the "Global Burden of Disease" (Whiteford et al. 2013). Promoting mental health is accordingly one of the central objectives of healthcare. This is also emphasized by the World Health Organization (WHO) in its current Mental Health Action Plan 2013–2020. Among other goals, this document calls for improvements in the monitoring of population health, in order to identify people at risk of developing diseases as early as possible, and help them to the greatest extent possible (WHO 2013).

Algorithms can contribute to such a monitoring process, and thus also to a reduction in the global disease burden. In this case, the algorithm described falls into the "predictive analytics" domain. Tools of this kind are defined by their ability to make real-time predictions about probable future events (Cohen et al. 2014). The results of recent studies, such as that by Reece and Danforth (2017), indicate that predictive analytics could also be used to identify signs of depressive disorders through the automated monitoring of social-media sites. The identification hereby takes place extremely quickly – as soon as any content is posted – and, as Reece and Danforth note, is also typically more reliable than, for example, general practitioners who make a first diagnosis without aids such as questionnaires. At least in comparision to this approach, the algorithms generate significantly fewer false-positive diagnoses. It is not clear from the project description whether fewer false-negative diagnoses were also rendered. Thus, it could be that the algorithm produces less convincing results in this respect. However, both features – real-time predictions and the avoidance of false-positive diagnoses – open up opportunities for improved care for people with mental illnesses.

Avoidable late diagnoses or misdiagnosed mental illnesses are not only a burden on the individuals, but also place extra strain on the healthcare system and the insurance mechanism's shared-risk pool. According to Germany's National Association of Statutory Health Insurance Funds (GKV-Spitzenverband), expenditures by the country's statutory health-insurance organizations for the treatment of mental illnesses totaled around €9.7 billion in 2016 (Wissenschaftliche Dienste Deutscher Bundestag 2017). Undiagnosed, depression can worsen to a point at which the sufferer's life is in danger. On the other hand, false-positive diagnoses may cause superfluous examinations and lead to unnecessary therapies. These latter translate into unnecessary burdens for the individuals in question as well as avoidable costs for the healthcare system (Reece and Danforth 2017).

Avoidable costs

Quality of care Monitoring

Global application possibilities

"Predictive analytics"

In addition, existing offerings could be expanded up to and including large-scale screenings (ibid.). Previous analysis procedures, particularly those requiring technical support, have been comparatively expensive, and often reach only a small portion of the group at risk. The use of predictive algorithms applied to content that is already widely shared on social-media sites opens up fundamentally new perspectives in this regard. Social-media sites are increasingly used by people belonging to many different population groups. This circumstance makes it possible to reach people who cannot be reached using previous approaches to care. For example, this includes people who, due to a lack of knowledge or a feeling of shyness, do not make contact with institutions or other providers offering the needed medical or psycho-social services.

The use of algorithms could also promote the improvements to mental-health care called for by the WHO on a global level. In many regions of the world, mental-health services are severely underfunded or simply unavailable. However, social-media sites are frequently used even there. The use of predictive algorithms for the early detection of mental illnesses such as depression would therefore be conceivable there as well, at least in principle (ibid.).

The opportunities cited are still associated with numerous challenges that must be considered from the individual-ethics and institutional-ethics perspectives, as well as from the societal and sociopolitical perspectives. The distinction between these perspectives is not to be understood as extensional in the sense of connoting three different applied ethics (with separate subject areas). Rather, in accordance with the points noted in chapter 5, it references three perspectives of responsibility that are neither reducible to one another nor eliminable (Göbel 2017: 49; Gutmann and Quante 2017: 105).

Individual-ethics challenges

From the perspective of individual ethics, it is necessary to ask who or what group of people, under which circumstances, would be allowed to use or possibly even should use a predictive algorithm for the analysis of social-media content. The user that posted the content? Their contacts? Exclusively psychologically trained professionals and physicians? Other entities that have an interest in being informed about possible depressive disorders at an early stage, such as employers or insurance companies?

In addition, it is necessary to ask what rights or obligations arise from the use of such predictive algorithms. From an ethical point of view, might social-media users have not only a right, but also an obligation to allow their own content to be analyzed for possible mental-illness risks? After all, it could be argued, not only could mental suffering be diminished as a result of such early detection, but treatment costs too could be reduced. Or are social-media users alone entitled to decide whether their content is to be analyzed by predictive algorithms for mental-illness risks? In our society, this latter question would clearly and emphatically be answered in the positive, with reference to the right to informational self-determination (grounded in Germany in the Federal Data Protection Act (BDSG) and the EU General Data Protection Regulation (GDPR)).

This right refers to a core ethical principle within our society: the principle that every person has a right to privacy. The protection of personal privacy is a political cornerstone of free democratic societies, and is enshrined at various levels, often in the form of legal provisions. Examples include Article 8 of the European Convention on Human Rights, which according to the prevailing legal interpretation also incorporates the right to informational self-determination and the protection of personal data (European Convention on Human Rights 1950, current 2010 edition; EGMR 26.3.1987 – 9248/81).

Individual consent

Informational self-determination

Protection of personal privacy

Social pressure

Insufficient anonymization

The right to the protection of personal privacy also applies when dealing with health matters, perhaps even to a particularly great extent (Schaar 2016). The use of predictive algorithms for the automatic analysis of content posted on Instagram or similar platforms would raise a number of questions regarding the protection of personal privacy. In the context of studies such as that by Reece and Danforth, users are informed and provide consent for their image content to be retrieved and analyzed. However, if procedures of this kind were to find their way into general healthcare practice, such consent would not be in place.

The automated evaluation of social-media content with predictive intent thus represents a highly problematic violation of personal privacy and the right to informational self-determination (Guntuku et al. 2017). This is true regardless of whether the content in question (images, etc.) has been posted in public forums or not. As a general rule, social-media users do not expect their data to be surveyed and analyzed for medical purposes. The public they are reaching out to is constituted by other social-media users, not medical professionals. Particularly in sensitive areas such as mental health, the scientific analysis of publicly accessible data is thus ethically problematic. This holds regardless of whether the data has not been, or has been only insufficiently anonymized. However, this all also applies if the data has been collected "only" with the intention of training an algorithm (Conway and O'Connor 2016). Even participants in the Reece and Danforth (2017) study expressed some discomfort about sharing their Instagram data with the researchers. Several subjects who had initially wanted to participate in the study ultimately decided against releasing their data, declining to participate due to concerns about personal-privacy protections (ibid.).

If the predictive algorithm were to be put into practice, it would be necessary to obtain the consent of the individual users on the platforms involved. One possibility might be an additional settings option that enabled users to decide whether the algorithm would be used; this might also indicate whether the users themselves and/or other people should be informed if the algorithm found the typical signs of an emerging depressive disorder. A solution of this kind might seem plausible at first glance. However, it would in practice run a high risk of failing due to internet users' typical patterns of behavior. Most social-media users agree to useful-sounding offers all too quickly, without being fully aware of the consequences.

In addition to reasons of principle, the right of informational self-determination also serves to protect people from harms or wrongs caused by other parties using information relating to them. Predictive algorithms for the detection of mental health issues could be used by employers to obtain information about a person's capacity to cope with stress, for example (Guntuku et al. 2017). Already today, some employers visit publicly accessible social-media accounts in order to develop a more comprehensive picture of job applicants (Büttner 2016). If a job application were to be rejected from consideration due to an analysis of social-media accounts, despite the person's professional suitability, the applicant would have been harmed in terms of her interest in the job offer.

People with mental illnesses or an increased risk of illness also may face discrimination and stigmatization outside of professional contexts. The individuals in question could therefore suffer significant harm if the information obtained using the algorithm were to be disclosed. While this would certainly be true if the information was retrieved without the knowledge of or against the will of the affected individual, other situations could also be problematic.

Even a supposedly voluntary disclosure of such information should be viewed very critically (Bauer et al. 2017). Social-media sites are often used in the company of other people, and the posted content is often viewed alongside others at school or in a cafeteria, for example. This aspect is of interest when considering the question of whether and how the users themselves should be informed about the results of the algorithmic analysis. An automatic notice providing information about the mental state of an Instagram or other social-media user could easily lead to a situation in which other people also obtained information about that mental state. However, even preferences preventing the automatic display of such a notice could in practice fail to provide effective protection. It is anything but unlikely that – if confronted with social pressure – even users who know better might request such information or authorize its provision in the presence of third parties (ibid.).

Institutional-ethics challenges

One important criterion for the legitimacy of institutional action is the ethical legitimacy of the objectives. The widespread deployment of predictive algorithms could effectively support healthcare institutions in fulfilling their social mandate. If the early recognition of depressive disorders was rendered more effective, it could contribute to improvements in care for people with mental illnesses (Reece and Danforth 2017). However, this would only be true if the algorithms employed for this purpose were embedded within a comprehensive care concept. Taken on its own, the simple use of algorithms here would not suffice. Due to their social mandate and associated welfare obligations, institutional healthcare actors have a responsibility not only to shape the framework conditions governing access to new diagnostic opportunities, but also to make appropriate therapeutic offerings available.

Simply informing social-media users about their risk of illness, without also making appropriate therapeutic services available, would probably do more harm than good. The automatic provision of information about a potential depressive disorder, if not embedded in additional related informational materials and not delivered by a trained psychological professional, would be more burdensome than helpful for the affected individual. At worst, this could intensify existing mental problems by producing uncertainty, especially if the individual declined to seek professional help. Thanks to the algorithm's "external perception" or categorical grouping, the user's self-image could be influenced in such a way as to trigger self-reinforcing feedback loops (Cornford et al. 2007). The diagnosis itself might thus intensify the problems that are meant to be addressed.

Ethical legitimacy of goals
Duty of care
Self-reinforcing feedback loops
"Dual use"
Right not to know
Significant security risks

In addition, it should not be assumed that users in general want to be informed about a possible depressive disorder. Automatically generated information about a potential impending depressive episode is the equivalent of an unsolicited diagnosis. However, in health matters, as long as no other person is in danger (for example, through a risk of contagion), there is a right not to know. Everyone should be able to decide for themselves whether and how they want to be informed about health risks (Conway and O'Connor 2016). Users should thus be able to decide for themselves whether they are to be informed about their risk of a depressive disorder.

Healthcare institutions face the challenge of creating framework conditions and structures that enable the use of predictive algorithms to be managed and monitored in a professional, ethically proper manner. In addition, there is an institutional responsibility to ensure that no unauthorized persons gain access to the content and results of the algorithm-based analyses. Institutions must shape the conditions of access and use for predictive algorithms in the interest of the users. From an institutional-ethics perspective, it is therefore critical to clarify who should be provided with algorithm-supported access to information about the mental states of social-media users.

The possibilities opened up by predictive algorithms are also generating considerable interest beyond the medical sector. According to estimates cited by Bauer et al. (2017), global investment in the field of automated emotion detection and analysis will make this a $22.65 billion market by the year 2020 (MarketsandMarkets 2016). The goal of such investment is both to detect and manipulate emotional states. For example, one objective is to make advertisements more successful, placing them so that they speak to viewers' emotions in a more direct, personalized manner, and thus trigger the impulse to buy. This kind of attempt at manipulation is legal; however, it runs contrary to the ethical principle of protecting and promoting self-determination (Glenn and Monteith 2014). Experts such as Bauer et al. (2017) have called for clear boundaries to be drawn in the use of algorithms of this kind: "There must be a clear distinction between the algorithmic findings from the practice of psychiatry, and commercial findings for profit, even though similar analytic approaches are used."

Algorithms that provide information about mental states or tendencies thus raise the question of how to approach a possible "dual use" technology – that is, an application that pursues not only ethically desirable, but also morally problematic goals. The question cannot be addressed at the individual level. Rather, it must be approached at least at the institutional level. Individual institutions have already experienced the fact that software offered with positive intentions can be associated with significant safety risks when brought into practical use. Algorithms offering early detection of depressive moods, for example, could also be used by scammers or hostile individuals to find people who might be particularly easy to victimize. The "Samaritans Radar" app, for example, was originally developed to help with the early detection of depressive disorders. Briefly deployed in a social media environment in 2014, it was deactivated after only a few days due to security concerns. They related to the fact that knowledge of a particularly vulnerable emotional state can also be used by stalkers and internet trolls in order to inflict further harm on the person in question (Samaritans 2014).

Further challenges could arise when public and private actors, each with different interests and areas of focus, begin to work together. Instagram, for example, is a privately owned company. While private entities may indeed be interested in improving healthcare services, they are not providers of healthcare services operating within an institutional framework designed for that purpose. Commercial providers may also have an interest in collecting information about the emotional state of users on Instagram and similar platforms in order to improve their products or create more successful advertisements. It seems improbable that algorithmically obtained findings regarding emotional states will be used exclusively for predictive-medicine purposes (Bauer et al. 2017). (Public) healthcare service providers may ultimately engage in cooperative ventures with (private) commercial algorithm providers; this then raises the question of whether and if necessary what conditions of use are needed to safeguard users' rights (particularly the right to informational self-determination and the protection of personal privacy).

Sociopolitical challenges

The algorithm-assisted analysis of social-media content also presents a challenge at the societal level. This is mainly due to the fact that such an activity systematically blurs the demarcation between healthcare and other areas of life. The question is whether this mixing of different spheres (photo sharing, prediction of mental illnesses) might promote a medicalization of everyday life (Gadebusch et al. 2017: 95 ff.). A medicalization of this kind would be problematic insofar as health – including mental health – is widely recognized as a major societal good, but by no means the greatest such good. Health as a good stands in competition with other societal goods such as personal privacy or freedom, for example.

The use of predictive algorithms on social-media platforms such as Instagram raises the issue of a possible shift of social values. Would it be an overemphasis of the value of health if predictive algorithms were used in areas of life not previously associated with health matters? What societal consequences might arise as a result? Could new health obligations possibly develop? For example, a general duty to have one's social network postings checked for signs of mental illness? How likely is it that such a readiness to be screened by an algorithm could be seen as responsible health behaviour, while refusal to do so would be seen as health neglect? These questions regarding potential societal developments cannot be easily answered today. However, it is critical to be sensitive to questions of this nature; it will ultimately be similarly important to hold a societal debate on how to handle these new possibilities.

Medicalization of life

Personal privacy

Freedom

Shifting social values

New health-related obligations

6.2 Voice as an indicator of mental illness

When investigating non-verbal manifestations of mental disorders, past studies typically employed expert analysis of facial expressions, gestures, and gaze as well as voice patterns (Scherer et al. 2013). However, these investigations can be time- and cost-intensive. Furthermore, today's screening technologies for mental illnesses are primarily based on filling out questionnaires. These often provide only a very rough assessment of a person's mental state. Moreover, the questionnaires do not take any quantitative or qualitative information about non-verbal behavior into account (ibid.). Newer approaches thus examine voice recordings in order to identify possible indicators of mental illness (ibid.). For example, Scherer et al. (2013) have developed a machine learning algorithm that investigates voice-quality characteristics as indicators of the presence of depression or post-traumatic stress disorder (PTSD).

The researchers' objective was to facilitate the development of automated techniques for the early recognition of mental illnesses, to be used in combination with structured questionnaires and a quantitative analysis of non-verbal behavior. The information collected from the questionnaires, in combination with the automatically analyzed and quantified behavior, could make the diagnosis process more effective. To this end, the authors examined the potential of a small number of parameters as indicators of mental disorders. The aim was to induce specific patterns of behavior in the persons being examined. The researchers referenced in particular a meta-analysis showing that study subjects suffering from depression showed reduced affective reactions to positive emotional stimuli and stronger affective reactions to negative emotional stimuli in comparison to control-group participants (Bylsma et al. 2008).

Wizard-of-Oz experiment
Screening technologies
Automated techniques
Artificial intelligence
Early detection

The underlying training data for the algorithm was made up of semi-structured interviews with 43 participants at the University of Southern California in the United States. It was a so-called Wizard of Oz experiment, in which the study subjects believed they were communicating with a purely virtual figure or an artificial intelligence; in fact, the system was being controlled by a real person. This allows a finished virtual system to be simulated realistically and adapted to the reactions of the study subjects during development. Before the interviews, the study subjects filled out the PTSD Checklist - Civilian Version (PCL-C), a screening instrument for PTSD symptoms, as well as the depression module of the Patient Health Questionnaire (PHQ-9), a screening instrument for depression. During the interviews, participants were initially asked general questions; then, they were asked to respond to a series of questions that typically evoke negative or positive feelings. For example, questions included: "What are some things that make you really angry?" or "What are things that improve your mood?"

All of the characteristics examined in the interviews pointed to the fact that study subjects with moderate to severe depression showed more tension than did non-depressed participants. Moreover, gender showed no influence on the values of any of the voice-quality characteristics examined. These results could also be shown for the PTSD vs. no PTSD groups; however, the effects here were less pronounced.

Overall, by using the machine-learning algorithm, the researchers were able to distinguish between participants with and without depression with an accuracy rate of above 75 percent, simply through evaluation of the interviews. The accuracy rate for subjects with PTSD was 72.09 percent. According to the researchers, this rate could be significantly improved if future versions of the algorithm were to incorporate additional visual aspects such as facial expressions, glances, gestures and posture.

Opportunities

The opportunities associated with the analysis of natural speech in combination with machine learning correspond broadly with the opportunities for automatic image-processing systems as described above. Here too, the concept can be expected to contribute to improving psychotherapeutic care. In addition, due to the increased effectiveness of the procedure, quantitatively more and more accurate diagnoses could be rendered in relatively shorter periods of time. One crucial difference from the algorithm discussed above consists in the fact that this technique is intended to be applied only by professional healthcare providers such as psychiatrists, psychologists or psychotherapists. The machine-based support in the (differential) diagnosis of depressive disorders and PTSD is meant to aid them in the collection of case-history information – with the goal of producing a more accurate and more reliable diagnosis. The study's authors stress that the automatic collection of diagnostically relevant information such as voice modulation, speech intensity, articulation and speech pauses is often superior to observations made by professionals in the field (Scherer et al. 2013). However, they also note that the system is not meant to replace professionals but rather to furnish them with new diagnostic tools.

New diagnostics tools

Increased accuracy and reliability of diagnoses

Better psychotherapeutic care

The service standards even of comprehensively trained, competent and experienced professionals could in this way be raised further, and the examination of voice recordings could be rendered comparatively cost-effective (ibid.).

Individual-ethics challenges

From an individual-ethics perspective, two groups of people are particularly relevant here as objects respectively addressees of moral considerations: people suffering from mental illnesses, and the trained professional psychologists or psychiatrists. As in healthcare more generally, the principle of doing no harm also applies here. It states that interventions aimed at restoring or preserving health must not themselves lead to harm (Beauchamp and Childress 1977). Diseased people are often particularly vulnerable; this is also true of people suffering from mental illness. If algorithms for processing natural speech are used in combination with machine-learning technologies, it is accordingly important to ensure that the patients are not unnecessarily burdened by this. Thus, the prospect of improving a differential diagnostic procedure through the use of algorithms does not automatically justify the use of the new technologies. It must also be assured that these modified procedures will not harm the patients.

