Received 20 November 2024, accepted 12 December 2024, date of publication 18 December 2024,
date of current version 31 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3519979

RESEARCH ARTICLE

Computing-Model and Computing-Hardware Selection for ICT Under Societal and Judicial Constraints

YANNIK N. BÖCK[1], **(Graduate Student Member, IEEE), HOLGER BOCHE**[1], **(Fellow, IEEE), FRANK H. P. FITZEK**[2], **(Senior Member, IEEE), AND GITTA KUTYNIOK**[3,4,5,6], **(Fellow, IEEE)**

[1]TUM School of Computation, Information and Technology, Department of Computer Engineering, Chair of Theoretical Information Technology, Technical University of Munich, 80333 Munich, Germany
[2]Deutsche Telekom Chair of Communication Networks, Technical University of Dresden, 01187 Dresden, Germany
[3]Bavarian AI Chair for Mathematical Foundations of Artificial Intelligence, Ludwig-Maximilians-Universität München, 80799 Munich, Germany
[4]Department of Physics and Technology, University of Tromsø, 9019 Tromsø, Norway
[5]Munich Center for Machine Learning, 80538 Munich, Germany
[6]German Aerospace Center (DLR), 82234 Weßling, Germany

Corresponding author: Yannik N. Böck (yannik.boeck@tum.de)

The work of Yannik N. Böck was supported in part by German Federal Ministry of Education and Research (BMBF) within the National Initiative on 6G Communication Systems through the Research Hub 6G-Life under Grant 16KISK002; and in part by BMBF Quantum Program QDCamNetz under Grant 16KISQ077, QuaPhySI under Grant 16KIS1598K, and QUIET under Grant 16KISQ093. The work of Holger Boche was supported in part by BMBF within the National Initiative on 6G Communication Systems through the Research Hub 6G-Life under Grant 16KISK002; in part by the Bavarian State Ministry of Science and the Arts through the Project ''Next Generation Artificial Intelligence (AI) Computing (GAIn)'' and German Research Foundation, Deutsche Forschungsgemeinschaft (DFG) within the Gottfried Wilhelm Leibniz Prize Program under Grant BO 1734/20-1; and in part by DFG Projects under Grant BO 1734/44-1, Grant BO 1734/35-1, and Grant STA 864/9-1. The work of Frank H. P. Fitzek was supported by German Research Foundation, DFG as part of Germany's Excellence Strategy—EXC2050/1—Cluster of Excellence ''Centre for Tactile Internet with Human-in-the-Loop'' (CeTI) of Technische Universität Dresden under Project 390696704. This work of Gitta Kutyniok was supported in part by the Bavarian State Ministry of Science and the Arts through the Project ''Next GAIn''; in part by the Konrad Zuse School of Excellence in Reliable AI (DAAD); in part by Munich Center for Machine Learning (BMBF); in part by German Research Foundation under Grant DFG-SPP-2298, Grant KU 1446/31-1, and Grant KU 1446/32-1; in part by the LMUexcellent, funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder; and in part the Hightech Agenda Bavaria.

ABSTRACT This article discusses a formalization of aspects of *Cyber-Sovereignty* (CyS) for *information and communication technology* (ICT), linking them to technological trustworthiness and deriving an associated paradigm for hard- and software design. The upcoming 6G ICT standard is considered a keystone within modern society's increasing interconnectedness and automatization, as it provides the necessary technological infrastructure for applications such as the *Metaverse* or large-scale *digital twinning*. Since emerging technological systems increasingly affect sensitive human goods, hard- and software manufacturers must consider a new dimension of societal and judicial constraints in the context of technological trustworthiness. This article aims to establish a formalized theory of specific aspects of CyS, providing a paradigm for hard- and software engineering in ICT. This paradigm is directly applicable in formal technology assessment and ensures that the relevant facets of CyS – specifically, the principle of *Algorithmic Transparency* (AgT) – are satisfied. The framework follows an axiomatic approach. Particularly, the formal basis of our theory consists of four fundamental assumptions about the general nature of *physical problems* and *algorithmic implementations*. This formal basis allows for drawing general conclusions on the relation between CyS and technological trustworthiness and entails a formal meta-thesis on AgT in digital computing.

