ARTIFICIAL INTELLIGENCE AND THE FUTURE OF THE DIGITAL WORK-ORIENTED SOCIETY

An outline for a holistic technology impact assessment

Thorben Albrecht and Christian Kellermann
October 2020

Factors for the use of AI at the workplace: technological performance of AI, the regulatory framework and whether AI can be integrated into production and labour processes.

An uncontrolled digitisation with AI on the top threatens to throw our work-oriented society off its current "balance". A laissez-faire policy would not be a good path towards a digitised work-oriented society.

An AI impact assessment should action on a suitable political-regulatory framework, transparency about objectives and operating modes of AI, participation and qualification of employees in the use of AI.

FRIEDRICH EBERT STIFTUNG

ARTIFICIAL INTELLIGENCE AND THE FUTURE OF THE DIGITAL WORK-ORIENTED SOCIETY

An outline for a holistic technology impact assessment

Content

1

INTRODUCTION

Digitisation has accelerated the transformation of labour. Even though work today is still prone to changes, there are certain periods during which the changes occur more rapidly. In this process, the implementation of new technologies plays an important role; however, this always takes place within the context of (and in interaction with) other social and cultural changes and circumstances. Digitisation should also be perceived in the milieu of new forms of globalisation, demographic development, changes in education, and, last but not least, a shift in working people's values.

A qualitatively new technological stage has begun under the label of "artificial intelligence" (AI) as part of the implementation of certain digital technologies used within the labour context in society. AI is the collective term for algorithms which can process large volumes of data and which are capable of learning and finding complex solutions autonomously. AI has the ability to self-optimise and can be characterised by an immanent complexity and non-transparent approaches ("Black Box"). This distinguishes AI from other instruments (including digital tools) that have been used thus far. At the company level, AI systems can organise, manage, and control labour. These systems can also create a self-organising and optimising structure of business and labour relationships, as is done, for example, on labour platforms. The application of AI systems in the labour context thus leads to qualitative and quantitative changes which, in turn, require an adaptation of the regulatory framework as well as of labour relationships.

Particularly on the macroeconomic and industry-wide level, a question has arisen – in addition to the issue of the new quality of work – of what quantitative employment impact would occur if AI as well as other digital technologies increasingly penetrated into business processes. However, the discussion about the potential impact on overall employment is very varied and inconsistent in Germany as well as in many other countries. The predictions of what effects AI involvement will have on employment vary strongly: this is also related to the prospects of putting AI to use and its potential.

In this paper, we would like to determine the main dimensions of a work-focused assessment of progressive digital technologies in general (and specifically of AI) as they are used in the digital work-oriented society. The goal is to define the necessity and scope of activities required for a comprehensive technology impact assessment. The first dimension includes the technology itself and its technical and economic potential. AI is part of digitisation processes, and the two of them are hard to separate. Therefore, we will first look at the development of AI and the technological limitations of its applicability in order to make the issue more tangible for further analyses in the labour context. The second dimension is the workplace itself. How is AI used in particular situations and what labour-related, political, and organisational changes does it affect? The third dimension is macroeconomic: will the automation potential lead to technological unemployment, or will AI bring about growth and employment for all and at a higher level? Or will it heighten the polarisation that the first waves of automation and digitisation have already set in motion? In our view, these three dimensions are fundamental for the continuation of the technology impact assessment in the course of ongoing digitisation, and they can guide areas of political action which will be elaborated on in the last section.

2

AI AND THE TECHNOLOGICAL LIMITATIONS OF ITS APPLICABILITY

The methods of AI were made possible by the connection between high-performance hardware, big data, and machine-learning procedures (algorithm-based processes). While the term "artificial intelligence" was coined in the 1950s, the most significant procedures used today have been developed only in the past few decades. These developments were made possible by advances in hardware (smaller and faster chips and thus more processing power) and software (better algorithms). The current application of AI goes beyond the previous forms of the machine processing of information and allows the computer-controlled analysis of increasingly complex environments and processes. The development in storage technology, processing power, and speed have outlined a qualitative advancement from what used to be a predominantly theoretical approach to application-oriented AI.

Despite the numerous contributions to the debate about "general-purpose AI", for example, in popular culture, AI has always reached its limitations when it comes to autonomous understanding and comprehension, that is, "intelligence" as opposed to pattern recognition. Even the most trivial correlations can quickly overload an AI system. Even with manageable data sets, the enormous data combination possibilities can quickly produce uncontrollable volumes. For now, such a data explosion phenomenon can only be addressed with the aid of heuristics and the calculation of probabilities. This means, however, that even though the results of such processes could be good enough to outperform humans, they are not flawless and some uncertainty about their performance will always remain. This does not principally contradict the notion of intelligent action, because human action is also uncertain by definition in that it produces incomplete information. In turn, this is one of the reasons for the limited practical applicability of individual AI systems.