Doctor-patient communication

Principle of doing no harm

Freedom of choice

Informed decision

In the study examined here by way of example, the subjects believed they were communicating with a machine. While this was not in fact the case, the procedure is intended to work this way for future applications. It will therefore need to be clarified whether communication with an entirely virtual interlocutor is more beneficial for or more detrimental to the group of patients in question. Should the communication with virtual interlocutors be experienced negatively, the expected diagnostic advantage would have to be weighed against this additional factor.

Of course, positive effects are also conceivable. Some traumatized persons might find it easier to open up to a machine. Louis-Philippe Morency is leader of a project at the University of Southern California's Institute for Creative Technologies (ICT). Speaking about the use of virtual reality in the treatment of soldiers with PTSD, he says, "We have an issue in the military with stigma, and a lot of times people feel hesitant talking about their problems ... A virtual counselling tool can alleviate some of this reluctance" (Leithead 2013). It will be accordingly important to ensure that patients have the choice. Those who need or prefer a personal, human-led interview should still be able to receive this, even if a differential-diagnostic algorithm is in practice available (Deutscher Ethikrat 2018: 275). Treating physicians, psychotherapists and psychologists should should thus not be tempted to automatically adopt new technologies without questioning all implications of their use.

Within the course of everyday practice, this may generate new doctor-patient communication demands. Protecting patients from harm, along with the need to respect their right to self-determination, demands that they be involved in any decisions regarding the use of diagnostic procedures. However, incorporating patients into decisions requiring the evaluation of advantages and disadvantages of procedures can at times be difficult, especially if patients suffer from mental illnesses. Treating physicians here face the not insubstantial challenge of ensuring that all patients are able to make an informed decision as to whether they can and want to engage in the "conversation" with the algorithm. The new technological developments can thus create new challenges both for patients and therapists, especially with regard to understanding that the new possibilities are only an option, not an obligation.

Institutional-ethics challenges

From the institutional-ethics perspective, the primary challenge here is in designing framework conditions for the psychotherapeutic practice that reflect the observations made above. In addition, institutional actors should ensure that patients' right to informational self-determination is protected in a way that accounts for the fact that the algorithm described above will capture and automatically process sensitive data. In this context, storing, managing and dealing with the data will produce additional challenges.

Psychologically trained professionals are not necessarily trained to ensure that electronically collected data is adequately protected against unauthorized access and use. Rather, it can be assumed that even relatively tech-savvy medical professionals are today often only sketchily informed about the technical possibilities associated with tracking and correlating patent data. Moreover, many may not fully understand the methodological foundations of algorithmic applications in the sense of evidence-based medicine. In this regard, experts such as Bauer et al. (2017) point to a fundamental shortcoming in medical education and training: "(P)hysicians and administrators need education with regular updates from independent sources, not vendors selling products," (ibid.: 8). Without appropriate education and training measures focusing on the use of algorithms, there is a risk that data will be processed or stored with only insufficient protection.

Data storage and management

Education and training measures

Cooperation with computer scientists

In this regard, it is also necessary to clarify which specific groups of people, under what conditions, will have access to the data and datasets. Various groups of people may be entitled to this access in the context of the therapy. This may even include research psychologists seeking new diagnostic and therapeutic approaches. Ultimately, the determining factor must be the patient's consent (Lipworth et al. 2017). However, not every data transfer made in the context of a treatment plan will be individually discussed with the patient. The exchange of information between physicians and assistants, for example, is typically secured by previous consent without the need for an explicit additional authorization. Here, the principle of implicit consent applies (Vollmann 2000: 38). However, the creation and introduction of speech-processing systems demands not only psychotherapeutic expertise, but often cooperation with computer scientists as well.From an institutional-ethics perspective, this thus raises the question of whether the computer scientists or providers of the speech-processing systems must be subject to the obligation of confidentiality if they are to have access to sensitive patient data.

Sociopolitical challenges

Socioculturally, similar questions arise here, as the ones that have been adressed in the context of predictive algorithms for early detection of depression by image analysis (see section 6.1). With regard to the automatic analysis of natural speech too, it is necessary to ask which contexts the technology can be legitimately used in. It is quite conceivable that speech-analysis systems might also be deployed in non-medical contexts, with the aim of obtaining the most comprehensive view possible of other people's emotional states. We already noted the risk of the associated medicalization of an increasing number of areas of life. In addition, as has also already been mentioned, a variety of groups of people could have an interest in obtaining information on the affected persons' emotional states, with these interests being potentially problematic from the affected persons' point of view. Commercial entities could also seek to use the algorithm for manipulative purposes (Bauer et al. 2017). Even today, significant effort is being expended to place online advertisements in the most favorable position possible, or to manipulate the way users act within online services so that this behavior becomes reflexive or even effectively externally controlled (Eyal 2014).

Legitimation

Medicalization Manipulative purposes

Problematic interests

A discussion regarding the developing possibilities for manipulation here has been going on for years within professional circles (Zeng et al. 2009). From an economic perspective, efforts of this kind may be understandable. From an ethical perspective, however, they represent a challenge to the extent that they undermine the human capacity for self-determination. Sociopolitically speaking, this thus produces a set of challenging questions: To whom and under what circumstances should the ability be given to use algorithms to "read" emotional states from natural speech? Similarly, who will be allowed to use the information rendered available in this way, under what conditions, and to what end? Regulations limiting the use of algorithms of this kind to appropriately trained professional physicians, for prognostic purposes only, would be quite conceivable. However, a societal and sociopolitical discussion on this issue has yet to emerge.

6.3 Data-based support for physicians' therapeutic decisions

So-called recommender systems constitute another area of use for algorithms. These offer the ability to guide users to content selected specifically for them on a personalized basis (Burke et al. 2011). Currently, they are often used in online stores to provide users with targeted product suggestions. These automatically generated but simultaneously personalized suggestions are based on the data held by the system relating to that specific user. Other data can also be incorporated, such as the purchasing behavior and product evaluations produced by other people displaying similar user behaviors. This kind of recommendation system operates on the assumption that users showing similar interests will also have similar preferences in the future (Brandl et al. 2015: 229).

To date, recommender systems have seen little use within the medical context. However, they have significant potential. One example is the therapy recommendation system developed by Gräßer et al. (ibid.). This is based on two different algorithm-based recommendation systems: a "collaborative recommender" and a "demographic-based recommender." Both recommendation algorithms employ users' previous explicit and implicit evaluations as an expression of preference. In this case, preference should primarily be understood as meaning that a patient has responded positively to a therapy. The "collaborative recommender" considers evaluations by other users; here, primarily the response of these patients to certain therapies. On this basis, predictions are made about the preferences of other people – that is, about whether and how various individual patients will respond to a given therapy. In addition, the algorithm evaluates data containing the results of all the patient's previous medical consultations. The hybrid "demographic-based recommender" incorporates other available patient-related data as well as this initial data.

Recommender systems

Personalized suggestions

"Collaborative recommender"

Machine-learning algorithms

Predictive accuracy and precision

"Demographic-based recommender"

The goal of such a clinical decision-support system is to predict what therapy or therapies are advisable for a certain patient at a particular point in time. The recommendation system developed by Gräßer et al., was tested using therapeutic recommendations for patients with psoriasis, a skin disease. The underlying data was based on the hospital records of 213 patients from the University Hospital's Clinic and Polyclinic for Dermatology in Dresden, Germany. The data as a whole encompassed 1,111 medical consultations by these patients. It included patient and therapy descriptions, demographic data, and information on health status, comorbidities and current treatments. The data was manually transferred from the clinical records into a digital database; incomplete or erroneous data was corrected or removed. The data was processed by the algorithms with the goal of recommending the potentially most effective therapy for each participating patient. In a previous prognosis step, forecasts had been made regarding probable individual outcomes for all available therapies that had not yet been used for each given patient.

The researchers examined the accuracy and precision of both algorithms' predictions. Here, the "collaborative recommender" showed better results than the "demographic-based recommender." This was in part due to the similarity calculation contained in the latter tool. This proved to be unfavorably influenced by comparatively irrelevant information, while more important factors were given too little influence. This is likely to be improved in the future, particularly through the use of methods for selecting and weighting characteristics. In addition, the therapy recommendation must incorporate additional information if no or only limited data regarding the patient's previously pursued therapies is available. For the system presented here, the researchers linked the two recommendation approaches in order to compensate for the disadvantage of missing information. This also helped overcome the disadvantages of the individual data-mining and machine-learning techniques. The combination of the two approaches showed the greatest predictive accuracy and precision.

Opportunities

The recommendation systems described primarily create opportunities with regard to improving patient safety and increasing the effectiveness of professional medical activities. Faster and more comprehensive access to relevant information allows physicians to make better decisions in a timely manner. Recommendation systems are expected to contribute to significant prognostic improvements, because they bring together systematically organized information from patient records and from the most recent medical publications, and automatically detect relevant patterns such as potential drug interactions (Rüping 2015). As a result, errors caused by a lack of information can be avoided, and patient safety can be improved. If the time spent by physicians in obtaining information is reduced, this can ideally create time for other medical tasks, such as engaging in more detailed patient consultations or caring for a greater number of patients.

Avoidance of errors
Increased effectiveness Patient safety
Digital support
New findings available more quickly

Such systems could help prevent human error caused by information deficits, while also reducing errors attributable to occasionally unavoidable concentration gaps, for instance due to fatigue (Lepri et al. 2017). Bonderman (2017) summarizes these benefits as follows: "(O)ne of the most useful functions of using artificial intelligence in this way was: there is no human error."

Digital support in the processing of clinically relevant information is also useful to medical professionals especially in light of the increasingly great quantity of new medical publications. According to recent estimates, a new publication in the cardiology field currently appears every 2.7 minutes on average, for example (Bonderman 2017). This makes it impossible for individual physicians, and even for larger teams, to keep up with all new publications that may be potentially relevant to an individual treatment plan. Systems providing data-based support for physicians' therapeutic decisions can thus contribute to integrating new research findings more quickly and comprehensively into medical practice. Patient care can be continuously updated, and kept in alignment with the most recent state of scientific knowledge.

Provided that experiences gained from individual practice are recorded in electronic health records and released for the purposes of scientific evaluation, new findings from the field can also be expected to flow back into scientific research more swiftly and robustly. A comprehensive data-matching process, as is made possible by the use of algorithms, can in some cases also bring novel scientific findings to light. These in turn benefit the patients themselves, because the classification of patients into specific risk groups can be improved (Gräßer et al. 2017).

Individual-ethics challenges

From the individual-ethics perspective, we first have to ask whether the decision-support systems are in fact enhancing medical safety in all application areas. If systems of this kind were to produce results inferior to those of human decision-makers in some areas, their use would be ethically problematic for reasons of patient safety, and the principle of harm avoidance would be violated.

The question of application safety must ultimately be resolved empirically. This must also entail an examination of how possible system errors are to be handled. Critics of current developments warn that treating physicians may lose their decision-making power over medical decisions in the future (Cohen et al. 2014). A worse case scenario would be that the system's recommendations would be followed even if they were inferior to an individual expert's decision. The probability of diagnostic and therapeutic error would in this case be increased rather than reduced. In cases of doubt, an algorithm would generate erroneous recommendations not only in single cases, but systematically and continuously. Physicians relying on digital data and algorithms might not notice such flaws in the system, or might do so only at a relatively late date (ibid.).

In addition, the individual-ethics perspective prompts questions regarding the impact of decision-support systems on the physician-patient relationship (ibid.). The use of such systems could strengthen the relationship between physicians and patients insofar as physicians are given assistance in the area of their competence, and are thus able to focus more closely on individual patients both professionally and personally. On the other hand, the relationship could be impaired if the treating physicians were to become no more than "vicarious agents" for the recommendation systems. However, patient trust in the competence of physicians is an important factor for successful therapy.

For their part too, physicians themselves may also tend to give more weight to the information provided by the system than to the information given to them by the patients in person. The patient could in a certain sense become a mere "data subject" – that is, regarded merely as the bearer of certain data characteristics. However, in the relationship between physician and patient, the doctor's attention to the patient as a person is of significant importance for a successful diagnosis and therapy (Cohen et al. 2014; Fischer et al. 2016).

Erroneous recommendations
Doctor-patient relationship
Data subject
“Automation bias”
Loss of competence
Values
Self-determination

Bauer et al. (2017) refer in this context to the risk of an “automation bias.” Physicians are subject to such a bias if they pay less attention to the individual observations of their patients then to the algorithmically generated information. Such a bias is understandable insofar as the evidence comes from systems that are continually supplied with large amounts of data. Nevertheless, valuable contextual knowledge, such as that obtainable solely by a physician in a patient consultation, could be given insufficient attention or even ignored entirely due to excessive focus on the algorithmically generated knowledge (Bauer et al. 2017; Glenn and Monteith 2014).

If the use of algorithms were to produce this kind of loss of human and professional competence, it would increase the risk of erroneous medical decisions as well as the risk of a fundamental loss of trust in the physician-patient relationship. This concern affects not only the medical field, but also the use of algorithms to expand or replace human activities more generally (Mittelstadt et al. 2016).

Moreover, algorithms cannot provide other relevant services. These include the human capability to respond to patients’ personal preferences, fears and convictions, for example. The authors of the study considered as an example here emphasize that the algorithm they programmed is intended to refer back to the patients’ individual preferences and values (Gräßer et al. 2017). However, other authors such as Fischer et al. (2016) rightly indicate in a similar context that it is always necessary to be familiar with the patients’ values and preferences at a personal level in order to translate a prognosis into medical advice. Only when patients’ personal values and interests are known can an individual choice between various trade-offs be made (Mittelstadt et al. 2016).

This individual choice is of great ethical significance due to the respect that must be shown for patients’ ability to engage in self-determination. From an individual-ethics perspective, it thus becomes of critical importance whether predictive or prognostic statements are based solely on clinical findings, or also take patients’ values and interests into account.

Institutional-ethics challenges

From an institutional-ethics perspective, the challenge here is once again that of designing institutional processes so that the various actors involved are supported in making ethically good decisions, and in acting accordingly. Decision-support systems raise the concern that future physicians, acting from liability-law concerns, might always tend to follow the algorithms’ recommendations, potentially even against their own better judgement (Cohen et al. 2014). It is likely, runs this reasoning, that in the case of harm due to either human or machine error, the deciding factor would come to be the different levels of justification demanded. That is, if an individual therapeutic decision by a treating physician or team were to deviate from the algorithm's recommendation and harm were to ensue, the justification for the decision would be more difficult to sustain than in the reverse case. In the future, if an erroneous medical decision were to be made, a plaintiff could refer to the recommendation provided by an algorithm as proof of the physician's misconduct (Cohen et al. 2014).

In addition, difficult questions of responsibility are raised here: As Lepri et al. (2017) note, "Ultimately, we need accountability in decision-support algorithms such that there is clarity regarding who holds the responsibility of the decisions made by them or with algorithmic support." This is not easily resolved. To a certain extent, the algorithm becomes a co-decision-maker. However, it is not an actor in the strict sense (Jaume-Palasí and Spielkamp 2017). The work of the developer also has an influence on how good a given algorithm's recommendations can be. Furthermore, responsibility here is also shared, as programmers often work in teams on the development of an algorithm.

In any individual case, it may be thus virtually impossible to determine who has made what error, and who should be held responsible in case a harm occurs (Mittelstadt et al. 2016). The issue is aggravated by the fact that the performance of algorithms also depends on many additional factors, including the quality of the data and correct handling by the user, computer scientists and programmers. It is accordingly extraordinarily difficult to arrive at a clear attribution of responsibility if mistakes are made. One putative solution to the problem is to reserve the liability for any final decision expressly to the treating physician (Cohen et al. 2014). The last say on individual decisions in medicine, as Bonderman (2017) points out, may not be given to an automated "expert system."

Allocation of responsibility Failures

Transparency and intelligibility

Knowledge of how algorithms function

"System bias" Specific training

However, this demand may be only provisorily feasible. It is important to bear in mind that algorithms are more likely to support some specific decisions, while hindering others: "Even if a human has formal responsibility for making the final decision, ... the system leaves only limited maneuvering room. It is thus rather unlikely that a human would revise an algorithm's preliminary decision or adopt it only in part" (Wagner et al. 2017: 12; translation by the authors of this report).

All systems containing machine-learning algorithms pose particular challenges here. In these cases, neither the programmers nor the users have an adequate understanding of or insight into the processes relevant to generating the recommendation. From an institutional-ethics perspective, a number of challenges relate particularly to the transparency and intelligibility of the algorithms' mode of functioning. Responsible clinical decisions require sufficient knowledge regarding the way the system functions to understand the advantages and limitations of the various recommendations, and to be able to evaluate and classify the proposals automatically generated in each specific situation (Rüping 2015).

This capacity can be partly fostered by specific training programs. To some extent, however, the problem persists even following the implementation of comprehensive training measures. Transparency requirements often fail not due to the user's particular lack of understanding, but due to the construction of the system within which the interpretability of the models may only partly be considered (Rüping 2015). As Amarasingham et al. (2016) note: "Transparency in healthcare predictive analytics must be carefully implemented. Instead of a one-size-fits-all approach, a transparency framework should be adaptable to tailoring a predictive model's prototype, complexity and users."

Unsupervised and semi-supervised machine-learning algorithms pose particular problems in this regard. They independently create new instructions and models that are then used to control other processes. The range of services performed by these systems therefore necessarily entails significant uncertainty with regard to how and why specific results are generated. Users can accordingly make no distinction between isolated flaws (in the sense of individual "system bugs") and systematic flaws, for example due to "system bias" (Mittelstadt et al. 2016).

Institutional actors thus face significant challenges with regard to enabling responsible action by individual actors. One critical task is to clarify the conditions under which physicians should be allowed to ignore recommendations issued by the system. Initial proposals in this regard have already been formulated. For example, Cohen et al. (2014) propose that a machine-generated recommendation for the treating physician should be revisable depending on the risks associated with the specific intervention, and on what is known about both the therapy's complication rates and the recommendation system's error rate.

Sociopolitical challenges

The issue of responsibility for (mistaken) medical decisions and/or mistakes by recommender systems also has a sociopolitical dimension, particularly with regard to the need to adapt existing liability law (Cohen et al. 2014). An additional and more fundamental question is that of which medical contexts and for what objectives recommendation systems should be used – or should be allowed to be used – in the future. For instance, should algorithms also be used to generate prognoses regarding critically ill patients' remaining lifespan and their probable quality of life? The U.S.-based company Aspire Health, which is partially financed by Google, is already using an algorithm to predict the likelihood with which patients will die in a week, in six weeks or a year. In addition, the system could calculate the costs of different treatment plans, and thus the total costs that will be incurred by an individual patient (http://aspirehealthcare.com/).