INDEX TERMS Sovereignty, accountability, transparency, explainability, integrity, computability, hardware models.

I. INTRODUCTION

The degree of interconnectedness and automatization in society has grown unprecedentedly throughout the past

The associate editor coordinating the review of this manuscript and approving it for publication was Barbara Masini.

thirty years. Further, the role of technology in everyday life is currently changing qualitatively. Given the recent technological progress, the engineering community envisions machine learning and data science to enable systems that will revolutionize healthcare, transportation, the labor market, and the realization of political systems. The upcoming 6G *information and communication technology* (ICT) standard is considered essential in this context, as it introduces the required infrastructure for applications such as the *Metaverse* or large-scale *digital twinning* [1]. To a great extent, the projected 6G services, e.g., low-altitude air-traffic control for drone systems, will rely on intelligent algorithms, controlling both virtual and physical components. As potentially momentous decisions will increasingly be made by machines rather than humans, hard- and software manufacturers must consider a new dimension of societal constraints. The related discussions on *machine ethics*, i.e., the questions of how algorithms should decide in certain situations, has been present for some time, mostly in the context of *trajectory planning* for autonomous driving [2], [3], [4]. Only recently, [5] highlighted that *trustworthiness* – an umbrella term for different relevant aspects of technology assessment – will soon become critical to ICT in general.

Ever since the introduction of *ChatGPT*, the potential risks of the coming generations of ICT have experienced a significant rise in attention within the public and legal domain. The emerging paradigms and legal regulations are often summarized by the term *Cyber Sovereignty* (CyS). The arguably most relevant contemporary examples are the *G7 Hiroshima Process on Generative Artificial Intelligence* [6] and the *European AI Act* [7]. As their names suggest, both specifically concern artificial intelligence rather than general ICT. This may be partly attributed to the subject's novelty and the recent developments in artificial intelligence, which have prompted a demand for regulatory measures in the public discourse. Nevertheless, both [6], [7] have direct relevance for near-future ICT systems, as these will rely on AI to perform decision-making tasks – including such within critical infrastructure.

In addition to their focus on AI, existing statements primarily address virtual aspects of ICT. That is, "soft" interactions between humans and ICT, in which the relevant ICT system does not control parts of the human's physical environment. Regarding near-future communication networks, however, the case of "hard" interactions, i.e., scenarios where software directly controls physical agents, has to be addressed. Further, existing statements tend to express societal requirements in abstract terms. Regarding trustworthiness, the relevant standards and regulations must ultimately define formal technological specifications. While this is already the case in traditional technology assessment, – for example, both of an aircraft's engines must each provide enough *thrust* such that the aircraft can stay *airborne* with just one of them – the rise of ICT as part of society's critical infrastructure presents research and development with the challenge of translating unprecedented societal constraints into technological ones.

This article aims to establish a formalized theory of specific aspects of CyS, providing a precise paradigm for hard- and software engineering in ICT. This paradigm ensures that the principles of *Algorithmic Accountability* (AgA), *Algorithmic Transparency* (AgT), and *Right to Explanation* (RtE) are satisfied. The basis of our framework is presented in Section III-A. There, we introduce the principles mentioned above and discuss their relationship in the context of CyS. The provided discussion distinguishes AgT as decisive in the engineering context. In Section III-B, we further introduce the concept of (formalized) *physical problems* by a scheme of three axioms, which is necessary in order to derive the pivotal paradigm. Section III-C discusses the relation between physical problems and their algorithmic implementations. The resulting fourth axiom can be considered the core characterization of AgT in the context of technological trustworthiness and concludes the axiomatic preliminaries. In Section IV, we apply the established framework to models of computing hardware, which entails the main contribution of our work: The meta thesis on AgT in digital computing and the resulting paradigm of hard- and software design in ICT. Finally, we provide prospects on the implications of our results and an overview of related research in Section V.