There is also uncertainty when predicting the further development of AI, which does not inhibit some people from granting AI too much potential in the work environment and endorsing it without restriction. As of today, genuine AI applications in most businesses are either very limited or are still being projected for implementation in the future, albeit mostly for only occasional and partly experimental use. Therefore, we are primarily interested in the development of specific areas of AI application in recent years to delineate the practical potential of the technology.

Machine learning is a relatively well-developed and significant application area of AI in diverse production fields. With the aid of machine learning, cross-sectional tasks in enterprises are easier to control using AI, for example: (1) higher-level recognition of actions which enables abstract recognition of comparable situations (pattern recognition) as well as (2) handling of extremely large data volumes. This means that machine learning makes it possible to use examples and observations to draw conclusions and make generalisations which, unlike memorising (immediate storage), can help propose and transfer solutions to various situations. Typical practical application areas include target-specific advertising and marketing, logistics, predictive maintenance, customer relationship management, and people analytics. For example, user data such as purchase or search behaviour across different platforms can help automatically produce similar offers. This is one of the reasons why the distribution phase of the supply chain (advertising, the Internet, and customer relationships) and machine learning are in an interdependent relationship.

In many other contexts, AI methods currently mark the technological boundary of digitisation in the business and labour environments. There is still a big gap between appreciating the potential of the technology and the specific application of AI in the workplace. In spite of this, it is conceivable that AI – and along with it the "intelligent" digitisation of the labour world in the foreseeable future – will increasingly and essentially change. (In a way, this is already taking place.) There has been evident development and application potential in the growing predictability of processes in such labour-related areas as production, services, administration, and agriculture. The precision level, however, has suffered mostly from the fact that AI as such is hard to gauge and measure, and many business and work processes (in which AI could be integrated) are similarly difficult to quantify, which results in their difficult coding.

Despite the lively debate about the potential of AI to fundamentally change our society and the working world, the specific effects of AI on the labour context are still under-researched, which is why the discourse tends to be anecdotal. Some of the anecdotes are presented by AI developers, whose aim it is to fully mechanise human intelligence: an ambitious and desirable goal for some, but an ambivalent and dystopian objective for others. Nils Nilsson, a pioneer in AI and robotics, defined this goal as "[the] complete automation of economically important jobs" (Nilsson 2005: 69). Nilsson proposed that an "employment test" could help measure what share of human work might be acceptably performed by an AI system. AI systems would only need to pass the same qualification tests as people are required to pass to be allowed to do special jobs. Current versions of such tests have come to the conclusion that in knowledge-based careers, AI is becoming increasingly competitive and has a more vertical outreach in enterprises (Webb 2019; Muro et al. 2019).

These and other tests, however, suffer from a whole range of methodological restrictions which can strongly reduce their validity. In particular, the performed jobs and tasks are frequently described only in keywords and are quickly dubbed as "redundant" when contrasted with the alleged capabilities of AI. Another factor is that not all adjustments are implemented in businesses merely because they can technically be done. Putting technological innovations into practice, including the area of AI technology, needs special considerations and requires actual limitations to be put in place. The following section will provide a systematic overview of the potential implementation (and its boundaries) of AI in a specific workplace. It seeks to provide an essential and progressive view of the differentiated understanding of the purpose of AI application in the labour environment as a whole, which targets the substitution of labour only in selected cases (see Section 4).

3

BUSINESS IMPLEMENTATION FOR FACING THE TENSION BETWEEN PROFITABILITY AND REGULATION

Several factors influence whether and to what extent AI is actually used in individual businesses.The first factor is the technological performance of AI.The second factor is what AI is allowed to do (what regulatory frameworks are in place: under what conditions, for what tasks, and with what requirements AI can be implemented).The second point includes ethical and regulatory limitations as well as the question of how much AI can be integrated into production and labour processes: i.e., its "integration capacity".Both of these factors establish a framework for a business decision on whether an investment in AI will be truly beneficial in a given case.The third factor (the issueof what the business calculation of costs and income will look like) is particularly hard to addressin the case of AI. The decision about the actual implementation of AI in businesses thus depends on what AI can do and what it is allowed to do, as well as what it can bring to the business in question.

Regardless of the frameworks, many business contexts make it difficult to gauge what investments in AI will actually mean for production and labour processes, and whether they could help achieve the desired progress in production. Because of the fact that the application of technology changes as much as technology changes itself, such decisions are hard to make for many businesses.The speed and scope of AI development in the coming years is wide open. With individual implementations, it is difficult to estimate how fast the operational adaptation to AI use will be applied and how long it will last.Moreover, the regulatory framework is a fast-changing variable because it is only in its initial stages in many areas.Finally, it is difficult to assess whether the technology can be successfully integrated into the work organisation of a business and whether it can actually improve the overall processes in production or the provision of services.This depends very much on "soft" factors, namely, whether the interaction between AI and people (colleagues, customers, and so on)will succeed and is sufficiently productive.If thisinteraction fails, negative consequences cannot be ruled out.