Should such systems be used to say something about the efficiency of different treatment methods, or to determine the prospect of a positive cost-benefit ratio for individual patients? And how far should the use of algorithms for prognostic purposes go? Should algorithms also provide decision support when it comes to issues such as whether life-support measures should be maintained or terminated?

The healthcare system is laboring under increasing cost pressures. Recommendation systems could also be introduced with the intention of reducing this pressure. Potential savings could result from the replacement of human expertise and labor, of course. Aside from such ethically potentially problematic savings, however, such systems might also be used to determine the best possible cost-benefit ratio among the available therapeutic alternatives. This possibility in turn must be distinguished from approaches that serve the purpose to decide which patients are “worth” a particular treatment, and which patients are not. This prospect raises particular concerns. For example, the announcement of the Aspire Health algorithm prompted media discussions as to whether costs for therapies at the end of life could be reduced (Beck 2016).

The use of algorithms to achieve this kind of comprehensive optimization of the cost-benefit ratio could undermine trust in the healthcare system over the long term, as this would set aside a core healthcare principle – specifically, the respect for human dignity. To a great extent, the respect for the dignity of every person is expressed by taking the patient’s right to self-determination into account. To be sure, when making decisions regarding the allocation of scarce resources, different treatment prospects and costs must be considered for reasons of fairness. However, these considerations should not come down to a healthcare system in which certain patients or groups of patients are systematically excluded from care (Brock 2003).

Cost-benefit balance

Human dignity Distributional justice

Confidence in the healthcare system

In the media-shaped public debate, worries have been raised that the use of algorithms in decision-support systems could promote a development of this kind (Lobe 2017). The degree to which such concerns are justified is impossible to foresee at this point. In any case, it is critical to open a societal discussion regarding the conditions under which physician decisions should be supported by algorithms, and those under which the use of such digital systems would not be allowed.

6.4
Prediction models for illnesses and drug effects

As healthcare systems become increasingly digitalized and networked, it is possible to gain access to a great quantity of new (patient) data. On this basis, algorithms can develop quantitative models that can be used to make increasingly precise predictions regarding the emergence of diseases or the occurrence of medication side effects and drug interactions.

The Deep Patient system developed by Miotto et al. (2016), based on a deep learning approach, is one example of a prediction model of this kind. This system draws on all the data about a person available within an electronic health record. This includes data on sociodemographic characteristics such as age and gender; clinical data such as diagnoses, prescribed medications or laboratory tests; and data on inpatient, outpatient and emergency treatments. However, difficulties in preparing such data appropriately represent a key challenge in enabling it to be processed by a machine. The data is often both heterogeneous and incomplete; moreover, it frequently contains both random errors and systematic distortions (ibid.). To address this problem, the researchers programmed a so-called deep neural network that automatically detects stable structures and regular patterns in the data. By using data prepared in this way, the system was intended to enable significant improvements in the prediction of illnesses and drug effects.

Quantitative models
Deep learning algorithms Monitoring
Probability-based prediction
Random-forest classification procedures

To test Deep Patient's effectiveness, the researchers used a comprehensive database of electronic health records maintained by New York's Mount Sinai Hospital. The information used dated back to 2003. Data from 1.2 million patients was included in the evaluation. Using the patients' clinical statuses as a basis, the goal was to predict the probability with which individuals among this group would be diagnosed with certain illnesses over the course of a year. In developing this probability forecast, a random forest classification procedure was used, which in turn was trained in advance for each disease of interest using a dataset containing information on 200,000 patients. To this end, all clinical data was pre-processed in order to obtain harmonized codes, for example for the various procedures and laboratory tests.

The results show that Deep Patient achieves significantly better results than other prognosis systems based on machine learning techniques. The system produced a high rate of accurate predictions particularly with regard to conditions such as prostate cancer or sickle-cell anemia. One key feature of Deep Patient is that unlike previous systems, it is not targeted at any specific condition; rather, it considers all available data about a patient's health status. According to the developers, the system could be used in the future to create personalized medication plans, make treatment recommendations and recruit subjects for clinical studies.

In so doing, Deep Patient could support clinicians in their day-to-day work, for example helping to monitor patients and automatically check whether the appearance of a given disease is probable in the near future on the basis of their current clinical status. In hospitals, the system could automatically identify patients at risk of developing a particular condition, and generate warning messages accordingly. The researchers believe that in order to represent patients in an optimal manner, and to enable correspondingly reliable prediction models, additional information such as insurance data, family background and social behaviors should be integrated into the system in the future.

Opportunities

Prediction models such as the Deep Patient approach described here combine a large quantity of different health-relevant patient data, clinical data and data stemming from other contexts. Holistic approaches of this kind are expected to enable significant progress toward the development of personalized medicine. Ideally, automated data analysis will in the future allow to identify individual health risks and the probable course of diseases significantly earlier, with a higher degree of certainty. This in turn opens up the prospect of more effective early interventions, tailored more precisely to the individual person and condition. The Deep Patient developers point specifically to potential improvements in the prediction of various severe conditions such as cardiovascular diseases, heart failure and chronic kidney disorders. They also note the potential of obtaining better information about possible drug interactions (Miotto et al. 2016).

Beneficence

Non-maleficence

Personalized medicine

Justice

Early, tailormade interventions

Other authors too associate the use of deep learning algorithms with significant opportunities for clinical decision-making, and stress the use of such technology brings personalized medicine and systems medicine into the realm of the possible (Fischer et al. 2016; Amarasingham et al. 2016; Yuste et al. 2017). In order to achieve this, a large quantity of very heterogeneous data needs to be automatically evaluated and combined. In addition to clinical data, additional data from the patients themselves – for instance, data collected by so-called fitness wearables or other sensor containing personal devices – can be processed. Systems that process and interlink these and a wide variety of other data promise direct prognostic benefits. They may furthermore be used they to track health developments even over long periods of time (Fischer et al. 2016).

Aside from providing immediate healthcare benefits to specific patients, the use of such data could help optimize existing medical treatment practices and even medical care overall. Moreover, the selection and recruiting of subjects for clinical studies could be placed on a more efficient footing. The algorithm-supported evaluation of a large quantity of health-relevant data could allow to be more selective in choosing the subjects most appropriate for a study, while excluding subjects that might be less appropriate or even endangered by participation.

In sum, this technological advance presents major opportunities for improvements relating to several of the basic principles of medical ethics, including that of beneficence, non-maleficence, and potentially even justice. Improvements with regard to the principle of beneficence are possible insofar as the use of deep learning algorithms facilitates the development of personalized medicine, which in turn is likely to enable better medical care for individuals. Non-maleficence comes into play here if possible drug interactions or criteria for excluding certain individuals from participation in studies can be identified. The principle of justice would be served if and to the degree that algorithms could contribute to addressing care bottlenecks, preventing a random and accordingly problematic allocation of scarce medical resources.

Individual-ethics challenges

Many of the challenges identified above with reference to algorithm-based recommendation systems also arise here. There is no need to revisit them anew. Instead, we focus on a further challenge posed here. From an individual-ethics perspective, the extensive collection and automated analysis of patient data raises questions particularly regarding the ability to respect the right to informational self-determination, as well as the ability to ensure that data protection requirements are met (Mittelstadt et al. 2016).

The patient data compiled here is mostly collected in a therapeutical context, and thus based on a therapeutical contract. Accordingly, also the informed consent to the processing of data is given within – and limited to – this specific context. If consent for further processing of the data is given, for example for use in the framework of clinical research, this release generally refers only to use within the comparatively clear context of the specific clinical setting. By contrast, the possibility that various pieces of data collected in specific medical contexts might be digitally linked together and analyzed – perhaps even decades later – is generally not something that patients are conscious of at the time of the data collection. Indeed, in the case of Deep Patient, this collection took place beginning in 2003. The comprehensive analysis of this data in a subsequent use of this kind cannot be legitimized through "implicit consent."

From the individual-ethics perspective, this problem of insufficient or even lacking consent to the processing of individual personal data could be resolved, at least in the current research phase, by anonymizing the data. However, the actual purpose of the system – to benefit the health of the individual patient – would be nullified as a result. In addition, it should be borne in mind that new data-linking and data-tracking technical capabilities make it virtually impossible to ensure that patient-related data has been completely anonymized (Cohen et al. 2014). This poses severe challenges with regard to the right to informational self-determination.

The prospect of patients being harmed due to the misuse of their data raises further challenges. Given that data owners are in principle (re-)identifiable, they could be harmed if their data was made available outside the research context. An individual could be subject to a number of disadvantages resulting from the disclosure of a current, past or even prospective health status. For example, it is conceivable that similar systems might be used by employers or insurers in order to gain information about the health of potential employees or customers (Bauer et al. 2017). The use of data in this way would not only contravene the affected individuals' right to informational self-determination, but could also directly harm them insofar as they were not offered a job, or were forced to pay higher insurance premiums, for example.

Informed consent

Right to informational self-determination

Data misuse

Anonymization and re-identification of data

The data owners may also be harmed if the comprehensive analysis of his or her datasets produces predictively relevant information regarding potential future diseases for which there is no treatment is available. It seems likely that as analytical functions are developed further, an increasing quantity of "incidental findings" will be produced (Fischer et al. 2016). In this case, the treating physician is put in a position where he or she needs to decide whether or not to inform the patient. Being notified regarding a future illness that is more or less likely to occur could be a significant burden for the individual thus affected. Additional diagnostic measures could become necessary if the individual – who is (still) healthy at the time the prognosis is issued – is to be discharged with the knowledge that he or she is expected to develop a condition for which there is no therapeutic treatment available yet. Under these circumstances, dealing with incidental findings would be an urgent and difficult ethical problem for patients, algorithm developers, physicians, hospitals and insurers (Drazin et at. 2013; Lipworth et al. 2017; Fischer et al. 2016).

Additional ethical questions are also raised here by the prospect of using data to draw conclusions regarding health-related and other characteristics of family members. For example, this could be the case if genetic markers for the detection of cancer were to be used as part of the analysis (Fischer et al. 2016). The right of the genetically related person to informational self-determination would clearly be infringed if they were identified in the course of the analysis without consent. In addition, if prognoses unfavorable to them were to become known, they too could be harmed if the data was distributed further.

Institutional-ethics challenges

The issue of dealing appropriately with various types of findings primarily arises out of the institutional-ethics perspective. Medical facilities are tasked with protecting and promoting patients' welfare, and also with safeguarding their right to informational self-determination. They must consequently develop appropriate procedures and concepts of protection in order to protect these patient interests.

From an institutional-ethics perspective, there is the additional question of when systems of this kind can be considered sufficiently secure. Experts note critically that in today's typically fragmented healthcare systems, the lack of data compatibility is often an impediment. Even the coding of diagnoses is often handled in different ways. Thus, algorithms must work with data that is coded using very different systems (Lipworth et al. 2017). The evaluation of these data by deep learning algorithms can lead to distortions. Patients may be reduced to a small part of their data, for instance (Miotto et al. 2016). Prematurely introducing a data-processing system into the medical practice poses both medical and social risks. For example, if immature systems were to be employed with the goal of accelerating workflows as a response to increased cost pressures in spite of the fact that deficiencies could endanger the health of patients (Cohen et al. 2014).

Other institutional-ethics questions relate to the responsibility to use the collected data properly. Health-care institutions have the responsibility to decide who is allowed to collect and analyze patient-related data, for what purpose, and over what time period. For example, as algorithms designed for comprehensive patient-record processing are developed and used, should institutions such as insurers and hospitals be authorized to share patient data with commercial providers?

Such questions are in no way trivial. A few years ago, it became known that clinicians were willing to relinquish patient data to commercial entities in sometimes irresponsible ways. A case analyzed by Powles and Hodson (2017), for example, featured the extensive exchange of data between the English Royal Free London NHS Foundation Trust and Google's DeepMind project. The patients were neither informed about this nor did it appear that the clinicians involved were aware that they were enabling a provider like Google DeepMind to release the data for a variety of other, partly commercial, applications (ibid.). This case indicates that there is still a strong need for education and training around these issues.

Proper use

Diverse security risks

Incidental findings

Immature systems

In addition to an obligation to provide protection of patients' right to informational self-determination - also against commercial interests - healthcare institutions have the obligation, to protect their patients' data from criminal access. Under certain circumstances, a centralized archive of large amounts of data, as is necessary for Deep Patient, makes digital systems particularly attractive for criminals. Developers of digital technologies already face significant challenges today with regard to ensuring data security. Various hospitals have already had the experience of having their systems hacked by criminals, with critical patient data "taken hostage." The hospitals' access to their data was restored only after the payment of large (Bitcoin) ransom sums (www.dw.com/en/hackers-hold-german-hospital-data-hostage/a-19076030).

In short, the use of increasingly large interlinked databases that can be used in a variety of ways with deep learning algorithms also generates diverse security risks.From an institutional-ethics perspective, developers and users thus face the challenge of clarifying who, and under what circumstances, is responsible for the security of the data undergoing processing. Parties bearing partial responsibility include the developers, the companies offering the algorithms, the medical institutions, and even the physicians and medical researchers themselves. It can generally be assumed that such entities are aware of potential hacker attacks. However, institutional policies for dealing with such events in accordance with the terms of the General Data Privacy Regulation have in some cases yet to be developed. Without such policies, individuals tasked with handling sensitive patient data responsibly would be obviously overburdened.

Sociopolitical challenges

Many of the socio-political questions already addressed in the context of other application examples arise again here. For example: How are novel liability-law questions to be dealt with? How much knowledge in what health-related areas is societally desirable? What "side effects" might be associated with the expansion of health-related knowledge ? And how should this knowledge be handled? How can (and should) society respond to the prospect of a progressive "medicalization" of increasingly numerous areas of life?

The emergence of Deep Patient and comparable applications renders this final question still more urgent, since everyday data is also fed into the system here, and linked to clinical data. Users of fitness wristbands are unlikely to see themselves as patients. However, combining and cross-referencing data opens up ever more predictive possibilities; this in turn could lead to a situation in which the distinction between health-related behavior and other aspects of day-to-day living becomes increasingly difficult to ascertain (Deutscher Ethikrat 2018: 120).

Socially desired knowledge "Side effects"

"Medicalization"

Discrimination risks

A trend of this kind could carry societally problematic developments in its wake. For example, it could lead to a situation in which more and more people feel obligated to collect their data on a comprehensive, ongoing basis, and make it available for health analysis. Anyone who chooses not to divulge extensive amounts of data for others' use, on whatever grounds, could as a result be subject to long-term disadvantages. The quality of these people's healthcare could deteriorate in comparison. In part, this would simply be because the informational base underlying preventive or therapeutic decisions about them would be relatively less robust. But they could also suffer disadvantages such as having to pay higher insurance premiums than people whose current and future health states can be comprehensively analyzed and treated accordingly (IBC 2017: 17 f.; Becker and Strammer 2016: 510 ff.).

A societal debate is needed to determine whether such discrimination against anyone who "refuses" the analysis of personal data is legitimate. The challenge is to determine under which conditions it is justifiable to treat people who are ready to provide personal data and those who decide otherwise, unequally. In this context, it would also be necessary to discuss how potential discrimination risks associated with membership in specific risk groups can and should be handled (IBC 2017: 17). As seen in many other contexts, classifying individual persons into specific groups can create discriminatory effects. A broad societal discussion on how these possible effects should be dealt with is necessary.

6.5 Activation and restoration of paralyzed persons' mobility

Millions of people worldwide suffer from paralyses caused by disturbances in the neural pathways between brain and muscles. Recently, these paralyses have been treated using neuroprostheses, which create an "electronic neural bypass", and thus route around the interrupted pathways in the nervous system. This involves the use of a number of algorithms, which will be described here using research by Bouton et al. (2016) as an illustration. In specific terms, this entails a system that registers intracortical signals (meaning within the cerebral cortex) and connects them with the muscle-activation function in real time, seeking in this way to restore mobility to paralyzed people.

Neuroprostheses

Mobility for paralyzed people

Robotics systems Machine-learning algorithms

The system is based on previous studies that showed that decoding intracortically captured signals helps in extracting and processing information about movements. If this information is linked with a robotics system, it gives people (or primates) the ability to control computers or robot arms solely by engaging in imaginary movements (Hochberg et al. 2012; Aflalo et al. 2015). For this purpose, a sensor (an implanted intracortical microelectrode array) must detect the activity from the motor cortex. An accordingly programmed machine learning algorithm then works to decipher the neural activity. This algorithm also helps to control the activation of the forearm muscles through the use of a newly developed neuromuscular electro-stimulation system. By using this system, the paraplegic test subjects were able to continually control six different wrist and hand movements, and were able to manage functional daily-life tasks. These results show the great potential for the use of machine learning algorithms in helping to activate muscles on the basis of intracortically captured signals.

Opportunities

The use of algorithms in brain-computer interfaces to decode movement-relevant brain patterns, thus rendering computer and robot arms controllable by thoughts. This technology may give people with mobility constraints the opportunity to significantly expand their radius of action and to become more independent of the support of other people (Hildt 2011). The gain in motor skills is considered to be suitable for significantly reducing the affected individuals' mental and social suffering (Nicolas-Alonso and Gomez-Gil 2012). The increased independence, in turn, may allow more privacy (Hildt 2011).

The research results suggest that in the future, it will be possible to tranfer movement impulses in real time to computer-controlled extremities that will be able to carry out even demanding fine-motor movements. The freedoms of movement and action thus obtained are of significant benefit to paralyzed individuals. They make it easier to cope with everyday life, increase the individual's abilities to self-determination and contribute to a fundamental improvement in the quality of life (Bouton et al. 2016).

Gain in motor skills
Reduction of mental and social suffering
Independence Privacy Scientific findings

The use of algorithms in brain-computer interfaces could additionally generate new knowledge regarding the functioning of the human brain. This in turn is expected to lead to numerous opportunities in medical practice more broadly, which could benefit people with severe brain lesions, for example. Overall, research into the human brain could derive significant advantages from the use of algorithms due to their accelerated data-processing capacities (Jordan et al. 2018).

Individual-ethics challenges

From the individual-ethics perspective, the use of algorithms in brain research and in practical medical applications raises numerous questions. These relate particularly to the protection from harm (Glannon 2014), the protection of personal privacy and the right to self-determination (Jebari 2012).

An recording of neuronal patterns requires an invasive intervention into the brain. Such interventions are associated with considerable risks (Hildt 2011). Infection risks as well as the risk of brain lesions have to be considered. Thus, such interventions are undertaken only after a rigorous risk-benefit-assessment, and require that the concerned person is comprehensively informed in the process of optaining informed consent. These ethical requirements are not specific to this kind of research, and thus do not require a separate explanation. The reference is relevant above all because the use of brain-computer interfaces (BCIs) on non-medical grounds, for instance for enhancement purposes, is very difficult to justify given the associated medical challenges and risks. At least today, any benefits to be expected are not sufficiently favourable in relation to the risks.