II. BACKGROUND AND RELATED WORK

Evidently, the principles of trustworthiness and CyS are essential for all forms of ICT. However, the relevant literature primarily considers trustworthiness and CyS in the context of AI and machine learning. Arguably, this is due to the following reasons. First, given the significant advances the science of machine learning has made throughout recent years, large parts of the engineering community consider AI a key enabler for near-future ICT technology. Accordingly, concerning the safety of such technologies, the performance and behavior of AI occur as a bottleneck. Further, the potential harm caused by faulty next-generation ICT systems is unprecedented. Thus, it seems natural to pay special attention to trustworthiness and CyS in the context of AI.

Second, the approach of AI technologies towards problem-solving is much in contrast with the one of classical software: The latter ideally implements a provably correct method to solve a given task, while the former requires the machine to find the solution to the given task on its own. Ensuring that a machine finds (or recognizes whenever it does not) a correct solution when prompted with a problem adds another layer of complexity over merely implementing a particular correct solution. The black-box-like nature of many machine-learning techniques further contributes to this issue. Ensuring trustworthiness and CyS for AI-based ICT systems is thus extra challenging as compared to other forms of ICT.

The interactions of contemporary AI-based technologies with their surroundings are almost entirely of an indirect nature, influencing human choices rather than acting within the physical environment. The goal of *artificial general intelligence* [8], [9] includes a transition to direct interactions, utilizing deep-learning-based sensing and decision-making for physical agents. Often, the resulting algorithms exhibit unreliable and non-robust behavior [10], [11], [12], [13], [14], such as the vulnerability of artificial neural networks to slight input perturbations. Among others, [11], [15], [16], [18], [19] have demonstrated how small deviances in an artificial neural network's input data can create drastic fluctuations in the generated outputs, even if the deviances are unnoticeable to a human observer. Research and development must remedy this lack of trustworthiness [20], [21], [22], [23], [24] before AI-based technologies are safe to use in systems that involve physical agents. Especially, AI-based technologies must satisfy and provide

- robustness, particularly against changes in environments or situations, noisy or incomplete data, and adversarial attacks;
- transparency and interpretability, offering clear justifications for and explanations of all critical steps in the decision-making process;
- fairness, ethical compliance, and privacy, avoiding biases and ensuring the equitable treatment of diverse user groups and the secure handling of sensitive information;
- safety and security, including protection from potential threats and preventing unintended harmful outcomes.

Failing to establish trustworthiness in deep learning systems means it is impossible to provide performance guarantees. Accordingly, situations may occur where these systems display unexpected and potentially harmful behavior. This concern is widely recognized, even outside the scientific community. Policymakers have proposed guidelines and regulations for the relevant technological systems, such as the *European AI Act* [7] and the *G7 Hiroshima Leaders Communique* [6]. The outlined requirements are categorized according to their severity, depending on the extent to which the technological systems in question are safety critical. Specifically, the European AI Act creates a well-defined legal framework that some policymakers consider a "blueprint" for future regulatory proposals.

Nevertheless, the issue of trustworthiness in deep learning persists, as the core methodology has stayed the same. Thus, one may question the satisfiability of the proposed requirements for deep-learning-based technological systems. Notably, AI techniques such as *expert systems* [25] follow an entirely different approach, which makes them less susceptible to the vulnerabilities of deep learning systems. Informally speaking, they aim to provide trustworthiness benefits "by design", at least to some extent. However, compared to deep learning systems, this benefit comes at the cost of performance.

In contrast to classical technology assessment, the problem of feasibility remains for the proposed requirements. In particular, societal principles and legal regulations such as Algorithmic Transparency, Algorithmic Accountability, and Right to Explanation [7] are of an abstract nature. It is not evident how these principles mirror on the implementation level. The present article aims to fill this gap by framing these principles in the context of computability theory. Notably, the suggested framework incorporates the hardware level. Formalizing societal principles and legal regulations in a technologically applicable manner – that is, in terms of a mathematical model – is a nontrivial task. However, ensuring a technological system's adherence to such regulations is more comprehensive than their formalization, as the technological system is itself subject to mathematical modeling. Thus, it is essential to investigate whether the hardware in question is "compatible" with the formalized form of the relevant regulations. Otherwise, there is no guarantee that the technological system will satisfy the regulations in practice, even if it does so in theory. Within the relevant literature, this point of view is (to the best of the authors' knowledge) unique to the present work and [26]. Further, we try to draw some general conclusions. Particularly, the existence of a trustworthy algorithmic solution to some engineering problem depends on the underlying computing model, and varying results can indeed occur. This phenomenon is crucial, especially regarding real-world physical processes or, more broadly, problems modeled and represented within continuous domains.