The tension between moving technological boundaries, unclear regulatory framework, and operational functionality makes the decision-making process concerning making economic investmentsvery uncertain.If a positive decision is nonethelessmade and an investment in AI is carried out, it might be driven by a general technological optimism (a desire to present oneself as a "front-runner"), or it might simply be acting upon the advice of often exaggerating consultancy institutes advocating for a technological future.As a result, even expensive and disastrous investments are quite possible (note, for example, the experience with the "CIM Ruins" [Computer-integrated Manufacturing]in the 1970s and 1980s).On the other hand, the productive potential of AI might not be utilised at all because of general uncertainty and scepticism.In both cases, an informed discussion and impact assessment of the application of AI technology can help reduce the uncertainties affecting the decision-making processesin businesses.

What are the expected outcomes of AI application in the working environment or directly in the workplace?In which areas does (or could) AI play a role?Ultimately, such questions are relevant if we take into consideration the fact that AI implementation will affect operational functionality (the "functionality of labour")and alter the qualitative conditions and the power relationships in the workplace and across anentire company.

Based on the current (qualitative) potential-focused analysis, it can be expected that AI implementation will likely affect cognitive routine tasks; in this sense, "routine" should be more broadly formulated to include progressive AI capabilities (see more on this in Section 4).At present, such tasks already encompass such things as the processing of standard cases in finance, insurance, medicine, health care, and many areas of the law.In these areas, AI plays a role that goes beyond merely evaluating data, making predictions, and producing diagnostics: it is also actively involved in research.However, the discussion about pattern recognition in the medical field has shown that a full-scale replacement of human decision-making isnot (yet) possible. At the same time, there has been enough evidence to prove that AI could be used in an assistant capacity provided this is done within reasonable limits.

The social and care services provided to people are another field of potential AI implementation which demonstrates its limitations.Construed as the physical agents of AI, robots are often the barometers of the degree of the automatisation of a production line or service as a consequence of digitisation and the related impact on labour and employment (cf. Dauth et al. 2017; Bessen 2018). AI can control a robotic system which in theory could carry out all kinds of assistance activities in a company or household. Image recognition, sensors, and actors (that is, all technologically complex construction units) are thus increasingly capable of fulfilling diverse human-like tasks. Nevertheless, personal services make it particularly evident that reasonable decisions and emotional tasks will still need to be largely performed by humans. One of the reasons for this is that AI cannot develop human emotions (simulations are possible, but have so far been qualitatively very limited in comparison with human emotional expressions). Another reason is that humans, as emotional beings, will accept reasonable information about themselves primarily from other people rather than from machines. For example, in the insurance business, where the assessment of insurance claims is fully automatised, the communication of rejected claims is normally carried out by people. Overall, we have to assume that activities which require emotional intelligence and empathy, as well as those that involve making ethical decisions, will largely remain in the human domain for some time.

The actual decisions that are made about AI application in a company or a workplace consequently influence the possible quantitative impact of AI on employment in a business, industry, or sector, and ultimately the overall economy. The following section discusses the quantitative effects on employment based on various assumptions. The conclusion is reached that neither the threatening scenario of technology-driven unemployment nor a new "Pareto efficiency" are likely. Nonetheless, the idea that AI and digitisation will lead to a further polarisation of labour, employment, income, and status is worthy of consideration: the less controlled the manner in which AI and digitisation enter the labour environment, the greater the polarisation.

4

GREATER POLARISATION IN THE LABOUR MARKET

The discussion about the quantitative impact of digital technology implementation on employment has been going on for several years now. Quite a few notable studies have come to the conclusion that the digitisation of the labour environment will result in massive upheavals in the form of job losses. This also applies (perhaps most of all) to AI and the growing use of machine-learning systems.

Methodologically speaking, the reasons for such results usually include the comparison of profession profiles in labour market statistics as well as assessments of technology potential based on AI development or the numbers of registered patents. This aids the argument that in the next two decades up to a half of all jobs could be eliminated due to digitisation, particularly as a consequence of using AI. Professions in transport, logistics, manufacturing, and services are among the most endangered. Low- and middle-income groups will beat the centre of the job-cutting process, and this is why technologically driven unemployment further catalysed by digitisation would accelerate existing polarisation. In addition to manual routine activities,machine learning and mobile (lightweight) robotics could perform cognitive tasks without firm specifications and thus bring about upheavals in the middle and upper strata of the workforce (Frey/Osborne 2013; 2017; Muro et al. 2019).

There is a consensus that digitisation can potentially lead to the substitution of labour. The only question is how great this potential is and what counterforces there will be if they are necessary. The difference between a direct disruption in the labour market and moderating effects is usually made by a "net calculation" across society (Arntz et al. 2017; Arntz et al. 2018; Dauth et al. 2017; Fuchs et al. 2018). Jobs or activities eliminated by digitisation are contrasted with new jobs in other areas. Growing productivity propelled by digitisation brings macroeconomic competitive advantages, added value, and ultimately employment effects (McKinsey Global Institute 2018; World Economic Forum 2018).