In addition to medical risks, additional risks must be considered relating to the affected individual's right to self-determination and the protection of their privacy. Automatically capturing brain activity represents a potentially significant encroachment on personal privacy and even into a person's self-understanding as an autonomous actor (Jebari 2012). Some years ago, studies have shown that the impulse to carry out a particular action – for instance, to push a button – is already detectible before the person in question has consciously made the decision to act. This short but sufficiently clear temporal lead shown by the brain's activity is today prompting discussions among neuroscientists and philosophers regarding the degree to which the brain has in some sense already made a decision before the person gains a clear idea of his or her own intentions (Eagleman 2016: 107 ff.). New technologies allow this brain activity to be captured and harnessed. In the absence of consent by the affected individual, insight of this kind would represent a significant overstepping of the bounds of personal privacy – and would thus clearly be illegitimate from an ethical point of view.

Neuroscientific studies furthermore indicated that registering and analyzing a person's brain activities could lead to significant changes in that individual's self-perception. The prospect that recording and automatically translating brain activities might influence a person's self-perception as an autonomous actor is of direct practical relevance. In a study on the use of BCIs by Gilbert et al. (2017), some subjects said they perceived a loss of control and felt alienated from their self, for example. Subjects receiving BCIs reported that they subsequently perceived their disease as being a constant presence in their lives, while it previously had been only episodically relevant (ibid.). Thus, these measures can have negative effects on subjects' or patients' self-perceptions, and on their perception of their state of health. Currently, before such an intervention is made, it cannot be known whether any such negative effect will occur, or how strong it will be if it does. Moreover, also positive experiences are possible. In fact, one subject reported the perception that the BCI had become a part of the self – an effect that was welcomed by this person (ibid.).

Potential for harm
Right to self-determination
Enhancement
Loss of privacy
Self-image

Neuroscience is a relatively young discipline, with accordingly many yet-to-be answered questions. This necessarily includes questions regarding the legitimacy of using algorithms to create and use BCIs.

Institutional-ethics challenges

From an institutional-ethics perspective, this use of algorithms once again raises the issue of responsibility. If the algorithms used in a BCI lead to harm for the patient, who bears the responsibility: the person or team that developed the algorithm, the robotics system or the medical professionals who utilized it? Control over this kind of system is shared. Thus, it is necessary to ask who is responsible for what failures (Gilbert et al. 2017).

Problems such as the differing assessments of the changes in self-perception produced by the BCIs indicate that the use of such technology is associated with new requirements for medical professionals. In order to ensure that patients have adequate information, medical institutions must define consent conditions specifically tailored to the use of novel technologies, and design appropriate explanatory materials and processes (Klein and Jeffrey 2016).

The issue of data protection produces additional institutional challenges. In the course of registering brain activity, highly sensitive data material will be processed. The unique character of this data makes it imperative to address the issue of who will be allowed to collect the data in question, with what purposes, and who will be allowed to save and use the resulting data sets for how long. Data security must therefore be taken particularly seriously, as the consequences of such drastically expanded knowledge about the functioning of the human brain cannot be foreseen today.

Sociopolitical challenges

On the societal and sociopolitical level, this example initially raises issues similar to those discussed in the context of other applications. Here too, it will be necessary to respond to new liability-law challenges with suitable regulations. For example, it would be necessary to clarify who would be liable if an accident was caused by an algorithm-assisted externally controlled robotics system (Mattia and Tamburrini 2015: 733 ff.). New challenges arise in terms of how to ensure equitable access to emerging medical opportunities. The use of such systems will presumably be associated with high costs well into the foreseeable future. Under what conditions should people with mobility limitations be provided with systems of this kind?

New demands Responsibility

Highly sensitive data

At the societal level, discussion is also necessary regarding whether this kind of technology should be – or should be allowed to be – used for non-medical purposes as well. A technology making machines directly controllable by neural activity could also be used for military purposes, for example. A direct connection between the brain and the machine being controlled could significantly increase pilots' reaction times during combat operations, for instance. The military use of BCIs is commonly considered to be "dual use," and has been discussed critically by ethicists engaged with the issue of robotics for some time (Kotchetkov et al. 2010; Burwell et al. 2017).

In addition to such questions of responsibility and justice, the use of algorithms in BCIs also poses particular challenges for society at large. The deeper understanding of brain activity associated with this application is likely to have profound consequences for the conception of man. The human being's self-conception as a person, in the sense of being an autonomous actor with a specific social identity, is already being increasingly shaken today by current neuroscientific findings (Burwell et al. 2017).

The use of algorithms is likely to accelerate the pace of such changes even further, as it creates a series of novel opportunities that will produce further uncertainty. How these challenges are to be addressed at the societal level is thus an increasingly urgent question. Should such research be rigorously regulated, and allowed only in the context of very specific application areas? Should support be provided for fundamental research , or for research with the objective of specific medical applications? To decide these questions, is debate necessary at a broad societal level? Or are such specific questions solely a matter for experts in accordingly specialized bodies?

Last but not least, the algorithm-supported use of BCIs fuels a fundamental question about the future of humanity, discussed for years under the rubric of transhumanism and post-humanism. Among the controversial questions here is whether the course of technological developments will ultimately result in a division in the human species into cyborgs and "conventional" people, and how such a development should be judged from an ethical perspective (Smart and Smart 2017; Hildt 2011).

In fact, the use of BCIs appears to make possible the fundamental abolishment of the divide between human and machine (Attiah and Farah 2014). The literature on the subject has already featured discussions as to whether we are witnessing the beginning of a fundamental transformation of the mankind (Burwell et al. 2017). For reasons of space, we cannot follow this discussion in greater detail here. However, it should be noted that the technical intervention in the human brain for the application described here is obviously significant. If developments toward the division of humanity into groups of technologically deeply transformed and untransformed people should proceed further, the societal and sociopolitical task of protecting the moral rights of all persons would become increasingly important. Individual identity, in the sense of bodily and mental integrity, as well as the individual capacity to make decisions and act on them, are key aspects of basic human rights and must accordingly be protected. Of course, such protection must be provided today as well with regard to the use of BCIs (Yuste et al. 2017).

One additional challenge relates to the question of whether the use of algorithm-assisted BCIs might ultimately result in stigmatization or discrimination against those who elect not to adopt these novel opportunities. Some people with mobility impairments might decide to forego such an intervention because they consider the associated risks to be too great. However, they might also make this decision because they are satisfied with their current level of capabilities. From their perspective, a change in their condition would would not count as a treatment but as an enhancement.

Liability law

Fair access Self-conception

Social identity

Divide between man and machine

This raises the question of whether the creation of new options of this kind might produce a discriminatory societal environment, or further intensify existing tendencies toward discrimination against people with impaired mobility. If this should be the case, society as a whole is asked to take countermeasures to ensure that the intended increase in self-determination capacities for people with mobility impairments will not instead result in new pressures and problematic social restrictions.

6.6 Alarm systems In residential communities for senior citizens

Algorithms play an important role in the development and use of so-called ambient assisted living systems (AAL), which are digital assistance systems that give people requiring care support in coping with their daily lives. Such systems are often used as a means of enabling the individuals concerned to lead a self-determined life in their own homes despite physical and/or cognitive disabilities (Braun et al. 2016).

An illustrative example of the use of algorithms in an assistance system of this kind is offered by the study project by Rantz et al. (2013). Here, the researchers developed an alarm algorithm that used a network of passive sensors installed in the home environments of residents living in a senior-citizen community in Columbia (United States) to automatically monitor the elderly individuals. The algorithm detects typical signs of illnesses and accidents, and can thus generate an automatic alarm. This makes it possible for nursing staff and doctors to intervene quickly, preventing or at least delaying severe deteriorations in health status or the ability to function. The researchers hope the system will enable senior citizens to remain in their own homes until the end of their lives despite impairments and illnesses, receiving support as necessary.

Ambient assisted living systems (AAL)

Coping with daily life

Automatic monitoring

Individualization

Algorithms are primarily used here to improve the early detection of illnesses, and to counteract the progression of chronic diseases. The data required is provided by sensors installed in the residents' homes. In the study by Rantz et al. (2013) this included infrared motion-detectors that monitored the presence and activity of the person(s) in the apartment, as well as sensors placed under bedsheets or in a chair that automatically registered pulse and breathing data. Thus, measuring instruments embedded in the environment served as a replacement for devices carried on the body or measurements carried out actively. One of the researchers' goals was to increase convenience for the residents. For example, the automatic sensors cause that residents no longer have to break their daily routine to enable medical data to be collected. However, this model also did a significantly better job of ensuring that the vital data was actually collected. Data was collected regardless of whether it occurred to residents to measure their pulse rate or other vital signs.

Using algorithms, an individual profile is created on the basis of this data. This individualization is important, as every person is unique. A "one-size-fits-all" model would not work. Thus, the algorithm for each person in the study was individually modeled, and further adjusted based on feedback from the physicians. If the activity and vital-signs patterns captured by the sensors deviate from the typical patterns established for this person, an alarm is triggered and the treating physician is contacted by email. He or she can then visit a secure website that displays the sensor data, and analyze whether any additional action needs to be taken.

Opportunities

Living safely for as long as possible in an accustomed environment is for many people an objective of considerable value. For them, living at home amounts to a better quality of life and a greater ability to lead a self-determined life even in old age. AAL projects like that described above are thus responding to an interest that is widespread and for many people of central importance. In this regard, sensors integrated into furniture and other everyday objects offer the opportunity to live largely unaffected by the technology, while nonetheless benefiting from the monitoring and analysis of vital-sign and activity data.

Self-determined life

Quality of life Individualization

Real-time detection of risks

Reduction in workloads

The use of algorithms enables such functions to be highly individualized, thus improving the systems' reliability. Deviations from individually typical vital-sign and activity data can be automatically detected in real time and identified as personal indications of a threatening illness or deterioration in health status, enabling swift intervention by automatically informed medical personnel. All of this could produce significant health advantages for the affected individuals. The authors of the study cited above also point out that interventions carried out as early as possible are often both particularly effective and less costly, as conditions receiving early treatment are often associated with less loss of function than those that are detected and treated relatively late (Rantz et al. 2013).

There are advantages for nursing staff and clinicians too, as they are relieved of the task of constantly checking in on the resident’s physical well-being (ibid.). Ideally, this will give them time for other tasks, such as conversations with the facility’s residents. Even over the long term, algorithms will presumably be unable to take over activities of this kind, which require the capacity for empathy (Bonderman 2017).

At the same time, the technology described enables a record to be captured of the professional staff members’ work. Having such a record may help persons in need of care receive care that is better attuned to their individual needs (Jaume-Palasí and Spielkamp 2017).

In addition to these direct effects, the use of algorithms also offers the opportunity to improve the system’s performance on an ongoing basis. Along with the individually collected data, additional clinical data can be incorporated into the system. Ideally, the services provided will be continuously reviewed and improved (Rantz et. al. 2013).

Individual-ethics challenges

From the individual-ethics perspective, the first question is one of how likely it is that the opportunities noted above will in fact be realized in practice. Contrary to the assumption that the use of AALs brings a better quality of life, for example, is the objection that the further development of sensor- and algorithm-supported systems might result in a problematic decline in the frequency of human contact in medical-care and nursing contexts. In this case, instead of the hoped-for improvement, a de facto deterioration in the quality of life of elderly people and those in need of care would occur (Friesacher 2010).

It is furthermore questionable whether the gains in security compensate for the associated losses of personal privacy and possible data-security risks. From an ethical point of view, continuous monitoring through the use of barely perceptible sensors, the creation of a highly detailed individual profile, and the automatic transfer of data analyses to specialized staffers or other persons (sic!) raise problems at every step. The combination of these factors could significantly exacerbate the challenges posed. From an individual-ethics perspective, one relevant question is whether affected individuals are sufficiently informed about the associated encroachments on their privacy. Similarly, the degree to which they have the ability to decline the various offerings, and the question of whether their interests can be persistently protected over the course of further developments, are also critical issues.

Personal privacy Loss of human contact
Informed consent
Invisibility of technology Digital skills
Data-security risks

Particular difficulties also arise from the fact that the individuals concerned will need a relatively significant amount of health knowledge and digital literacy in order to understand the mechanisms underlying the monitoring system (Glenn and Monteith 2014; Lepri et al. 2017). Comprehending new technologies is very difficult for many people. This is particularly true for people with cognitive impairments, for instance due to dementia-related conditions. This raises complex ethical questions with regard to the need for the individuals in question to be able to give their informed consent to the installation of the AAL technologies (Novitzky et al. 2015).

In addition, the integration of sensors into everyday objects could over the long term lead to a loss of awareness of the fact that data is being collected and evaluated by algorithms on an ongoing basis (Yuste et al. 2017). In a certain sense, the technology and its functions become invisible, merging into the everyday environment (Monteith and Glenn 2016). The capacity to make situational decisions regarding the degree of desired privacy can hereby be significantly impaired. Such problems could potentially be addressed by designing the various functions in such a way that they can be individually activated or deactivated. However, individual control options of this kind would again require significant competencies on the part of the user. In order to be protected against possible harm, the user must be able to understand the possible consequences of deactivating the sensors. This capacity cannot always be presupposed. Thus, the question remains of who should perform needed or desired system adjustments if the individual concerned lacks sufficient competence for this task.

Institutional-ethics challenges

For healthcare institutions, the establishment of algorithm-supported AALs poses the already-noted challenge of determining who is to be responsible for harm resulting from system failures or incorrectly system use (Hofmann 2013). For example, who is liable if a person in need of care is not helped, or receives assistance too late due to a system malfunction or mistaken analysis (Friesacher 2010)?

Work description Institutional responsibility
Training of skilled workers
Liability issues Misuse

The possibility that the data collected and analyzed might be misused furthermore raises questions of liability. For example, information about the routine activities of elderly people could fall into the hands of criminals. Burglars could use the movement-data information in order to rob homes in the absence of their residents. Con artists could claim to be reacting to a system alarm in order to gain entrance to the residence. Older people, particularly when they have cognitive impairments, are particularly prone of falling victim to scammers. Individual-level solutions would be unlikely to successfully address challenges relating to the security of people using algorithm-supported AALs. To tackle these risks appropriately, institutional measures will also be necessary.

Institutions are also faced with the challenge of meeting new demands on the occupational profile of medical and nursing professionals. The introduction of digitalization into caregiving contexts is significantly changing job descriptions, and generating new requirements with regard to skills and competences. Nursing professionals have often chosen their careers with the aim of doing “something with people.” New requirements, particularly those related to the operation of technical systems, could wind up being diametrically opposed to the original motivation for entering the chosen profession (Friesacher 2010). This raises the question of what institutional measures can and should be taken in order to support skilled workers in their professions so that individuals needing care are optimally served without any neglect of the legitimate interests of the people providing that care.

In order to avoid a loss of jobs and a decline in quality due to digitalization, it will be necessary to train staff to meet the new requirements. As training institutions have a responsibility to prepare their students properly for the labor market, such entities would thus need to adapt existing curricula to the new professional demands.

Sociopolitical challenges

As previously noted, the use of the systems presented here would require accepting a loss of privacy. The prospect of living longer and more securely in one's own home could usher in a societal change with regard to the value accorded to privacy. Given the expected quality-of-life improvements, data privacy with regard to certain aspects of private and professional life could come to be increasingly less valued. Such a development would ultimately be a societal decision, in the sense of "voting with the feet". That is, the more that people are willing to equip their houses with sensors and let their individual activity patterns and vital signs be analyzed by algorithms, the more likely the concept of privacy as an aspect of data protection would diminish in importance.

This kind of societal development may be comprehensible and legitimate in itself. However, it would be problematic if, as a result, those who value their personal privacy were required to give it up. If this were to happen, one of the original core goals of algorithm-supported AALs – that of increasing the capacity for self-determination – would effectively be turned on its head. In order to avoid such a paradox, a fully informed societal debate must be initiated regarding the possible short, medium and long-term consequences of omnipresent sensors and continuously analyzed activity patterns and personal profiles.