The issue of understanding and verifying the behavior of software systems has been relevant ever since the birth of computer science itself. However, it has recently gained newfound attraction following the rise of machine learning and artificial intelligence. As indicated above, this is arguably due to the contrasting approaches to problem-solving of techniques such as deep learning on the one hand and classical software on the other: The latter ideally implements a provably correct method to solve a given task, while the former requires the machine to find the solution to the given task itself.

In both cases, verifying or even proving the correctness of a software system is a complex but, to some extent, viable task. In the simplest case, the verification process consists of running the software for a suitable list of inputs. The verifier needs to choose this list so that the desired result for each of its entries is known and can later be compared to the actual output. Note that, albeit presumably simple, finding a suitable list of inputs is highly nontrivial and sometimes impossible. While typical throughout all areas of software development, this technique is especially relevant in the field of machine learning. For recent examples in the context of ICT, see e.g. [27], [28]. *Model checking*, on the other hand, seeks to verify (in the sense of deductive theorem proving) a system's model against its formal specification [29]. Note that in contrast to the above, this form of verification corresponds to proving the correctness (with regards to the intended task) of the software system in a mathematical sense; the underlying mathematical framework is called *(modern) type theory*, c.f. e.g. [30]. In more abstract terms, this process aims to reduce the correct operation of the software to the correct operation of compiler and hardware. If the latter is provided, the software system (provably) behaves as intended. However, model checking is a highly intricate technique, which, in many practically relevant cases, is infeasible for sheer complexity. Further, even if complexity is disregarded, there exists a fundamental mathematical limitation to model checking: Gödel's incompleteness theorems imply that even if a piece of software is de facto correct (in the sense that for all possible inputs, the software provides the desired output), it is not necessarily possible to prove its correctness.

Initially, the rise of deep learning was not accompanied by novel validation techniques, as validation through basic testing was deemed sufficient given the achieved performance. Nevertheless, alongside the contemporary spread of AI software usage in everyday life, issues such as *adversarial examples* [11] indicating a discrepancy between intended and actual behavior emerged. In response, software developers applied attuned learning procedures – such as *adversarial training* [31], [32] – aiming to prevent the undesired behavior or *explainability* techniques trying to make the software's (decision) process transparent to the user [33], [34], [35], [36], [37]. However, these attempts have only been partially effective and mostly limited to narrow use cases. Moreover, newfound results suggest that it is impossible to prevent undesired behavior *in general* [38], [39]. Informally speaking, one may expect that for every attuned learning procedure, there exists a new set of "rogue inputs" that provoke the software to act in an unforeseen way. Consequently, the problem of appropriate benchmarking that ensures 'good' operation in complicated real-world settings persisted, eventually leading researchers to propose the concept of deep learning models *interpretable by design* [40]. Instead of focusing on the transition from benchmarking to real-world application, the design of these models a priori includes specific interpretability objectives, which developers can later refine by integrating insights gained from targeted tests and explainability techniques [41], [42]. Ethical and moral aspects form another highly relevant blindspot of conventional benchmarking. For example, deep learning systems are prone to biased decision-making and violations of users' privacy rights [43], [44], [45]. Research and development has tackled these challenges symptomatically through various technical modifications [46], [47]. Nevertheless, the underlying ethical and legal concerns remain unaddressed [47], [48], [49], [50], [51].

At the latest, when future technological systems employ machine-learning techniques in a safety-critical context, the trustworthiness problem must be solved. Like traditional technology assessment, intelligent software requires reliable certification standards [52], [53], [54], taking societal and legal considerations into account. In this sense, the present paper aims to contribute a 'trustworthy by design' framework that precedes classical and specific deep learning validation methods.