A common assumption in the scenario of technology-driven unemployment is that primarily very simple tasks can be automated. Nonetheless, the higher the routine degree of an activity, the greater the potentialis for its substitution. Profound research has clearly shown that no routine is like any other; after all, even the simplest routine tasks are incorporated into work processes as well as the whole organisation and cannot be simply broken up and structured anew. Many of these simple routine tasks are valuable precisely because they require practical knowledge and experience that is hard to transfer and formalise. In other words, it is something that AI has not been able to sufficiently reproduce so far. A static, isolated, and separable understanding of routine work often does not do justice to the tasks in question and in turn only exaggerates the realistic capabilities of AI. The capacity of human labour is, on the contrary, determined by means that are qualitative and context-dependent and cover a broad spectrum of activities (Pfeiffer/Suphan 2015). Knowledge that is not formalised but "tacit" (Autor 2015) goes beyond formal qualifications and involves such human senses as intuition, gut feeling, and emotions. It also comprises general knowledge and common sense, i.e., precisely the degree of understanding which AI and machine learning still have great lengths to reach. If the labour capabilities are not part of this perception (if that is even possible), it will quickly transpire that the pendulum has not yet swung to the AI side and that there is still a long journey ahead to achieve simple and informal human capabilities in the labour context.

The binary opposition between routine and non-routine tasks is therefore very limited and encourages a premature definition of digital technologies (including AI tools) as instruments of human replacement. It is evident, however, that the existing processes and organisation of work are becoming more efficient owing to AI: driver less transport systems, man-robot-collaboration (cobots), smart glasses, 3D-print and additive manufacturing, digital assistance systems, enterprise resource planning, digital twins, and other innovations are increasingly becoming part of the company-level and industry-wide division of labour. Such systems have far-ranging effects on individual jobs and are accompanied, in many cases, by a concentration of workflow and workload (Dispan/Schwarz-Kocher 2018).In addition to the qualitative aspects of these alterations in labour, this development also leads to possible quantitative changes. But drawing conclusions based on specific (negative) quantitative employment impacts does not do justice to the inherent complexity of routine tasks.

On the other hand, the potential for substitution is accompanied by positive quantitative employment expectations. Up to the individual burn-out limit,work intensification is connected with a growth in productivity. Digitisation and the implementation of AI require investments to be made, and these investments can mean either more or less employment.In an optimistic scenario, the implemented digital technology will lead to a technological upgrade and change the capital structure in a business. Rising overall employment is thus a consequence of the growing demand for a concrete type of capital. The increase in production demand in industries providing the inputs for this type of capital leads to growing employment in the economy as a whole.

Furthermore, investments in digitisation technology alter the cost structure and thus also the relative competitiveness of companies. Businesses that cut their costs thanks to digitisation can lower their prices and increase the demand for their products and services correspondingly, provided there is constant demand from other sectors of the economy (Arntz et al. 2018). Consequently, the output of the investing businesses increases and produces new income in the form of wages, profits, and capital income. An important aspect in this scenario is the competitiveness effect as well as the division of earnings into capital- and labour-related income.Rising productivity reduces the production costs for automated activities, which can lead to a growth inprofits or an increased demand for work in non-automated activities. Corresponding simulations have shown that this is often followed by long-term surges in the overall demand for the workforce in the economy (Fuchs et al. 2018).

Looking at these and similar considerations more closely, it becomes apparent that they strongly rely on spill-over effects from one sector to another. However, they are much more complex than is commonly assumed. Many problems emerge; if capital productivity increases in relation to work, technological innovations are generally labour-saving,meaning that it is rather unlikely that growing productivity would result in increased average wages. If labour is replaced by technology faster than new labour is created, then technology replaces work; and no increase in labour demand in other sectors is necessarily created. Ultimately, digitisation decisions made by a business will always be the result of a calculation of relative factor prices, i.e., relative prices for all necessary production factors. It is therefore not unlikely that digitisation will be accompanied by sinking average wages, because the substitution effect leads to a decreased labour demand (Acemoglu/Restrepo 2018). Ultimately, the outcome will be a growing inequality first on the labour market and then in the whole economy (Korinek/Stiglitz 2017). This context also highlights the need for a regulatory framework or a redistribution of profits from innovations, because otherwise the overall outcome from technology-driven innovation can be negative for society as a whole when compared with the situation before the innovation (Acemoglu 2019).

Even if higher wages become a reality as a result of digitisation in an industry or the economy as such, this does not mean that potential purchasing power will materialise in new purchasing activities in either the same or other business areas. A lot depends on other factors, such as the way digitisation creates a new structure of economic demand (for instance, via new purchasing patterns), how high or low earners profit from it, and what savings rates currently shape the economy. Until now it has been empirically observed that higher qualified labour profits from digitisation and human activity aremostly complemented by it. There is flexible demand for the corresponding products and services provided by the workforce; at the same time, however, there is inflexible labour supply in these areas (i.e., a shortage of skilled workers). Activities performed by workers with lower qualification profiles paint a different picture. The demand for manual activities is relatively inflexible as far as their price is concerned; if the price of manual activities drops due to digitisation, the demand for them does not rise correspondingly. We are now experiencing the preliminary stage of "Polarisation 4.0" (cf. Autor et al. 2017; Autor/Salomons 2017).