Change in societal values

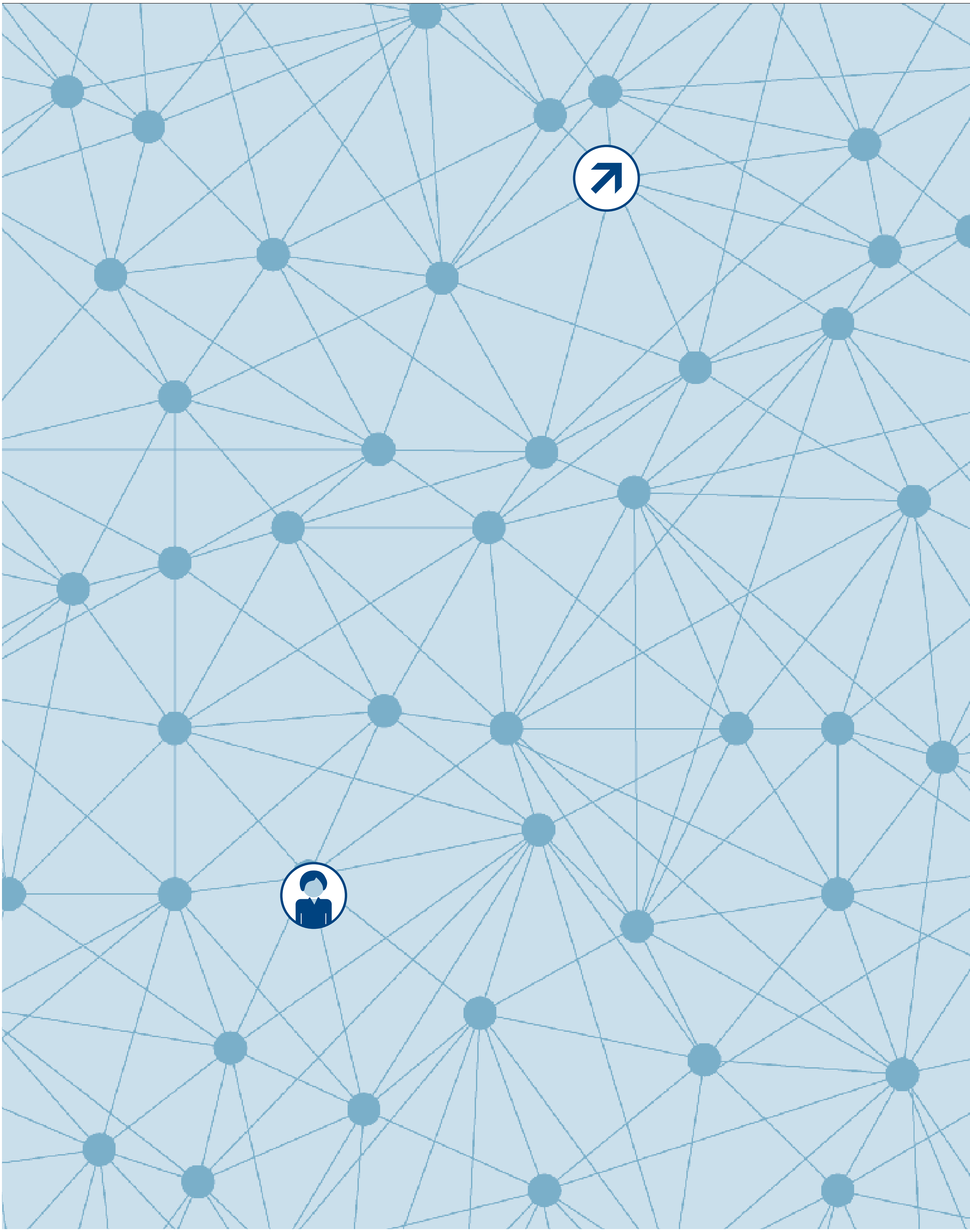
Personal privacy vs. security

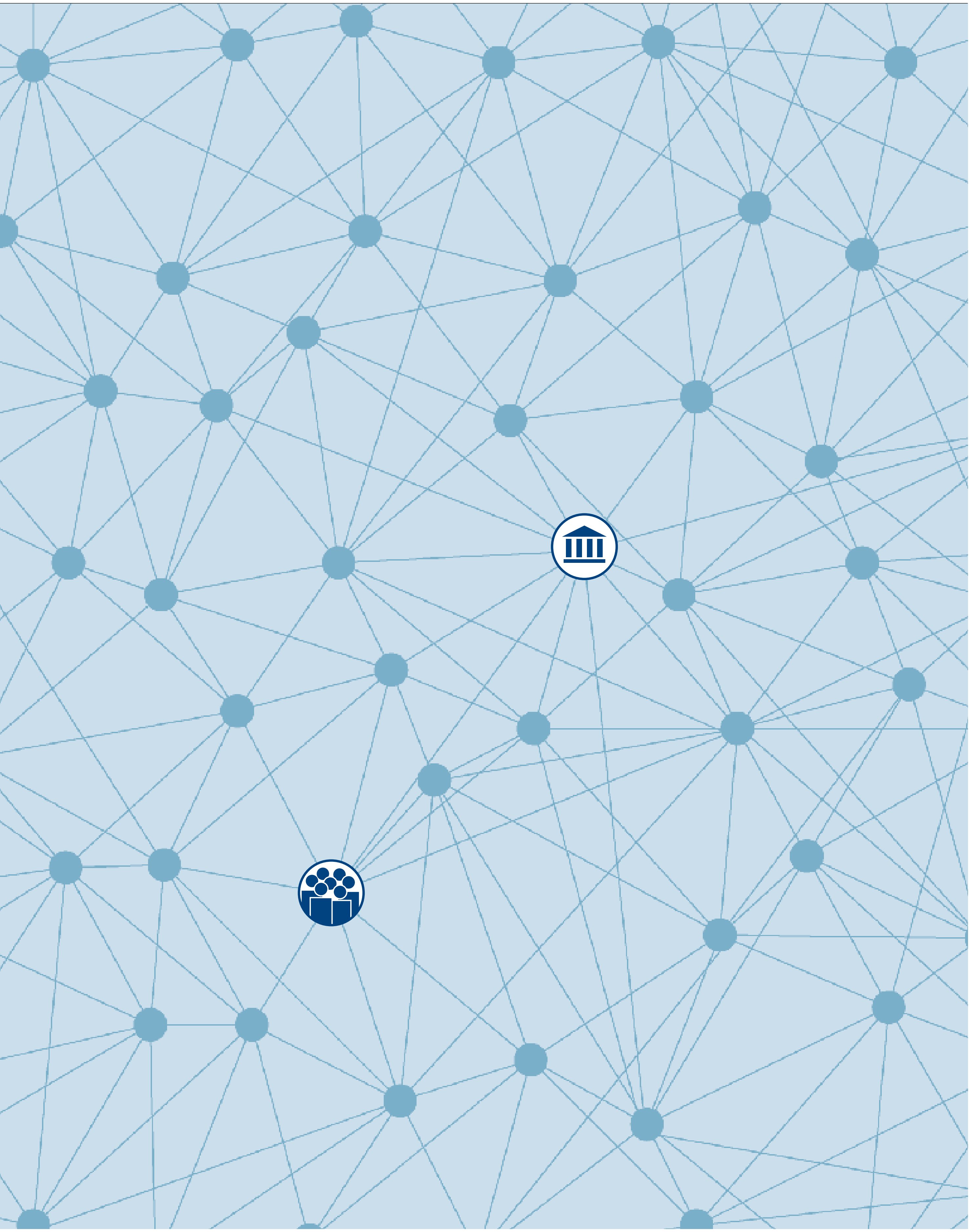
Fair access Healthcare system based on solidarity

Replacement of human contact and care

Additional societal challenges relate to the question of who, and under what conditions, should have access to the various digital services. The questions raised above relating to fair access are also relevant here. Due to demographic change, there will presumably be an increasingly large number of people in need of direct and preventative health care. Many will also be interested in living in their own homes for as long as possible. At least for the foreseeable future, it is unlikely that digital services can be made available for everyone who might be interested (Friesacher 2010) From a sociopolitical perspective, this poses the challenge of deciding what criteria will be used to inform decisions regarding the provision of care. Should the allocation of such services simply follow market principles, costly systems might be used only by well-off individuals, at least for the foreseeable future.

However, as the sector develops further, this circumstance could be reversed – that is, over the long term, algorithm-assisted care in AALs may ultimately prove to be less costly than care provided by human nursing staff and clinicians. Critics warn that AALs could increasingly replace human contact and care (Hofmann 2013). Either way, society faces the challenge of deciding who, under what conditions, should have access to the resources in question (care by humans or a machine-based system). In both cases, allocation solely on the basis of free-market principles would be difficult to reconcile with the principles of social solidarity underlying the healthcare system.





7 Outlook for the future

The algorithms used in healthcare settings are being continuously improved. Moreover, it can be assumed that this development will continue to accelerate. In the future, algorithms are expected to make an important contribution to the further development of personalized or precision medicine, which could entail a number of specific new applications.

With the help of powerful machine-learning algorithms, for example, it will become possible to sequence and analyze the human genome in a cost-effective and rapid way. This in turn will enable treatments that are significantly more targeted and individually tailored. In addition, the study of very large samples will become possible. The examination of patients at all levels in order to create an integrated picture of all processes across all biological and even non-biological levels – from the genome, epigenome and proteome to the organelles, and even to the behavior and biomechanics of the organism as a whole – is delivering wholly new knowledge regarding the diagnosis and treatment of diseases (Deutscher Ethikrat 2018). It will become possible to analyze the contribution of a very large number of variables to the emergence of specific diseases (Wired 2017). For example, it might become possible to render diagnoses and prognoses simply by analyzing the breath alone (De Witte 2017). In this way, cancers such as lung cancer could be detected more quickly, and potentially also treated earlier.

As the healthcare sector becomes increasingly networked, an ever-increasing body of information will be available to physicians for the purposes of diagnosis and treatment. It can be assumed that efficient algorithms will be directly connected to the hospital information systems of the future, directly providing physicians with diagnoses for individual patients (Zukunftsinstitut 2012). Algorithms will not only produce differential diagnoses, but also suggest diagnostic tests, while simultaneously significantly reducing the quantity of superfluous tests (Obermeyer and Emanuel 2016).

In addition, algorithms are enabling a number of areas of work to be outsourced to external providers. Numerous tasks that today are carried out in hospitals by multi-person teams will in the future be able to be carried out quickly and efficiently by vendors around the world. This includes tasks such as laboratory analyses and pathology work (ibid.). For example, in a diagnostic context, algorithms break down an X-ray into millions of individual variables. Other data, for example from insurers, can also be incorporated into the diagnostic process (ibid.). Clinics can reallocate the resources saved in this way to other innovations or use them for improvements in patient care (Zukunftsinstitut 2012).

Today, algorithms in expert systems often do no more than apply rules to data, for example when identifying interactions between medications (Bublak 2016). However, machine-learning algorithms have the capacity to derive rules from data. As a consequence, it can be expected that machines will surpass human experts with regard to diagnostic accuracy in the near future. Thus, the degree to which the future work of radiologists or anatomical pathologists will be taken over by machines, at least to a significant extent, remains unclear (ibid.). In addition, high-performance algorithms will in the future increasingly identify psychosocial factors relevant to the prevention, development and progression of diseases, while additionally discerning new epidemiological relationships (Langkafel 2015).

Machine-learning techniques will help improve prognostic capacities, for example regarding cancer patients' life expectancies (Obermeyer and Emanuel 2016). Algorithms will enable the use of thousands of predictor variables that will be drawn directly from electronic health records or insurers' claims databases. In contrast, today's prognosis models are limited to just a few variables (ibid.). The incorporation of a larger number of variables will facilitate significantly better care planning for patients with severe illnesses, for example (ibid.).

In the future, algorithms may additionally provide increasingly accurate forecasts of when a patient with a specific condition will die. This could lead to recommendations regarding appropriate end-of-life treatments, but – more problematically – could also be used for decisions regarding the allocation of scarce resources (Lobe 2017).

The use of algorithms promises an increase in patient safety; unlike humans, algorithms do not make mistakes if they have been programmed or trained appropriately from both the technical and ethical perspectives. They don't need to sleep, and are functional at all hours of the day and night (Obermeyer and Emanuel 2016). They can simulate the effect of drugs using a patient's available vital parameters, thereby avoiding ineffective medications. In addition, potential drug interactions or side effects can be calculated in advance using algorithms. Experts predict that this will help significantly increase treatment success rates, while also reducing costs (Zukunftsinstitut 2012).

Algorithms are also expected to be used for monitoring purposes with increasing frequency. Algorithms connected to a hospital's information system could monitor every patient in a clinic around the clock. Upon registering deviations in a given patient's measured vital signs, direct countermeasures could be taken, such as automatically increasing the dosage of a medicine from an intravenous drip dispenser, or informing the on-duty staff members. Thus, it is likely that algorithms will not only recommend actions in the future, but will also trigger them independently, for example in intensive-care settings (Langkafel 2015). They are expected to take on numerous anesthesia- and intensive-care-related tasks (Obermeyer and Emanuel 2016).

Algorithms are also increasingly widely used in the context of telemedicine. Camera images (e.g., a person's overall appearance or a patient's changed complexion) can be directly evaluated so as to generate recommendations for further treatment.

Algorithm-supported monitoring systems will also become more common outside of medical facilities. In the future, increasing numbers of people will integrate such systems into their "smart home" applications – whether to monitor existing conditions or in hopes of preventing the development of new illnesses. Algorithms will then calculate a person's or patient's optimal daily routine; for example, they might remind the individual to take medications, or autonomously make an appointment with a physician (Zukunftsinstitut 2012). Generally speaking, algorithms will be able to help integrate and evaluate increasing amounts of data from a person's everyday living environment (individual health-analytics approach). This may provide a better understanding of and improve the predictability of disease risks.

The analysis of this data – perhaps in conjunction with other information from the social or occupational environment – may ultimately facilitate faster and more targeted forms of disease and epidemic prevention (Langkafel 2015). Machine-learning algorithms will detect patterns in collected environmental and vital-sign data and link them with critical health events (Wired 2017). This will enable diseases posing a threat to be detected at an early stage, enabling swift intervention. Monitoring of this kind can also be used in cases involving mental illnesses such as depression (ibid.). This could help reduce the suicide rate, among other benefits.

The trend toward self-tracking will continue to grow in this context. In the future, many people will collect a large amount of health-related data about themselves as a matter of course. Algorithm-supported smartphone apps will then immediately process the collected data, and offer recommendations on topics such as diet or time for fitness. In addition, using previous behavior as a basis, they might calculate the optimal strategy for how that specific person could remain motivated to reach a particular goal, such as a desired ideal weight. A wide range of possible uses are conceivable here.

In addition to the fields of application already described, intelligent software will use algorithms to evaluate the large amounts of information produced by scientific publications more quickly, and in a much more targeted manner. The information relevant to a diagnosis and therapy will be filtered out and directly integrated into the treatment process (Langkafel 2015). In this regard, algorithms have the potential to reverse the usual path taken in scientific work; by automatically examining large amounts of data for relationships, they have the capacity to discover correlations that were not the specific subject of the search. In so doing, the path no longer runs from a previously formulated hypotheses to the search for and analysis of data. Rather, it runs from the data to the formulation of a new hypothesis.

Overall, algorithms and well-prepared data sets will in the future help us understand the development of diseases with increasing clarity, and positively influence the aging process. Disease treatment will thus increasingly give way to disease prevention, or to new ways of improving various conditions' progression (De Witte 2017).

8 To-dos and research needs

If the opportunities promised by algorithms in the healthcare sector are to be realized in practice, and the associated challenges addressed appropriately, a number of actions at a variety of levels will be necessary.

Medical professionals such as physicians and caregivers must be comprehensively prepared for the technical innovations, a process that must in particular ensure that they are trained in the necessary technical skills. The specialist personnel must understand how algorithms function, and how the results of algorithm-based decision-support systems should be interpreted. In the design of predictive algorithms, and particularly in the case of decision-support systems, it will often be necessary to deal with various risk assessments, such as the probability of contracting a disease. Working with risk calculations is as challenging for experts as it is for laypeople. For the systems to be able to provide meaningful support, medical professionals must be able to interpret the automatically generated results correctly. Moreover, they must be able to convey this interpretation to patients properly, and in an understandable way.

Various approaches for automated decision support are currently being discussed in the literature. A distinction has been made between "opt-out" and "opt-in" models, with another option being systems that automatically point out recommended decision paths (Cohen et al. 2014). The advantages and disadvantages of these models each must be analyzed and discussed further. In this regard, it is also necessary to clarify how "incidental findings" will be handled – that is, findings that are not related to the original research question, but which are nonetheless of significance for the health of the person being examined and their relatives.

Those working in the healthcare sector must develop skills in the practical use of the new technologies; however, the capability to reflect critically on their use must also be promoted. All people involved in their development, as well as those affected by their use, should understand the opportunities, risks, strengths and weaknesses of the various technical options within each of the settings where such techniques may be deployed. Before the launch of any informational campaigns, a survey should be conducted regarding the attitude of the population (e.g., the general population, as well as patient groups, physicians, nurses, and healthcare professionals within other areas such as insurers) toward the use of algorithms in medical care and other health-related areas. This would allow the fears and concerns of individual groups to be better addressed.

A comprehensive discussion of the new technologies' advantages and disadvantages may well lead to a greater degree of technical acceptance. Such a conversation would allow fears and prejudices to be reduced, and it could be clearly communicated that there is no intention of replacing medical professionals with computers. This is not a conflict pitting humans against machines. Instead, algorithms in healthcare can be important and useful tools that make human professionals' work significantly easier. The goal is one of meaningful cooperation between humans and machines. To this end, a multi-level governance approach could be adopted to help strengthen people's trust in the technology and convey the sense that it can be used safely. For example, a study by Wang et al. (2016) showed that use of a machine-learning algorithm to support pathologists in the diagnosis of metastasized breast cancer led to a reduction in the error rate from more than 3 percent to less than 1 percent.

From a technical point of view, the establishment of uniform standards for data collection and processing is critical. This would enable the various systems to be compatible with one another without requiring the data to undergo a fundamental, time-intensive pre-processing stage. For this reason, the underlying data to be used for algorithms' calculations, such as electronic health records, should be maintained carefully and be complete. It would also be useful to raise awareness among medical-institution employees with regard to the impact of missing or incorrectly entered data. Ensuring system interoperability through data-exchange standards is additionally important in order to avoid making users dependent on certain providers (Hahn and Schreiber 2018: 340).

With regard to the development of algorithms, clear rules of conduct should be established for the programming process. For example, programmers might make a (voluntary) commitment to be as transparent as possible regarding how their algorithms function, and with respect to what data is to be captured and processed. However, it will also be necessary to explain how potential problems created by industrial espionage or hacker attacks could be resolved. To be sure, complete transparency in the medical-products industry will be difficult to realize. The consideration of ethical issues should be incorporated into design and development processes from the very beginning, for example through the use of interdisciplinary teams.

With regard to specific applications of algorithms in the healthcare sector, a number of additional actions are needed. In order to be able to use machine-learning algorithms for diagnostic purposes, uniform standards for the diagnosis of certain disorders must be established. Currently, as in the cases of sepsis or rheumatoid arthritis, these standards are often unclear (Obermeyer and Emanuel 2016). This makes it much more difficult to train an algorithm for use in a practical setting. In addition, ways to prevent the appearance of spurious correlations should be found. The greater the amount of data involved, the greater the risk of such spurious correlations. In the case of large, quite heterogeneous data sets, they can emerge simply by chance (Fasel and Meyer 2016: 9). To minimize this risk, special techniques such as machine-learning algorithms optimized for a particular system must be used (ibid.).

In addition, distortions in the underlying data must be identified and eliminated. Finally, the use of algorithms in healthcare should take place in an evidence-based manner. That is, any support provided by algorithm-based systems should be based on proof that such techniques are both effective and efficient. In this regard, the use of algorithms should also be evaluated from a health-economics perspective, with the aim of creating a basis for the fair allocation of resources.

As algorithms are put into practical use, for instance in the clinical context, more must be done with regard to ensuring informed patient consent and the protection of patients' personal privacy. Patient disclosure and consent processes must in the future include

Recommended actions

In order to realize the practical opportunities offered by the use of algorithms in healthcare, while appropriately addressing the associated challenges, a number of actions must be taken on various levels.

Fulfilling technical prerequisites

- Establish uniform standards for the collection, processing and exchange of data
- Create compatible systems
- Maintain underlying data (e.g., electronic health records) conscientiously, while ensuring it is as complete as possible

Strengthening cooperation between humans and machines

- Promote technical skills, for instance through training programs
- Promote critical reflections on the technology: Algorithms as a useful tool, conflict between humans and machines
- Conduct informational campaign focusing on new technologies, with aim of diminishing fears and prejudices
- Sensitize medical staff to impact of incorrectly entered data
- Create multi-level governance structures with aim of strengthening trust and security

Ensuring informed consent

- Explain use of algorithms as part of patient disclosure and consent process
- Conduct research on additional opportunities and risks, and explain these accordingly
- Develop new, appropriate explanatory materials
- Clarify issues relating to protection of the right to informational self-determination
- Clarify requirement that algorithms must function in a transparent manner

Ensuring algorithms are programmed in an ethical manner

- Use interdisciplinary teams to plan and develop algorithms, taking possible risks and consequences into account
- Define clear rules of conduct for programming algorithms

Using algorithms reliably and safely

- Establish unified standards, for instance for the diagnosis of specific diseases
- Work to prevent appearance/discovery of spurious correlations
- Interpret and convey the results of decision-support systems (probabilities/risks) correctly
- Develop rules for handling incidental findings
- Clarify legal compatibility of various types of algorithms and applications with the European General Data Protection Regulation, which has been in effect since May 2018

Creating networked structures

- Establish integrated care structures in clinical practices (e.g., closer interconnections between outpatient and inpatient treatment)
- Promote cooperation between research institutions
- Promote interactions between different disciplines (e.g., computer science and ethics)
- Initiate interdisciplinary projects tasked with developing and evaluating new algorithm-assisted systems, and with translating them into innovations in care
- Initiate an interdisciplinary discussion around the social consequences of current developments, helping to shape a public dialogue

Developing an ethics of algorithms and a means of algorithm oversight

- Create appropriate structures for reviewing healthcare-sector algorithms by type and purpose of application, if necessary with the aim of certifying them and providing oversight
- Focus on clarifying the ethical rules that should guide the development and behavior of machine-learning algorithms
- Monitor deep learning algorithms for discriminatory effects

information on what algorithms are being employed as a part of the diagnostic and therapeutic measures, and what risks and opportunities are entailed by this use. At present, neither risks nor opportunities are sufficiently explained; nor have adequate approaches been developed for the creation or use of explanatory materials (Cohen et al. 2014). Given the new possibilities for processing personal data, the specific practical requirements for obtaining informed consent must be clarified. Naturally, this must go beyond a simple formal disclosure or explanation (ibid.).

If a wealth of data from different people is to be processed using algorithms, this poses fundamental issues with regard to protecting the right to informational self-determination, both at the individual level and beyond. Various authors have already called for additional research on this issue: "Further work is required to describe how privacy operates at group level" (Mittelstadt et al. 2016).

From a legal point of view, there is a need for clarification regarding the degree to which various kinds of algorithms and application areas (monitoring, data matching, generation of working medical hypotheses, etc.) are compatible with the European General Data Protection Regulation, which has been in effect since May 2018.

At the institutional level, additional integrated care structures should also be created to support the incorporation of algorithms into clinical practice. For example, this should include deeper interconnections between outpatient, inpatient and rehabilitative treatment. In addition, research institutions must also be more closely networked in order to promote interactions between different disciplines such as clinical medicine, epidemiology, computer science and ethics. This would facilitate the creation of interdisciplinary projects working to develop and evaluate new algorithm-assisted systems, and to translate them into innovations in care.

Overall, an interdisciplinary societal discussion is necessary regarding the probable social consequences associated with current trends and the appropriate response to them. For example, this might deal with issues of responsibility; the need for national, regional and local regulations (and governance); the need to update the curricula used in medical and engineering courses. In addition, a public dialogue should be initiated regarding the opportunities and risks associated with the use of algorithms in healthcare more generally. This must go beyond the simple provision of information to laypeople by experts in the scientific, ethics, legal and sociopolitical fields. As a discussion, it should also include all involved stakeholders.

Overall, there is a growing need to deal with the questions posed by an independent "ethics of algorithms," as well as with the issue of algorithm oversight. Instead of focusing on human action, an ethics of this kind – in the sense of a subdiscipline of machine ethics – would be oriented toward ethical rules intended to apply to all machine-learning systems and processes. In particular, it would be a matter of determining – and then programming – rules through which moral norms could be recognized and ethical dilemmas resolved. Initial approaches to this problem have already been developed within the philosophical community (e.g., Anderson et al. 2004; Mitchell et al. 2009). For application in the medical domain, they would have to be further developed and concretized, and if necessary integrated into an institutionalized system of algorithm oversight.

Bibliography

Abouelmehdi, K., Beni-Hssane, A., Khaloufi, H., and Saadi, M. (2017). Big data security and privacy in healthcare: A review. *Procedia Computer Science,* 113 (2019), 73–80.

Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-Jackson, S., Aisen, M., Heck, C., Liu, C., and Andersen, R. A. (2015). Neurophysiology. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science,* 348 (6237), 906–910.

Amarasingham, R., Audet, A.-M. J., Bates, D. W., Cohen, G., Entwistle, M., Escobar, G. J., Liu, V., Etheredge, L., Lo, B., Ohno-Machado, L., Ram, S., Saria, S., Schilling, L. M., Shah, A., Stewart, W. F., Steyerberg, E. W., and Xie, B. (2016). Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges. *eGEMs (Generating Evidence & Methods to improve patient outcomes),* 4 (1), 3.

Anderson, M., Anderson, S. L., and Armen, C. (2004). Towards Machine Ethics. AAAI-04 Workshop on Agent Organizations: Theory and Practice.

Antes, G. (2016). Ist das Zeitalter der Kausalität vorbei? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen,* 112 (1), 16–22.

Anzaldi, L. J., Davison, A., Boyd, C. M., Leff, B., and Kharrazi, H. (2017). Comparing clinician descriptions of frailty and geriatric syndromes using electronic health records: a retrospective cohort study. *BMC geriatrics,* 17 (1), 248.

Arbeitskreis Versorgungsforschung beim wissenschaftlichen Beirat der Bundesärztekammer (2004). Definition und Abgrenzung der Versorgungsforschung. www.definitionen.versorgungsforschung.net/.

Attiah, M. A., and Farah, M. J. (2014). Minds, motherboards, and money: futurism and realism in the neuroethics of BCI technologies. *Frontiers in Systems Neuroscience,* 8 (86), 1–3.

Bakker, E., Tian, K., Mutti, L., Demonacos, C., Schwartz, J.-M., and Krstic-Demonacos, M. (2017). Insight into glucocorticoid receptor signalling through interactome model analysis. *PLoS Computational Biology,* 13 (11), e1005825.

Banerjee, S. (2017). Automated interpretable computational biology in the clinic. A framework to predict disease severity and stratify patients from clinical data. *Interdisciplinary Description of Complex Systems,* 15 (3), 199–208.

Bauer, M., Glenn, T., Monteith, S., Bauer, R., Whybrow, P. C., and Geddes, J. (2017). Ethical perspectives on recommending digital technology for patients with mental illness. *International journal of bipolar disorders,* 5 (1), 6.

Beauchamp, T. L., and Childress, J. F. (1977). *Principles of Biomedical Ethics.* 6th edition. Oxford 2008.

Beck, M. (2016). Can a death-predicting algorithm improve care? *Wall Street Journal.* www.wsj.com/articles/can-a-death-predicting-algorithm-improve-care-1480702261.

Becker, A. S., Blüthgen, C., Mühlematter, U., and Boss, A. (2018). Medicina ex Machina: Machine Learning in der Medizin. *Praxis,* 107 (1), 19–23.

Becker, K., and Strammer, Y. (2016). Sensorbasierte Gesundheitsservices für mehr Fitness im Alltag. In S. Müller-Mieltz and T. Lux (Eds.): *E-Health-Ökonomie,* Wiesbaden, 501–516.

Bonderman, D. (2017). Artificial intelligence in cardiology. *Wiener klinische Wochenschrift,* 129 (23-24), 866–868.

Bouton, C. E., Shaikhouni, A., Annetta, N. V., Bockbrader, M. A., Friedenberg, D. A., Nielson, D. M., Sharma, G., Sederberg, P. B., Glenn, B. C., Mysiw, W. J., Morgan, A. J., Deogaonkar, M., and Rezai, A. R. (2016). Restoring cortical control of functional movement in a human with quadriplegia. *Nature,* 533 (7602), 247–250.

Brady, E. S., Leider, J. P., Resnick, B. A., Alfonso, Y. N., and Bishai, D. (2017). Machine-Learning Algorithms to Code Public Health Spending Accounts. *Public health reports,* 132 (3), 350–356.

Brandl, P., Aschbacher, H., and Hösch, S. (2015). Mobiles Wissensmanagement in der Industrie 4.0. In A. Weisbecker, M. Burmester and A. Schmidt (Eds.): *Mensch und Computer,* 225–232.

Braun, A., Kirchbuchner, F., and Wichert, R. (2016). Ambient Assisted Living. In F. Fischer and A. Krämer (Hrsg.): *eHealth in Deutschland.* Berlin, Heidelberg, 203–222.

Braun, V. (2018a). Diagnostik. Pschyrembel Online. www.pschyrembel.de/diagnostik/K05UB/doc/.

Braun, V. (2018b). Prognose. Pschyrembel Online. www.pschyrembel.de/Prognose/K0HR4/doc/.

Brock, D. (2003). Ethik und Altersrationierung in der Medizin: Ein konsequentialistischer Standpunkt. In G. Marckmann (Ed.): *Gesundheitsversorgung im Alter.* Stuttgart, 89–115.

Broy, M., and Spaniol, O. (1999). VDI-Lexikon Informatik und Kommunikationstechnik. VDI-Lexikon Informatik und Kommunikationstechnik. Berlin, Heidelberg.

Bruce, M. L., and Hoff, R. A. (1994). Social and physical health risk factors for first-onset major depressive disorder in a community sample. *Social psychiatry and psychiatric epidemiology,* 29 (4), 165–171.

Bublak, R. (2016). Macht Big Data Mediziner arbeitslos? *ÄrzteZeitung* 31.10.2016. www.aerztezeitung.de/praxis_wirtschaft/e-health/article/922490/zukunft-medizin-computer-besseren-aerzte.html.

Büttner, R. (2016). Effizientes Recruiting von Fachkräften im Web 2.0 (Efficient Recruiting 2.0): Hoch-automatisierte Identifikation und Rekrutierung von Fachkräften durch Analyse internetbasierter sozialer Netzwerke mittels intelligenter Softwareagenten. Abschlussbericht zum BMBF-Forschungsprojekt. Munich. www.prof-buettner.com/downloads/buettner2017b.pdf.

Bundesministerium für Bildung und Forschung (2018). Mathematik fürs Ohr. *Newsletter* 89. www.gesundheitsforschung-bmbf.de/de/mathematik-furs-ohr-7827.php.

Burke, R., Felfernig, A., and Göker, M. H. (2011). Recommender Systems: An Overview. *AI Magazine* (32) 3, 13–18.

Burkhardt, M. (2017). Vorüberlegungen zu einer Kritik der Algorithmen an der Grenze von Wissen und Nichtwissen. In A. Friedrich, P. Gehring, C. Hubig, A. Kaminski and A. Nordmann (Eds.): *Technisches Nichtwissen,* 55–68.

Burwell, S. Sample, M., and Racine, E. (2017). Ethical aspects of brain computer interfaces: a scoping review. *BMC Medical Ethics,* 18, 60.

Bylsma, L. M., Morris, B. H., and Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical psychology review,* 28 (4), 676–691.

Carruthers, H. R., Morris, J., Tarrier, N., and Whorwell, P. J. (2010). The Manchester Color Wheel: development of a novel way of identifying color choice and its validation in healthy, anxious and depressed individuals. *BMC medical research methodology,* 10, 12.

Chakraborty, C. (2017). Chronic Wound Image Analysis by Particle Swarm Optimization Technique for Tele-Wound Network. *Wireless Personal Communications,* 96 (3), 3655–3671.

Chandrasekhar, S., Laxminarayana, G., and Chakrapani, Y. (2017). Novel Hybrid Segmentation Techniques for Cardiac Image Processing in Remote Health Care Monitoring Systems. *Journal of Medical Imaging and Health Informatics,* 7 (6), 1153–1159.

Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., and Lo, B. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health affairs (Project Hope),* 33 (7), 1139–1147.

Conway, M., and O'Connor, D. (2016). Social Media, Big Data, and Mental Health: Current Advances and Ethical Implications. *Current opinion in psychology,* 9, 77–82.

Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., Bearden, C. E., Cecchi, G. A. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry,* 17 (1), 67-75.

Cornford, C. S., Hill, A., and Reilly, J. (2007). How patients with depressive symptoms view their condition: A qualitative study. *Family Practice,* 24 (4), 358–364.

DeBoer, E. M., Swiercz, W., Heltshe, S. L., Anthony, M. M., Szefler, P., Klein, R., Strain J., Brody, A. S., and Sagel, S. D. (2014). Automated CT scan scores of bronchiectasis and air trapping in cystic fibrosis. *CHEST,* 145 (3), 593–603.

Dente, C. J., Bradley, M., Schobel, S., Gaucher, B., Buchman, T., Kirk, A. D., and Elster, E. (2017). Towards precision medicine: Accurate predictive modeling of infectious complications in combat casualties. *Journal of Trauma and Acute Care Surgery,* 83 (4), 609–616.

Deutsche Gesellschaft für Humangenetik e.V. (2013). Stellungnahme der Deutschen Gesellschaft für Humangenetik zu genetischen Zusatzbefunden in Diagnostik und Forschung. www.gfhev.de/de/leitlinien/LL_und_Stellungnahmen/2013_05_28_Stellungnahme_zu_genetischen_Zufallsbefunden.pdf.

Deutsche Gesellschaft für Public Health e.V. (2018). Was ist Public Health? www.deutsche-gesellschaft-public-health.de/informationen/public-health/.

Deutscher Ethikrat (2018). Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung. Stellungnahme (30. November 2017), Berlin 2018.

Deutsches Referenzzentrum für Ethik in den Biowissenschaften (2018). Medizinische Forschung. www.drze.de/im-blickpunkt/medizinische-forschung-mit-minderjaehrigen/module/medizinische-forschung.

De Laat, P. B. (2017). Algorithmic Decision-Making based on machine learning from big data: Can transparency restore accountability? *Philosophy & Technology,* 1–17.

De Witte, B. (2017). Algorithms will outperform doctors in just 10 years time. *Dataconomy.* http://dataconomy.com/2017/10/18532/.

Divita, G., Carter, M., Redd, A., Zeng, Q., Gupta, K., Trautner, B., Samore, M., and Gundlapalli, A. (2015). Scaling-up NLP Pipelines to Process Large Corpora of Clinical Notes. *Methods of Information in Medicine,* 54 (6), 548–552.

Dörn, S. (2018). Digitale Technologien. In S. Dörn (Ed.): *eXamen.press. Programmieren für Ingenieure und Naturwissenschaftler.* Berlin, Heidelberg, 333–370.

Drazin, D., Spitler, K., Cekic, M., Patel, A., Hanna, G., Shirzadi, A., and Chu, R. (2013). Incidental finding of tumor while investigating subarachnoid hemorrhage: ethical considerations and practical strategies. *Science and engineering ethics,* 19 (3), 1107–1120.

Duden (n.d.). Rehabilitation. Duden online. www.duden.de/rechtschreibung/Rehabilitation

Eagleman, D. (2016). *The Brain. The Story of You.* New York.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature,* 542 (7639), 115–118.

Europarat (1950). Europäische Menschenrechtskonvention. EMRK. www.menschenrechtskonvention.eu/

European Group on Ethics in Science and New Technologies (2018). Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. Brussels.

Eyal, N. (2014). *Hooked: How to Build Habit-Forming Products.* New York.

Fasel, D., and Meier, A. (2016). Was versteht man unter Big Data und NoSQL? In D. Fasel and A. Meier (Eds.): *Edition HMD. Big Data: Grundlagen, Systeme und Nutzungspotenziale.* Wiesbaden, 3–16.

Fechter, T., Adebahr, S., Baltas, D., Ben Ayed, I., Desrosiers, C., and Dolz, J. (2017). Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. *Medical Physics,* 44 (12), 6341–6352.

Fischer, P., and Hofer, P. (2008). *Lexikon der Informatik.* Berlin, Heidelberg.

Fischer, T., Brothers, K. B., Erdmann, P., and Langanke, M. (2016). Clinical decision-making and secondary findings in systems medicine. *BMC medical ethics,* 17 (1), 32.

Fraunhofer Institut (2018). Algorithmen für die Leberchirurgie – weltweit sicherer operieren. www.fraunhofer.de/de/presse/presseinformationen/2018/Mai/algorithmen-fuer-die-leberchirurgie-weltweit-sicherer-operieren.html.

Friesacher, H. (2010). Pflege und Technik – eine kritische Analyse. *Pflege & Gesellschaft,* 293–367.

Gadebusch Bondio, M., Gordon, J. S., and Sporing, F. (Eds.) (2017). *Medical Ethics, Prediction, and Prognosis: Interdisciplinary Perspectives.* Oxford.

Gilbert, F., Cook, M., O'Brien, T., and Illes, J. (2017). Embodiment and Estrangement: Results from a First-in-Human "Intelligent BCI" Trial. *Science and engineering ethics,* 1–14. DOI: 10.1007/s11948-017-0001-5.

Gkoulalas-Divanis, A., Loukides, G., and Sun, J. (2014). Publishing data from electronic health records while preserving privacy: a survey of algorithms. *Journal of Biomedical Informatics,* 50, 4–19.

Glannon, W. (2014). Ethical issues with brain-computer interfaces. *Frontiers in Systems Neuroscience,* 8, 136.

Glenn, T., and Monteith, S. (2014). New measures of mental state and behavior based on data collected from sensors, smartphones, and the Internet. *Current psychiatry reports,* 16 (12), 523.

Göbel, E. (2017). *Unternehmensethik. Grundlagen und praktische Umsetzung.* 5th revised and updated edition. Constance.

Goldspiel, B. R., Flegel, W. A., DiPatrizio, G., Sissung, T., Adams, S. D., Penzak, S. R., Biesecker, L. G., Fleisher, T. A, Patel, J. J., Herion, D., Figg, W. D., Lertora, J. J. L., and McKeeby, J. W. (2014). Integrating pharmacogenetic information and clinical decision support into the electronic health record. *Journal of the American Medical Informatics Association,* 21 (3), 522–528.

Gombolay, M., Jessie Yang, X., Hayes, B., Seo, N., Liu, Z., Wadhwania, S., Yu, T., Shah, N., Golen, T., and Shah, J. (2016). Robotic Assistance in Coordination of Patient Care. *The International Journal of Robotics Research,* 0278364918778344.

González, G., Ash, S. Y., Vegas-Sánchez-Ferrero, G., Onieva, J., Rahaghi, F. N., Ross, J. C., Díaz, A., San José Estépar, R., and Washko, G. R. (2018). Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography. *American Journal of Respiratory and Critical Care Medicine,* 197 (2), 193–203.

Grant, K. A., Apffelstaedt, J. P., Wright, C. A., Myburgh, E., Pienaar, R., De Klerk, M., and Kotze, M. J. (2013). MammaPrint Pre-screen Algorithm (MPA) reduces chemotherapy in patients with early-stage breast cancer. *South African Medical Journal,* 103 (8), 522–526.

Gräßer, F., Beckert, S., Küster, D., Schmitt, J., Abraham, S., Malberg, H., and Zaunseder, S. (2017). Therapy Decision Support Based on Recommender System Methods. *Journal of Healthcare Engineering,* 2017, 8659460.

Guntuku, S. C., Yaden, D. B, Kern, M. L., Ungar, L. H., and Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavior Sciences,* 18, 43–49.

Gutmann, T., and Quante, M. (2017). Individual-, Sozial- und Institutionenethik. In I.-J. Werkner and K. Ebeling (Eds.): *Handbuch Friedensethik,* Wiesbaden, 105–114.