The automatic interconnection between digitisation, boosts to productivity, and a growing (macroeconomic) demand for labour, as implied by several involved actors, turns out to be rather dubious. It still remains to be seen whether digitisation and the increasing use of AI will have positive or negative effects on labour and employment. A lot depends on what direction is taken and the relevant regulatory framework. A laissez-faire policy would not be a good path towards a digitised work-oriented society. On the contrary, uncontrolled digitisation with AI on the top threatens to throw our work-oriented society off its current "balance". In the end, there will be (too) few winners and (too) many losers.

5

AN OUTLINE FOR A HOLISTIC TECHNOLOGY IMPACT ASSESSMENT

In the foreseeable future, however, it seems that we will not be addressing the fundamental question of whether people will be working or not. Technology-driven unemployment, being a potentially comprehensive and long-lasting state of affairs, principally involves a scenario that is itself rooted in technological progress, just like a new "Pareto efficiency" for labour and employment.

Nonetheless, in order to achieve and maintain a socially, economically, and environmentally balanced work-oriented society within such a highly dynamic process as digitisation, it will not suffice to evaluate the impact of technology implementation and fine-tune the corresponding measures by means of regulation. The rapid development of core digital technologies, with AI and machine learning at the current technological frontier, requires a continuous and holistic approach to labour and the socio-economic dimension of technology (Kellermann/Obermauer 2020).

As was mentioned above, the relevant aspects of such an approach have yet to be defined. This is harder than it might seem at first sight, since the development – be it technological, corporate, or macroeconomic – is very dynamic. Defining the first discussed aspect – technology and the prediction of its potential – is difficult, if not impossible, even for AI developers. Both software and hardware keep on hitting severe barriers of further development. Developing them beyond these barriers is not impossible, but such advancement is a kind of wild card which calls for a whole number of detailed solutions in diverse areas and labour contexts. Besides, it cannot be ruled out that many individual types of artificial intelligence will establish a form of AI that will be capable of performing activities aimed at its initially defined purpose as a sort of general-purpose intelligence (Russel 2019). At present, however, this seems to be a rather unlikely scenario.

It is important to continuously and carefully observe all technological developments in society. An open and independent technology impact assessment which is focused on specific uses as well as social impact thus remains a fundamental prerequisite for the tangible evaluation of the potential of technology, including its social impact. In this regard, the establishment of an AI observatory by the German Ministry for Labour and Social Affairs is a step in the right direction[1].

In addition to technological advances, the ongoing transformation of labour, industry, and individual workplaces and labour contexts has to be continuously assessed. Representative evaluations of AI implementation in the labour context have been scarce in Germany as well as in Europe. One exception is a joint research project between the American IT corporation IBM and the German services union Ver.di, which has been commissioned to enable the application of Watson-AI for IBM customers (IBM 2019). The project specifically pursues the question of what effects AI implementation could have for services activity.

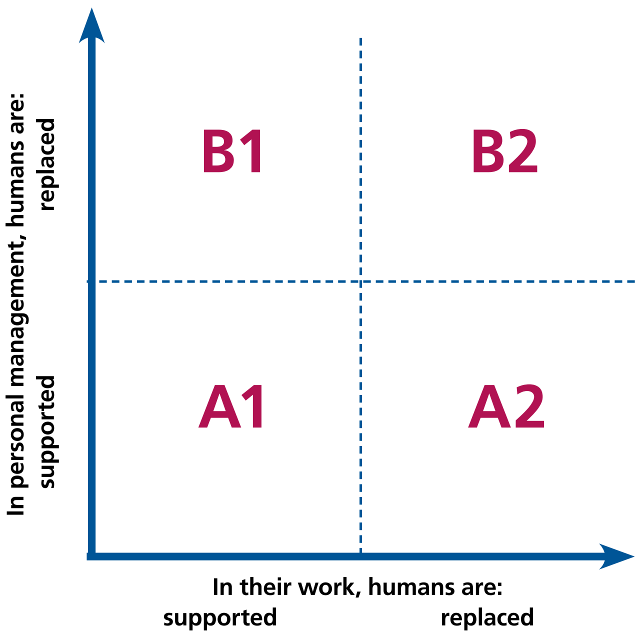
In order to make a broadly defined and labour-focused technology impact assessment, it is necessary to categorise and continuously gauge the application forms of AI in the labour context. In the process, the primary aim should not only be to observe the technical differences between various AI systems or determine the degree of technological advancement, but also to understand the role which AI plays (or should play) in a company. This role will then define different demands on the regulation and transparency of AI as well as framework conditions for the required participation and competence of workers and their representation in the business.