Haahr-Raunkjær, C., Meyhoff, C. S., Sørensen, H. B. D., Olsen, R. M., and Aasvang, E. K. (2017). Technological aided assessment of the acutely ill patient – The case of post-operative complications. *European Journal of Internal Medicine,* 45, 41–45.

Hahn, H., and Schreiber, A. (2018). E-Health. In R. Neugebauer (Ed.): *Digitalisierung.* Berlin, Heidelberg, 321–345.

Hanton, C., Kwon, Y.-J., Aung, T., Whittington, J., High, R. R., Goulding, E. H., Schenk, A. K., and Bonasera, S. J. (2017). Mobile Phone-Based Measures of Activity, Step Count, and Gait Speed: Results from a Study of Older Ambulatory Adults in a Naturalistic Setting. *JMIR mHealth and uHealth,* 5 (10), e104.

Hecker, D., Döbel, I., Rüping, S., and Schmitz, V. (2017). Künstliche Intelligenz und die Potenziale des maschinellen Lernens für die Industrie. *Wirtschaftsinformatik & Management,* 9 (5), 26–35.

Hildt, E. (2011). Brain-computer interaction and medical access to the brain: individual, social and ethical implications. *Science and Engineering Ethics,* 4, 3.

Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., Haddadin, S., Liu, J., Cash S. S., Smagt, P. v. d., and Donoghue, J. P. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature,* 485 (7398), 372–375.

Hofmann, B. (2013). Ethical challenges with welfare technology: A review of the literature. *Science and Engineering Ethics,* 19 (2), 389–406.

International Bioethics Committee (IBC) United Nations Educational, Scientific and Cultural Organization (UNESCO) (2017). Report of the IBC on Big Data and Health, Paris, 15. Sept. 2017. http://unesdoc.unesco.org/images/0024/002487/248724e.pdf.

Iqbal, E., Mallah, R., Rhodes, D., Wu, H., Romero, A., Chang, N., Dzahini, O., Pandey, C., Broadbent, M., Stewart, R., Dobson, R. J. B., and Ibrahim, Z. M. (2017). ADEPt, a semantically-enriched pipeline for extracting adverse drug events from free-text electronic health records. *PloS one,* 12 (11), e0187121.

Jaume-Palasí, L., and Spielkamp, M. (2017). Ethics and algorithmic processes for decision making and decision support. AlgorithmWatch Working Paper No. 2. Berlin.

Jebari, K. (2012). Brain Machine Interface and Human Enhancement – an ethical review (Appendix 2, ETHENTECH, Final report, August 2012). www.kth.se/polopoly_fs/1.333040!/Menu/general/column-content/attachment/Ethentech%20Appendix%2002.pdf.

Jernite, Y., Halpern, Y., Horng, S., and Sontag, D. (2013). Predicting Chief Complaints at Triage Time in the emergency department. NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare.

Jordan, J., Ippen, T., Helias, M., Kitayama, I., Sato, M., Igarashi, J., Diesmann, M., and 'Kunkel, S. (2018). Extremely Scalable Spiking Neuronal Network Simulation Code: From Laptops to Exascale Computers. *Frontiers in neuroinformatics,* 12, 2.

Kasperkevic, J. (2015). Google says sorry for racist auto-tag in photo app. *The Guardian* 1 July 2015. www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app.

Kim, S. J., Lim, G. J., Cho, J., and Côté, M. J. (2017). Drone-Aided Healthcare Services for Patients with Chronic Diseases in Rural Areas. *Journal of Intelligent & Robotic Systems,* 88 (1), 163–180.

Klein, E., and Jeffrey, O. (2016). Informed consent in implantable BCI research: Identification of research risks and recommendations for development of best practices. *Journal of Neural Engineering,* 13 (4), 43001–43011.

Kotchetkov, I. S., Hwang, B. Y., Appelboom, G., Kellner, C. P., and Connolly, E. S. Jr. (2010). Brain-computer interfaces: military, neurosurgical, and ethical perspective. *Neurosurgical Focus,* 28 (5), E25.

Kraemer, F., van Overveld, K., and Peterson, M. (2011). Is there an ethics of algorithms? *Ethics and Information Technology,* 13 (3), 251–260.

Krumholz, H. M. (2014). Big Data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. *Health Affairs,* 33 (7), 1163–1170.

Kruse, R., and Borgelt, C. (2015). *Computational Intelligence: Eine methodische Einführung in Künstliche Neuronale Netze, Evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze.* 2nd Edition. Wiesbaden.

Lämmel, U., and Cleve, J. (2012). *Künstliche Intelligenz: Mit 51 Tabellen, 43 Beispielen, 118 Aufgaben, 89 Kontrollfragen und Referatsthemen.* 4th updated edition. Munich.

Landon, B. E., Keating, N. L., Barnett, M. L., Onnela, J.-P., Paul, S., O'Malley, A. J., Keegan, T., and Christakis, N. A. (2012). Variation in patient-sharing networks of physicians across the United States. *JAMA,* 308 (3), 265–273.

Langkafel, P. (2015). Auf dem Weg zum Dr. Algorithmus? Potenziale von Big Data in der Medizin. *Aus Politik und Zeitgeschichte* 11–12, 27. www.bpb.de/apuz/202246/dr-algorithmus-big-data-in-der-medizin?p=all.

Leithead, A. (2013). Almost human: Lab treats trauma with virtual therapy. *BBC News,* Los Angeles 27 May 2013, www.bbc.co.uk/news/magazine-22630812.

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., and Vinck, P. (2017). Fair, Transparent, and Accountable Algorithmic Decision-making Processes. *Philosophy & Technology,* 84 (3), 488.

Lippert-Rasmussen, K. J. (2016). Is health profiling morally permissible? *Medical Ethics,* 42, 330.

Lipworth, W., Mason, P. H., Kerridge, I., and Ioannidis, J. P. A. (2017). Ethics and Epistemology in Big Data Research. *Journal of bioethical inquiry,* 14 (4), 489–500.

Lobe A. (2017). Müssen Roboter-Ärzte einen hippokratischen Eid ablegen? *Spektrum der Wissenschaft Online* 13 June 2017. www.spektrum.de/kolumne/muessen-roboter-aerzte-einen-hippokratischen-eid-ablegen/1463833.

Mainzer, K. (2016). *Künstliche Intelligenz – Wann übernehmen die Maschinen?* Berlin, Heidelberg.

Mantzaras, G., and Voudrias, E. A. (2017). An optimization model for collection, haul, transfer, treatment and disposal of infectious medical waste: Application to a Greek region. *Waste management,* 69, 518–534.

Marketsandmarkets.com (2016). Emotion Detection and Recognition Market by Technology (Bio-Sensor, NLP, Machine Learning), Software Tool (Facial Expression, Voice Recognition), Service, Application Area, End User, and Region – Global Forecast to 2021. www.marketsandmarkets.com/Market-Reports/emotion-detectionrecognition-market-23376176.html; www.marketsandmarkets.com/Market-Reports/emotion-detection-recognition-market-23376176.html.

Mattia, D., and Tamburrini, G. (2015). Ethical issues in brain–computer interface research and systems for motor control. *Handbook of Neuroethics.* Dordrecht. 725–740.

Mayer-Schönberger V., Cukier K., and Mallett, D. (2013). *Big Data. Die Revolution, die unser Leben verändern wird.* 2nd edition. Munich.

Miotto, R., Li, L., Kidd, B. A., and Dudley, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports,* 6, 26094.

Mitchell, A. J., Vaze, A., and Rao, S. (2009). Clinical diagnosis of depression in primary care: A meta-analysis. *The Lancet,* 374(9690), 609–619.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society,* 3 (2), 205395171667967.
Mittelstadt, B. D., and Floridi, L. (2016). The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics,* 22, 303–341.
Monteith, S., and Glenn T. (2016). Automated Decision-Making and Big Data: Concerns for People with Mental Illness. *Curr Psychiatry Rep,* 18:112.
Mortensen, E. L., Jensen, H. H., Sanders, S. A., and Reinisch, J. M. (2001). Better psychological functioning and higher social status may largely explain the apparent health benefits of wine: a study of wine and beer drinking in young Danish adults. *Archives of internal medicine,* 161 (15), 1844–1848.
Moyer, J. W. (2016). After computer hack, L.A. hospital pays $17,000 in bitcoin ransom to get back medical records. *The Washington Post* 18 February 2016. Accessed on 20 July 2018 atwww.washingtonpost.com/news/morning-mix/wp/2016/02/18/after-computer-hack-l-a-hospital-pays-17000-in-bitcoin-ransom-to-get-back-medical-records/?noredirect=on&utm_term=.706ec777deb4.
Myles, P. S., Bellomo, R., Corcoran, T., Forbes, A, Peyton, P., Story, D., Christophi, C., Leslie, K., McGuinness, S, Parke, R., Serpell, J., Chan, M. T. V., Painter, T., McCluskey, S., Minto, G., and Wallace, S. (2018). Restrictive versus Liberal Fluid Therapy for Major Abdominal Surgery. *The New England Journal of Medicine,* 378 (24), 2263–2274.
Nahrstedt, H. (Ed.) (2018). *Algorithmen für Ingenieure.* Wiesbaden.
Napoli, A., Glass, S. M., Tucker, C., and Obeid, I. (2017). The Automated Assessment of Postural Stability: Balance Detection Algorithm. *Annals of Biomedical Engineering,* 45 (12), 2784–2793.
Nicolas-Alonso, L. F., and Gomez-Gil, J. (2012). Brain Computer Interfaces, a Review. *Sensors,* 12 (2), 1211–1279.
Novitzky, P., Smeaton, A.F., Chen, C., Irving, K., Jacquemard, T., O'Brolcháin, F., O'Mathúna, D., and Gordijn, B. (2015). A review of contemporary work on the ethics of ambient assisted living technologies for people with dementia. *Science and Engineering Ethics,* 21 (3), 707–765.
Nürnberger, S., and Bugiel, S. (2016). Autonome Systeme. *Datenschutz und Datensicherheit,* 40 (8), 503–506.
O'Brolcháin, F., and Gordijn, B. (2015). Ethics of Brain–Computer Interfaces for Enhancement Purposes. In J. Clausen and N. Levy (Eds.): *Handbook of Neuroethics.* Dordrecht. 1207–1226.
Obermeyer, Z., and Emanuel, E. J. (2016). Predicting the Future – Big Data, Machine Learning, and Clinical Medicine. *The New England Journal of Medicine,* 375 (13), 1216–1219.
Pfannstiel, M. A., Krammer, S., and Swoboda, W. (Eds.) (2017). *Digitale Transformation von Dienstleistungen im Gesundheitswesen III. Impulse für die Pflegepraxis.* Wiesbaden.
Poplin, R., Varadarajan, A. V., Blumer, K., Liu, Y., McConnell, M. V., Corrado, G. S., Peng, L., and Webster, D. R. (2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering,* 2 (3), 158–164.
Porcalla, A., Barshteyn, N., Snyder, S., and Bhattacharya, M. (2017). An Innovative, Collaborative, and Strategic Approach to Proactively Evaluate and Update Drug Interactions Based on Prescribing Information of Newly Approved Medicinal Products. *Therapeutic Innovation & Regulatory Science,* 51 (6), 780–786.
Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., Watts, B., Flashman, L., and McAllister, T. (2014). Predicting the risk of suicide by analyzing the text of clinical notes. *PloS one,* 9 (1), e85733.
Powles, J., and Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology,* 7 (4), 351–367.

Pschyrembel W., and Dornblüth, O. (2002). Rehabilitation. *Pschyrembel Klinisches Wörterbuch.* 258th edition. Berlin.

Pschyrembel Online (2017) Prädiktion. www.pschyrembel.de/Pr%C3%A4diktion/H0J25.

Pschyrembel Online (2018a) Prävention. www.pschyrembel.de/Pr%C3%A4vention/K0HLD/doc/.

Pschyrembel Online (2018b). Therapie. www.pschyrembel.de/therapie/K0MEH/doc/.

Pschyrembel Online (2018c). Pflege. www.pschyrembel.de/pflege/K01BX/doc/.

Pschyrembel, W., Braun, R. N., and Engst, R. (Eds.) (2002). Rehabilitation. *Pschyrembel Klinisches Wörterbuch.* 259th edition. Berlin.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Ball, R.L., Langlotz, C., Shpanskaya, K., Lungren, M. P., and Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *arXiv,* 1711.05225v3.

Rantz, M. J., Scott, S. D., Miller, S. J., Skubic, M., Phillips, L., Alexander, G., Koopman, R. J., Musterman, K., and Back, J. (2013). Evaluation of health alerts from an early illness warning system in independent living. *CIN: Computers, Informatics, Nursing,* 31 (6), 274–280.

Rasche, C. (2017). Digitaler Gesundheitswettbewerb: Strategien, Geschäftsmodelle, Kompetenzanforderungen. In M. A. Pfannstiel, P. Da-Cruz and H. Mehlich (Eds.): *Digitale Transformation von Dienstleistungen im Gesundheitswesen I. Impulse für die Versorgung.* Wiesbaden, 1–30.

Reece, A. G., and Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science,* 6 (1), 157.

Rimscha, M. v. (2017). *Algorithmen kompakt und verständlich: Lösungsstrategien am Computer.* 4th revised edition. Wiesbaden.

Rüping, S. (2015). Big Data in Medizin und Gesundheitswesen [Big data in medicine and healthcare]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz,* 58 (8), 794–798.

Samaritans (2014). Samaritans Radar. www.samaritans.org/how-we-can-help-you/supporting-someone-online/samaritans-radar.

Schaar, P. (2016). Schutz der Privatsphäre im Gesundheitswesen. Zu den rechtlichen Grenzen der Überwachung in der Medizin. *Zeitschrift für medizinische Ethik,* 62, 31–39.

Scherer, S., Stratou, G., Gratch, J., and Morency, L.-P. (2013). Investigating Voice Quality as a Speaker-Independent Indicator of Depression and PTSD. *Interspeech,* 847–851.

Schmidt, M., Schmidt, S. A. J., Sandegaard, J. L., Ehrenstein, V., Pedersen, L., and Sørensen, H. T. (2015). The Danish National Patient Registry: a review of content, data quality, and research potential. *Clinical Epidemiology,* 7, 449–490.

Schubert, S., and Schwill, A. (2011). Didaktik der Informatik. In S. Schubert and A. Schwill (Eds.): *Didaktik der Informatik.* Heidelberg, 1–30.

Smart, A., and Smart, J. (2017). *Posthumanism: Anthropological insights. Anthropological insights.* Toronto.

Solymosi, A., and Grude, U. (2017). *Grundkurs Algorithmen und Datenstrukturen in JAVA: Eine Einführung in die praktische Informatik.* 6th updated and expanded edition. Wiesbaden.

Somanchi, S., Adhikari, S., Lin, A., Eneva, E., and Ghani, R. (2015). Early Prediction of Cardiac Arrest (Code Blue) using Electronic Medical Records. Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, 2119–2126.

Statista (2018). Anzahl der monatlich aktiven Facebook-Nutzer weltweit vom 3. Quartal 2008 bis zum 1. Quartal 2018. https://de.statista.com/statistik/daten/studie/37545/umfrage/anzahl-der-aktiven-nutzer-von-facebook/.

Stiftung Datenschutz (2017). *Neue Wege bei der Einwilligung im Datenschutz – technische, rechtliche und ökonomische Herausforderungen. Handlungsempfehlung.* Leipzig.

Stusser, R. J., and Dickey, R. A. (2013). Quality and cost improvement of healthcare via complementary measurement and diagnosis of patient general health outcome using electronic health record data: research rationale and design. *Journal of Medical Systems,* 37 (6), 9977.

Szucs, T. D., Puri, D., and Blank, P. R. (2014). The use of the EVITA algorithm for clinical assessment of novel agents used in prostate cancer, metastatic melanoma, and systemic lupus erythematosus. *European Journal of Clinical Pharmacology,* 70 (8), 983–990.

Tafelski, S., Nachtigall, I., Deja, M., Tamarkin, A., Trefzer, T., Halle, E., Wernecke, K. D., and Spies, C. (2010). Computer-assisted decision support for changing practice in severe sepsis and septic shock. *The Journal of International Medical Research,* 38 (5), 1605–1616.

Tartof, S. Y., Sy, L. S., Ackerson, B. K., Hechter, R. C., Haag, M., Slezak, J. M. Luo, Y., Fischetti, C. A., Takhar, H. S., Miao, Y., Solano, Z., Jacobsen, S., and Tseng, H.-F. (2017). Safety of Quadrivalent Meningococcal Conjugate Vaccine in Children 2-10 Years. *The Pediatric infectious disease journal,* 36 (11), 1087–1092.

TU Berlin (2017). Modulares Gangstudio NeuroReha “GSNeuro.” www.ige.tu-berlin.de/bemobil/forschung/teilprojekt_b.

Vallot, D., Caselles, O., Chaltiel, L., Fernandez, A., Gabiache, E., Dierickx, L., Zerdoud, S., Bauriaud, M., and Courbon, F. (2017). A clinical evaluation of the impact of the Bayesian penalized likelihood reconstruction algorithm on PET FDG metrics. *Nuclear Medicine Communications,* 38 (11), 979–984.

Van Solm, A., Hirdes, J. P, Eckel, L. A., Heckman, G. A., and Bigelow, P. L. (2017). Using standard clinical assessments for home care to identify vulnerable populations before, during, and after disasters. *Journal of emergency management,* 15 (6), 355–366.

Vollmann, J. (2000). *Aufklärung und Einwilligung in der Psychiatrie. Der Beitrag zur Ethik in der Medizin.* Darmstadt.

Wagner, B., and Vieth, K., on behalf of the Bertelsmann Stiftung (2017). *Teilhabe, ausgerechnet.* Gütersloh.

Wang, D., Khosla, A., Gargeya, R., Irshad, H., and Beck, A. H. (2016). Deep Learning for Identifying Metastatic Breast Cancer. arXiv preprint arXiv:1606.05718.

Weicker, K. (2015). *Evolutionäre Algorithmen.* 3rd revised and expanded edition. Wiesbaden.

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., Charlson, F. J., Norman, R. E., Flaxman, A. D., Johns, N., Burstein, R., Murray, C. J. L., and Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet,* 382 (9904), 1575–1586.

World Health Organization (WHO) (2013). Mental health action plan 2013-2020. Geneva. http://www.who.int/mental_health/publications/action_plan/en/.

Wilder, B., Suen, S.-C., and Tambe, M. (2018). *Preventing Infectious Disease in Dynamic Populations under uncertainty.* Association for the Advancement of Artificial Intelligences. www.cais.usc.edu/wp-content/uploads/2017/11/aaai_tb_final.pdf.

Wired (2017). Behandlungen und Technologien der Medizin der Zukunft. www.wired.de/collection/science/medizin-der-zukunft.

Wissenschaftliche Dienste Deutscher Bundestag (2017). Kurzinformation. Ausgaben für die Behandlung psychischer Erkrankungen WE 9-3000-041/17. www.bundestag.de/blob/531100/4f36889840d1105fc512aaf45fe12b40/wd-9-041-17-pdf-data.pdf.

Wolf, B., Scholze, C., and Friedrich, P. (2017). Digitalisierung in der Pflege – Assistenzsysteme für Gesundheit und Generationen. In M. A. Pfannstiel, S. Krammer and W. Swoboda (Eds.): *Digitale Transformation von Dienstleistungen im Gesundheitswesen III. Impulse für die Pflegepraxis.* Wiesbaden, 113–135.

Wu, X., Zhu, X., Wu, G.-Q., and Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering,* 26 (1), 97–107.

Xu, G., Song, A., and Li, H. (2011). Adaptive Impedance Control for Upper-Limb Rehabilitation Robot Using Evolutionary Dynamic Recurrent Fuzzy Neural Network. *Journal of Intelligent & Robotic Systems,* 62 (3-4), 501–525.

Yokoi, A., Yoshioka, Y., Hirakawa, A., Yamamoto, Y., Ishikawa, M., Ikeda, S.-I., Kato, T., Niimi, K., Kajiyama, H., Kikkawa, F., and Ochiya, T. (2017). A combination of circulating miRNAs for the early detection of ovarian cancer. *Oncotarget,* 8 (52), 89811–89823.

Yuste, R., Goering, S., Arcas, B. A. Y., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., Specker Sullivan, L., Teicher, M., Wasserman, D., Wexler, A., Whittaker, M., and Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature,* 551 (7679), 159–163.

Zeng, Z., Pantic, M., Roisman, G. I., and Huang, T. S. (2009). A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 31, 39–58.

Zukunftsinstitut GmbH (2012). HEAG 2040 – Die Stadtwirtschaft von morgen. Darmstadt. www.zukunftsinstitut.de/fileadmin/user_upload/Publikationen/Auftragsstudien/HEAG-2040-zukunftsstudie_2.pdf.

Legal basis:

BDSG – Bundesdatenschutzgesetz (BDSG-neu). www.gesetze-im-internet.de/bdsg_2018/index.html#BJNR209710017BJNE001900000.

DSGVO – Datenschutz-Grundverordnung. https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679&from=de.

Annex

Application	Function	Author	Year	Source
Public health today				
Identification of unwanted side effects when taking antipsychotics and antidepressives (GB)	Evaluation of electronic health records using an NLP algorithm (natural language processing) ("ConText algorithm")	Iqbal et al.	2017	PLoS ONE
Identification and evaluation of drug interactions for medications containing Ritonavir as an active ingredient (USA)	Evaluation of package informational inserts and clinical data, as well as information from post-marketing studies using algorithms in the context of a pharmacokinetic simulation model	Porcalla et al.	2017	Therapeutic Innovation & Regulatory Service
Identification of side-effects of a conjugate vaccine against meningococcal disease in children between two and 10 (US)	Evaluation of hospital electronic health records and emergency practices using an automated case-identification algorithm	Tartof et al.	2017	The Pediatric Infectious Disease Journal
Identification of geriatric systems and investigation of healthcare-service usage by older adults (US)	Evaluation of electronic health records using a natural language processing (NLP) algorithm	Anzaldi et al.	2017	BMC Geriatrics
Public health in the future				
Estimate of public healthcare expenditure (US)	Estimate of expenditure, using a machine-learning algorithm, based on public health-services data	Brady et al.	2017	Public Health Reports
Identification of persons vulnerable to the prioritization and allocation of resources during a catastrophe (CAN)	Decision-support algorithms based on the "Ontario Resident Assessment Instruments – Home Care (RAI-HC)" database	Van Solm et al.	2018	Journal of Emergency Management
Healthcare provision processes (health services research) today				
Dissemination of innovations and new processes through physician networks (US)	Fruchterman-Reingold algorithm used to analyze virtual doctor networks or so-called patient-sharing networks (PSNs) on the basis of Medicare data, with aim of revealing de facto regional network structures among physicians	Landon et al.	2012	JAMA

Application	Function	Author	Year	Source
Healthcare provision processes (health services research) in the future				
Planning for the collection, transfer and disposal of infectious medical waste (GRC)	Use of a genetic algorithm within an optimization model to calculate the optimal location for waste-processing plant, as well as optimal travel route for special garbage-collection trucks	Mantzaras and Voudrias	2017	Waste Management
Planning of logistics centers and flight routes for transporting medicines to patients in rural areas using drones (US)	Algorithm for the calculation of the optimal number and locations of logistics centers, as well as the optimal flight routes for the drones (preprocessing algorithm)	Kim et al.	2017	Journal of Intelligent & Robotic Systems
Scheduling appointments in physician practices	Automatic assignment of long and short appointments in a physician's practice using an algorithm embedded in an optimization model, on the basis of the reason for the consultation and other patient data	Becker et al.	2018a	Praxis
Medical research today				
Protecting data in the context of scientific publications (US, UK)	Anonymization of structured patient data from medical databases using data-protection algorithms	Gkoulalas-Divanis et al.	2014	Journal of Biomedical Informatics
Identification of research-relevant health events (DK)	Evaluation of Danish National Patient Register using search algorithms	Schmidt et al.	2015	Clinical Epidemiology
Medical research in the future				
Quantification of bronchiectasis incidents in children with cystic fibrosis as a possible end point of a clinical study (US)	Automated quantification of bronchiectasis incidents using CT scan of children, using an iterative algorithm	DeBoer et al.	2014	CHEST
Prediction / risk profiling today				
Prediction of hypersensitivity reactions due to drug interactions (US)	Clinical decision support (CDS) algorithms decide, on the basis of electronic health record data, whether overseeing physician should approve use of a particular medication	Goldspiel et al.	2014	Journal of the American Medical Informatics Association
Evaluation of new active pharmaceutical ingredients for prostate cancer, metastatic melanoma and systemic lupus erythematosus (CH)	EVITA algorithm (Evaluation of Pharmaceutical Innovations with regard to Therapeutic Advantage) generates a benefit-risk score on the basis of study findings	Szucs et al.	2014	European Journal of Clinical Pharmacology
Prediction of fertile days in a menstruation cycle for pregnancy planning or contraception (UK)	Algorithm for calculating fertile data in a menstruation cycle on the basis of data entered in the "Natural Cycles" app, such as body temperature and cycle length	Natural Cycles	2018	https://www.naturalcycles.com/de

Application	Function	Author	Year	Source
Prediction / risk profiling in the future				
Predictive clinical tool for predicting the action of glucocorticoid hormones in the body (GB)	Use of an algorithm (semi-quantitative signal transduction score flow algorithm) to analyze the action of glucocorticoid hormones	Bakker et al.	2017	PLoS Computational Biology
Precise risk assessment and automated observation of adverse post-operative events in intensive care units (DK)	Automated observation of patient vital signs using computer algorithm, with connected alarm system for adverse post-operative events and automated intervention recommendations	Haahr-Raunkjær et al.	2017	European Journal of Internal Medicine
Warning system for complications and co-morbidities among patients in general-practitioner practices (US)	Evaluation of electronic health records using deep-learning algorithm	Miotto et al.	2016	Scientific Reports
Prediction of the likelihood of disease, the severity of disease, and the identification of new breast-cancer biomarkers (US)	Evaluation of the Breast Cancer Wisconsin (Diagnostic) Data Set using a machine-learning algorithm	Banerjee	2017	Interdisciplinary Description of Complex Systems: INDECS
Prediction / risk profiling in the future				
Estimation of the suicide risk among U.S. military veterans (US)	Computer-supported textual analysis of unstructured clinical notes in U.S. Veterans Administration files, using a machine-learning algorithm	Poulin et al.	2014	PLoS ONE
Prediction of cardiovascular conditions and risk factors (US)	Analysis of retinal scans using a deep-learning algorithm for the identification of cardiovascular risk factors such as age or blood pressure	Poplin et al.	2018	Nature Biomedical Engineering
Prediction of depression among social-media users	Evaluation of the characteristics and metadata of user-uploaded images on Instagram, using a machine-learning algorithm and facial-recognition algorithm	Reece and Danforth	2017	EPJ Data Science
Prediction of psychosis within adolescent risk groups (US)	Automated analysis of voice recordings of adolescents using a machine-learning algorithm	Corcoran et al.	2018	World Psychiatry
Early detection and subtype determination for ovarian cancer (JPN)	Predictive model with predictive algorithm using various miRNAs	Yokoi et al.	2017	Oncotarget

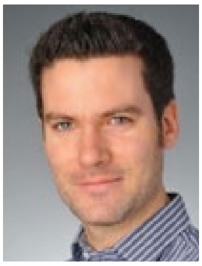
Application	Function	Author	Year	Source
Diagnostics today				
Early detection and classification of skin cancer (US)	Evaluation of skin images using a deep-learning algorithm	Esteva et al.	2017	Nature
Diagnosis of breast cancer in early-stage patients (ZAF)	Evaluation of a genetic test (70-gene profile) using the MammaPrint Pre-Screen Algorithm (MPA)	Grant et al.	2013	South African Medical Journal
Verifying the presence of an indwelling catheter in veteran-hospital patients, in the context of diagnosing a catheter-associated urinary-tract infection (US)	Evaluation of textual clinical information using a natural language processing (NLP) algorithm	Divita et al.	2015	Methods of Information in Medicine
Diagnostics in the future				
Diagnosis of a patient's primary complaints after hospitalization (US)	Evaluation of electronic health record using a natural language processing (NLP) algorithm with a focus on patients' vital signs, based on hospital staffers' recorded assessment of patients' physical state	Jernite et al.	2013	NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare
Diagnosis of skull-brain traumas and brain concussions (US)	Automated assessment of posture stability through evaluation of Microsoft Kinect® sensor data, with help of an algorithm (balance-error detection algorithm)	Napoli et al.	2017	Annals of Biomedical Engineering
Improvement of image quality in positron emission tomography (PET) and CT scans for cancer patients (FRA)	Improved signal-noise ratio through use of an algorithm (Bayesian penalized likelihood algorithm)	Vallot et al.	2017	Nuclear Medicine Communications
Diagnosis of frailty phenotype among elderly people	Use of algorithms to evaluate data from older individuals' mobile phones, such as number of steps, gait speed, etc.	Hanton et al.	2017	JMIR mHealth and uHealth
Wound diagnosis in the context of telemedicine treatment (IND)	Use of a particle swarm optimization algorithm (PSO algorithm) to improve image quality and the accuracy and efficiency of medical wound diagnosis	Chakraborty	2017	Wireless Personal Communications
Diagnosis of breast cancer (CH)	Evaluation of mammography scans using a deep-learning algorithm	Becker et al.	2018	Praxis
Diagnosis of pleural effusions and intrapulmonary signs of an active pulmonary tuberculosis in HIV patients (CH)	Evaluation of chest X-rays using a deep-learning algorithm	Becker et al.	2018	Praxis
Decision support for physicians when rendering diagnoses (US)	Diagnosis algorithm as an automated decision-support system, based on electronic health record data	Stusser and Dickey	2013	Journal of Medical Systems
Diagnosis of chronic obstructive pulmonary disease (COPD) among smokers (US)	Evaluation of CT scans using a deep-learning algorithm	González et al.	2017	American Journal of Respiratory and Critical Care Medicine

Application	Function	Author	Year	Source
Therapies today				
Therapy for septic shock among intensive-care patients (DE)	Use of a computer-supported decision-support system that includes algorithms to provide physicians with evidence-based diagnostic and therapeutic recommendations, taking into account complex standardized procedures	Tafelski et al.	2010	The Journal of International Medical Research
Therapies in the future				
Surgical planning for esophageal cancer (DE)	Test of a random-walk algorithm for segmenting esophageal tumors	Fechter et al.	2017	Medical Physics
Analysis of wounds in the context of telemedicine treatment (IND)	Use of a particle swarm optimization algorithm (PSO algorithm) for the segmentation of wound areas in medical image data	Chakraborty	2017	Wireless Personal Communications
Data-based support system for physicians' therapeutic decisions (DE)	Use of predictive algorithms to assess various therapies by forecasting patients' individual reactions to different options, based on patient data drawn from clinical databases. Example focuses on psoriasis, an autoimmune condition	Gräßer et al.	2017	Journal of Healthcare Engineering
Planning radiotherapy among cancer patients with head and neck tumors (UK)	Use of Google's "DeepMind" artificial intelligence to segment healthy structures and structures in need of therapy, using body scans as a basis	Powles and Hodson	2017	Health and Technology
Prognosis in the future				
Predicting complications such as bacteremia and pneumonia in severely injured patients (US)	Use of an algorithm to evaluate data from tissue, serum and wound-fluid samples to support clinical decision-making	Dente et al.	2017	Journal of Trauma and Acute Care Surgery
Prediction of acute respiratory diseases and mortality among smokers (US)	Analysis of CT scans using a deep-learning algorithm	González et al.	2018	American Journal of Respiratory and Critical Care Medicine
Early-warning system for critical cardiac events (IND)	Use of the Chan-Vese algorithm to evaluate cardiac magnetic resonance images, with goal of identifying changes in left ventricle in cases of aortic stenosis	Chandrasekhar et al.	2017	Journal of Medical Imaging and Health Informatics
Prediction of cardiac arrests among hospital patients (US)	Algorithm-supported early-warning system that can predict a code blue up to four hours before its occurrence on the basis of demographic information, hospital-admissions data, vital-sign values and laboratory measurements contained in electronic health records	Somanchi et al.	2015	Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining

Application	Function	Author	Year	Source
Rehabilitation today				
Training therapies for Parkinson's patients with bicycle ergometer and acoustic feedback in the form of music (DE)	Use of algorithms to carry out automated speed analysis of music and velocity changes in real time	Wolf et al.	2017	Digitale Transformation von Dienstleistungen im Gesundheitswesen III
Rehabilitation in the future				
Rehabilitation robot for upper-extremity motion therapy after a stroke (CN)	Use of an impedance control algorithm to produce a dynamic relationship between human and machine, facilitating individualized adjustment of the force- and movement-support functions	Xu et al.	2011	Journal of Intelligent & Robotic Systems
Gait analysis and gait-phase detection for prosthesis wearers, for the purposes of calculating what electrostimulation is necessary at what stimulus locations (DE)	Use of algorithms to process information from leg-mounted Bluetooth sensors to measure acceleration, angular velocity and magnetic field	TU Berlin	2017	https://www.ige.tu-berlin.de/bemobil/forschung/teilprojekt_b
Nursing care today				
Algorithm-controlled warning system in a home for senior citizens (TigerPlace) (US)	Use of algorithms to report deviations from daily routines (sensor-based measurements in the living environments), by providing alarms to caregivers	Rantz et al.	2013	CIN: Computers, Informatics, Nursing
Nursing care in the future				
Computer-supported decision-support system for nursing staff on a maternity ward (US)	Before being deployed, an algorithm learns the difference between good and bad decisions. On this basis, it offers recommendations regarding how the ward's nursing staff and rooms should be distributed in order to ensure optimal care	Gombolay et al.	2016	The International Journal of Robotics Research

Authors

ceres
cologne center for
ethics, rights, economics, and social sciences
of health



Marc Jannes, M.Sc., studied health economics at the University of Cologne. From 2010 to 2015, he worked as a student and research assistant at the Research Unit Ethics at the University Hospital Cologne. Since 2015 he works as a research associate at ceres. Marc Jannes is the coordinator of the ceres research area Health and Society in a Digital Age and scientific coordinator of the ceres expert group Digital Transformation as well as co-author of the study Digital Self-Determination. His research is focused on technology use and digitization in health care, online health information and digital self-determination as well as the assessment of quality of life.

Dr. phil. Minou Friele is a philosopher specialised in practical ethics. Since 2016 she has been working as a researcher, coordinator and lecturer at the cologne center for ethics, rights, economics, and social sciences of health (ceres) and the Ethics Research Center of the University Hospital in Cologne. Previously she had been working at the Institute of Science and Ethics at the University of Bonn; the Institute of Philosophy and the Institute of Medical Sociology at Heinrich-Heine University in Dusseldorf, and the European Academy of Technology and Innovation Assessment in Ahrweiler; and had spent time as a visiting research fellow at the Uehiro Center for Practical Ethics in Oxford, UK and at Harvard University, Cambridge MA, USA. Her current research focusses on ethical questions related to current developments in medicine and healthcare, such as genome editing, public health issues, and, last but not least, digitization.

Christiane Jannes (née Noack), M.A. completed her bachelor's degree in health economics at the University of Cologne, and her masters in medical management at the University of Duisburg-Essen. From 2011 to 2016, she worked as a student and research assistant at the Research Unit Ethics and at ceres. Since June 2016, she has served there as a research associate and as coordinator for the “PerF-ICF” project in the area of health services research. Her research focuses on digitization in healthcare, care for prematurely born children, and the International Classification of Functioning, Disability and Health (ICF).

Christiane Woopen is managing director of ceres (Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health), professor of Ethics and Theory of Medicine and head of the Research Unit Ethics at the University of Cologne. In the context of her national and international research projects, she studies issues including how to maintain quality of life across an entire life, genome editing, neuroethics, reproductive medicine, and digital change particularly within the healthcare sector. She has also served as a policy advisor, including as chair of the German Ethics Council (2012-2016), president of the Global Summit of National Ethics/Bioethics Committees (2014-2106), chair of the European Group on Ethics of Science and New Technologies (since 2017), and co-speaker of the Data Ethics Commission of the Federal Government of Germany (since 2018).

Bertelsmann Stiftung activities on this topic

The Digital Patient
This project addresses the opportunities and risks associated with digital change in the healthcare sector. It develops concepts and recommendations regarding digitalization in the service of health **(Blog at: blog.der-digitale-patient.de)**

Ethics of Algorithms
This project addresses the fundamental societal consequences of algorithm-based decision-making. It aims to contribute to the development of algorithmic systems in a way that leads to more social inclusion for all. **(Blog at: algorithmenethik.de)**

Legal notice

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telephone +49 5241 81-0
www.bertelsmann-stiftung.de

Managing editor
Uwe Schwenk

Authors
Dr. Minou Friele
Christiane Jannes
Marc Jannes
Prof. Dr. Christiane Woopen

Translation
Dr. Barbara Serfozo, Berlin

Design
Dietlind Ehlers

Photos
© peshkova, flareimages, Photographee.eu – stock.adobe.com; Minou Friele; Medizin-FotoKöln; Reiner_Zensen

Addresses | Contact

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telephone +49 5241 81-0

Dr. Thomas Kostera
Senior Expert
Program Improving Healthcare – Informing Patients
Telephone +49 5241 81-81204
Fax +49 5241 81-681204
thomas.kostera@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de

BertelsmannStiftung

BertelsmannStiftung

Algorithms in healthcare