Such a technology assessment impact needs adequate categorisation and tools which can be used on the company level as well as across the industry. The categorisation will initially require the use of two aspects of AI implementation in the workplace in order to set up a matrix describing the consequences of implementing AI as part of the labour organisation (Albrecht 2020). The process will have to distinguish between a "horizontal" dimension, i.e., whether AI should serve as a tool to support or replace the work of employees (A), or whether it should be implemented as an instrument of human resources management (B). Examples in the A category include chatbots used to communicate

1 https://www.denkfabrik-bmas.de/projekte/ki-observatorium

with customers or to evaluate an insurance claim. Examples in the B category may include the use of AI to produce shift schedules, put teams together, assess performance, or select applicants. There is also a "vertical" dimension to the process in who makes the final decision. Here, it is important to differentiate whether the decision-making should remain within the competence of people (1) or be shifted to AI (2).

As far as the role that could be assumed by AI on a corporate level is concerned, there are the following four implementation modes: The technology can be a tool that will either support employees (A1) or partly replace them (A2). Additionally, the technology can also support the decisions made in human resources management (B1) or autonomously take over the process (B2).



The different categories can help clearly define the various requirements for the framework conditions under which AI can be implemented for use in such a way that does justice to socially and economically sustainable work design (best labour practice). It can be assumed that the (technological) regulatory demand in A1 is rather small, while it is quite high in B2. For example, work platforms present an extreme case of the B2 category, because ultimately the entire "business organisation" will be more or less taken over autonomously by AI.

This categorisation can contribute to the establishment of technology impact assessment practices for AI implementation at the workplace. This seems necessary because the advancement of social institutions, including the workplaces in question, should not be driven solely by technologically disruptive developments. More institutions should be established where the application of new technologies can be aligned with social norms and demands as well as with individual workplaces and the role of labour in the work-oriented society as a whole.

Four action levels are of particular importance: (1) a suitable political-regulatory framework, (2) transparency about objectives and operating modes of AI or technology application in specific cases, (3) participation of employees in the implementation and use of the technologies, and (4) employee competence to deal with AI constructively and critically.

(1) As far as AI implementation in numerous areas of the labour world is concerned, it will be necessary to carry out inspections (and ensure further development) of the existing regulations and institutions. An essential area is data protection. The European General Data Protection Regulation (GDPR) introduces such important principles as "privacy by design" and "privacy by default", and it defines the rules that shall also apply for all categories of AI systems and that shall be put into practice. In addition, it makes room for more specific regulations for national lawmakers when it comes to employee data protection due to the special interdependence of employees. Such special protection rights should become a necessary option, especially for AI implementation within the B1 and B2 matrixes. Of course, the regulation also affects other areas, such as the adaptation of labour protection, labour law, and the rights of the participating workers. An example of regulating this area would be the prohibition of fully automated decisions made about layoffs.

(2) Another matter of essential importance for the implementation of AI is the transparency of its goals and functionalities, so far as this is possible with self-learning systems. Such a transparency primarily requires the setting up of criteria for use in data collection and evaluation as well as establishing what decisions would be taken on the basis of this data (categories A1 and B1) or made by it (A2 and B2). The most important question here is whether people would control technology or vice versa. It is not imperative to fully exclude decisions made by AI; however, it is necessary to create a system in which humans do not become a mere "appendage" to a machine-made order and will always keep the higher decision command in both the overall process as well as in critical individual situations. This issue is particularly sensitive in cases when AI is used in processes of personal management (B1 and B2). Even though these decisions are made by (human resources) management "only" upon recommendations made by AI (B1), the problem is that it is frequently impossible for the management itself to follow the AI-made recommendation. This is partly because many AI providers do not allow any insight into their systems. It is also because self-learning systems make it difficult to understand and follow their recommendations as long as no "explanation function" is made available.

(3) The implementation of AI systems in specific national business and labour contexts also means that AI has to be integrated into the respective stakeholder structures. In Germany, this means that the employees and the representatives of their interests should participate in this process from early on (in other European countries, different AI-re-

lated participation rules are in place). It is a challenge to make sure that such an early form of participation will actually and transparently take place. In practice, employers often lack access to the necessary information which they ought to make accessible for their employees and their representatives because the AI providers do not make the information transparent. Further demand for regulation has arisen from the fact that AI systems are learning in the business environment and, as opposed to non-learning IT systems, are constantly altering the technological basis. This dynamic character of learning systems is of great significance to all stakeholders and has so far been underestimated as far as regulation is concerned. Agreements concluded between social and business partners concerning AI systems will need to change their nature: in the place of constantly applied rules, regular consultations between social partners and new centralised and decentralised conflict solution mechanisms will need to be planned.

(4) Such "process agreements" may not be completely new for trade unions and workers' representatives; however, in order to address the specific challenges of AI implementation, it will become necessary to continuously educate workers in technical, legal, and especially cultural issues. This makes it possible for trade unions to enter a new field: industry-wide know-how focused on AI issues could help train and support workers' representatives and employees. They could, for instance, participate in the certification of AI applications regarding behaviour and performance control functions, at least with standard applications. Such a certification could help avoid an inhibited implementation of AI systems by facilitating the above process of employee representatives' involvement, even though it could not replace it due to the mentioned permanent changeability of the systems.

Finally, employee qualifications will be crucial for the successful introduction and use of AI systems. It has to be ensured that employees are qualified to operate AI systems and are ready to assume new tasks, if necessary, when the tasks they have been performing thus far are replaced by AI. The qualification measures ought to extend to the operation of business- and activity-focused systems as well as to a fundamental understanding of the system in place, including the logic behind it. They should also mediate the new capabilities which will arise through the altered activities during the use of AI systems.

The four action levels described above contribute to a new social framework, which could settle the disruptive and dynamic technology development within a predictable and transparent institutional context. The social objective is to apply such an institutional framework to prevent a loss of control over the technology. It also aims to give AI and digitisation the necessary room for advancement and application, while at the same time allowing one to calculate the impact of technology on a workplace, the affected professions, and the overall economy, and to establish reliable control mechanisms on various levels. Otherwise, there is the danger that the work-oriented society will suffer a greater polarisation in income groups, and thus a much greater polarisation in society as a whole. Should we fail to regulate digitization in the context of work on the abovementioned levels, there will be a substantial threat to the working society as a whole. The technological dynamic needs to be merged with the normative issue of how, by what means, and why we hang on to a division of labour – and work for each other's ends.

BIBLIOGRAPHY

Acemoglu, D. (2019): Elizabeth Warren's Bold Ideas Don't Go Far Enough. Project Syndicate. Cited from: (https://www.project-syndicate.org/commentary/good-jobs-agenda-us-by-daron-acemoglu-2019-12).

Acemoglu, D. / Restrepo, P. (2018). Artificial Intelligence, Automation and Work. NBER Working Paper No. 24196. Cited from: (https://www.nber.org/papers/w24196.pdf).

Albrecht, T. (2020): Regulation, Transparenz, Beteiligung und Befähigung – Voraussetzungen für den erfolgreichen Einsatz Künstlicher Intelligenz am Arbeitsplatz. (Regulation, Transparency, Participation, and Competence: Prerequisites for theSuccessful Implementation of AI in the Workplace). In: Jens Nachtwei & Antonia Sureth (2020). Sonderband "Zukunft der Arbeit", HR Consulting Band 12, Berlin: Veröffentlichungsreihe für Qualitätssicherung in Personalauswahl und -entwicklung (in Vorbereitung).

Arntz, M. / Gregory, T. / Zierahn, U. (2017). Revisiting the Risk of Automation. In: Economics Letters, No. 159, pp. 157–160.

Arntz, M. / Gregory, T. / Zierahn, U. (2018). Digitalisierung und die Zukunft der Arbeit: Makroökonomische Auswirkungen auf Beschäftigung, Arbeitslosigkeit und Löhne von morgen. (Digitisation and the Future of Labour: Macroeconomic Impacts on Employment, Unemployment, and the Wages of Tomorrow). Bundesministerium für Forschung und Entwicklung (BMBF), Mannheim. Citedfrom: (http://ftp.zew.de/pub/zew-docs/gutachten/DigitalisierungundZukunftderArbeit2018.pdf).

Autor, D. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. In: Journal of Economic Perspectives, No. 29(3), pp. 3–30. Cited from: (https://economics.mit.edu/files/11563).

Autor, D. / Dorn, D. / Katz L.F. / Patterson, C. / Van Reenen, J. (2017). Concentrating on the Fall of the Labor Share. In: American Economic Review Papers & Proceedings, No. 107(5). pp. 180–185. Cited from: (https://economics.mit.edu/files/12544).

Autor, D. / Salomons, A. (2017). Robocalypse Now – Does Productivity Growth Threaten Employment? In: ECB Forum on Central Banking – Investment and Growth in Advanced Economies – Conference Proceedings. Frankfurt am Main: European Central Bank, pp. 119–128. Cited from: (https://www.ecb.europa.eu/pub/pdf/other/ecb.ecbforumcentralbanking2017.en.pdf).

Bessen, J. (2018). AI and Jobs: The Role of Demand.In: NBER Working Paper Series, No. 24235. Cited from: (https://www.nber.org/papers/w24235.pdf).

Dauth, W. / Findeisen, S. / Südekum, J. /Wößner, N. (2017). German Robots – The Impact of Industrial Robots on Workers. In: IAB-Discussion Paper, 30/2017. Cited from: (http://doku.iab.de/discussionpapers/2017/dp3017.pdf).

Dispan, J. / Schwarz-Kocher, M. (2018). Digitalisierung im Maschinenbau. Entwicklungstrends, Herausforderungen, Beschäftigungswirkungen, Gestaltungsfelder im Maschinen- und Anlagenbau. (Digitisation in Machine Engineering: Development Trends, Challenges, Employment Impact, and Design Fields in Machine and Equipment Construction).In: Working Paper Forschungsforderung, No. 094. Cited from: (https://www.boeckler.de/pdf/p_fofoe_ WP_094_2018.pdf).

Frey, C.B. / Osborne, M.A. (2013). The Future of Employment: How Susceptible are Jobs to Computerisation? Oxford Martin School (OMS) Working Paper.University of Oxford. Cited from: (https://www. oxfordmartin.ox.ac.uk/downloads/academic/The_Future_ of_Employment.pdf).

Frey, C.B. / Osborne, M.A. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? In: Technological Forecasting & Social Change. No. 114, pp. 254–280.

Fuchs et al. (2019): Zuwanderung und Digitalisierung. Wie viel Migration aus Drittstaaten benötigt der deutsche Arbeitsmarkt künftig? (Immigration and Digitisation. How Much Migration from Third Countries will the German Labour Market Need in the Future?). Bertelsmann Stiftung. Cited from: (https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Migration_fair_gestalten/IB_Studie_Zuwanderung_und_Digitalisierung_2019.pdf).

IBM (2019): KI-Studie über die Folgen für Arbeitnehmende und Arbeit. (An AI Study on the Consequences for Employees and Labour). Cited from: (https://www.ibm.com/de-de/blogs/think/2019/09/17/watson-ki-studie/).

Kellermann, C. / Obermauer, R. (2020). Von der Würde der Arbeit in digitaler und klimaneutraler Zukunft. (On the Dignity of Work in a Digital and Climate-neutral Future). spw 238

Korinek, A. / Stiglitz, J.E. (2017). Artificial Intelligence and Its Implications for Income Distribution and Unemployment. In: National Bureau of Economic Research. Cited from: (https://www.nber.org/papers/w24174.pdf).

McKinsey Global Institute [MGI] (2018). Notes from the AI Frontier: Modeling the Impact of AI on the World Economy. Cited from: (https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy).

Muro, M. / Maxim, R. / Whiton, J. (2019). Automation and Artificial Intelligence: How Machines are Affecting People and Places. Metropolitical Policy Program. Cited from: (https://www.brookings.edu/wp-content/uploads/2019/01/2019.01_BrookingsMetro_Automation-AI_Report_Muro-Maxim-Whiton-FINAL-version.pdf).

Nilsson, N. (2005). Human-level Artificial Intelligence? Be Serious!American Association for Artificial Intelligence. Cited from: (http://ai.stanford.edu/~nilsson/OnlinePubs-Nils/General%20Essays/AIMag26-04-HLAI.pdf).

Pfeiffer, S., / Suphan, A. (2015). Der AV-Index. Lebendiges Arbeitsvermögen und Erfahrung als Ressourcen auf dem Weg zu Industrie 4.0 (The AV Index: Living Work Capacity and Experience as Resources on the Way to Industry 4.0). Working Paper 2015 No. 1 (final version of adraft originally published on 13 April 2015). Cited from: (https://www.sabine-pfeiffer.de/files/downloads/2015-Pfeiffer-Suphan-final.pdf).

Russell, S. (2019). Human Compatible: AI and the Problem of Control. Allen Lane.

Webb, M. (2019). The Impact of Artificial Intelligence on the Labour Market. Stanford University Working Paper.

World Economic Forum (2018). The Future of Jobs Report (2018). Cited from: (http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf).

ABOUT THE AUTORS

Thorben Albrecht is former State Secretary of the German Ministry of Labour and Social Affairs. He was a member of ILO's Global Commission on the Future of Work.

Christian Kellermann is assistant professor at the HTW – University of Applied Sciences in Berlin and is the general manager of Denkwerk Demokratie. We would like to thank Kajsa Borgnäs for valuable comments and references.

IMPRESSUM

Friedrich Ebert Stiftung,
Regional Project on Labour Relations and Social Dialogue

Responsible:
Matthias Weber, Director Regional Project
www.fes-socialdialogue.org

ARTIFICIAL INTELLIGENCE AND THE FUTURE OF THE DIGITAL WORK-ORIENTED SOCIETY

An outline for a holistic technology impact assessment

→

Several factors influence whether and to what extent AI is actually used at the workplace and whether investments pay of: The technological performance of AI, the regulatory framework in place and whether AI can be integrated into production and labour processes.

→

The automatic interconnection between digitisation, boosts to productivity, and a growing (macroeconomic) demand for labour is rather dubious. A lot depends on what direction is taken and the relevant regulatory framework. A laissez-faire policy would not be a good path towards a digitised work-oriented society. On the contrary, uncontrolled digitisation with AI on the top threatens to throw our work-oriented society off its current "balance". In the end, there will be (too) few winners and (too) many losers.

→

An AI impact assessment should look at two dimensions: First, whether AI should serve as a tool to support or replace the work of employees (A) or whether it should be implemented as an instrument of human resources management (B). Second, whether AI supports human decision-making (1) or takes over decision-making (2). It should guide action on a suitable political-regulatory framework, transparency about objectives and operating modes of AI, the participation and qualification of employees in the use of AI.

FRIEDRICH
EBERT
STIFTUNG