III. MODEL BUILDING: FROM ABSTRACT PRINCIPLES TO FORMAL CRITERIA

A. PRINCIPLES OF CYBER SOVEREIGNTY

To the best of the authors' knowledge, no widely accepted definitive characterization of CyS exists. The same holds true for its subordinate principles AgA, AgT, and RtE. Subsequently, we provide brief conceptual definitions that serve as a reasonable basis for a formal analysis in the context of ICT.

From Section I, recall that technological trustworthiness serves as an umbrella term for different aspects of technology assessment. For a comprehensive description, we again hint at [5]. We refer to the aspect of trustworthiness that is relevant within the scope of this article as *integrity*. In contrast to CyS, the notion of (technological) integrity emerged in the engineering context and has a direct interpretation therein. One of this article's pivotal contributions is establishing a link between CyS and trustworthiness. For didactic reasons, we will thus define integrity alongside CyS, AgA, AgT, and RtE and return to the definition in Section III-C.

Cyber Sovereignty (CyS): CyS refers to the ability of a group of individuals – commonly a nation – to jointly decide upon and regulate the usage of ICT technologies within the group without being dependent or influenced by parties outside the group. This ability encompasses rights and responsibilities on the individual level, as well as rights and responsibilities on the level of the group as a whole. In particular, it encompasses the group's ability to jointly decide which ICT solutions are legitimate for a specific engineering problem.

Algorithmic Accountability (AgA): AgA refers specifically to the legal regulations of which party, individual, or possibly system is to be held accountable for harm or losses resulting from algorithm-based decision-making. The latter especially applies to decisions that are deemed faulty. Primarily, AgA is a judicial concern. As a prerequisite, a cyber-sovereign group is responsible for establishing conclusive regulations on which algorithm-based decisions are considered faulty in a specific context.

Algorithmic Transparency (AgT): AgT requires that the factors determining the result of an algorithm-based decision be visible to the legislator, the operator, the user, and other affected individuals. Upon proper formalization, it is a property directly applicable in the engineering context. Within a cyber-sovereign group, specific ICT systems may be legitimate only if the employed algorithms satisfy transparency.

Right to Explanation (RtE): RtE refers to the right of an individual affected by an algorithm-based decision to know the entirety of factors and their case-specific expressions that lead to the decision. The term refers to the possible right in legal and the general right in a philosophical sense. Again, a cyber-sovereign group is responsible for establishing conclusive regulations on the extent of this right within their jurisdiction.

Integrity: In engineering, integrity refers to a state where a technological system resides within its design-based margin of operation. For ICT applications that include decision-making for physical agents, integrity requires that the relevant algorithm correctly captures the state and (physical) dynamics of the agent in question. Like AgT, integrity is a property directly applicable in the engineering context upon proper formalization.

As indicated, a cyber-sovereign group is responsible for establishing clear regulations regarding AgA and RtE. In turn, AgT is a requirement for both AgA and RtE. Following the definitions above, it is possible to provide RtE if and only if the relevant algorithm satisfies AgT. The interplay of AgT and AgA emerges from the observation that without AgT, it is impossible to differentiate between faulty and intended algorithm-based decisions. Recall that, as a prerequisite of AgA, a cyber-sovereign group is responsible for establishing conclusive regulations on which algorithm-based decisions are considered faulty in a specific context. The agreed-upon legal regulations determine the entirety of visible factors relevant to a specific decision-making task. On the other hand, the decision-making by an intransparent algorithm incorporates factors that are invisible to the legislator per definition. If an intransparent algorithm is considered legitimate, any resulting decision can be arbitrarily attributed to some invisible factor, leaving no means of verification. Consequently, the operator is either penalized without proof of their delinquency, or the distinction between faulty and intended decisions is rendered infeasible altogether.

While the principles of AgA and RtE may have had little practical relevance in the early days of ICT, their importance is evident given the current advances in technology and engineering. From Section I, recall that *hard* human-technology interactions – that is, interactions in which an intelligent system controls the physical environment of humans – will be an essential part of the envisioned 6G services. As indicated in Section I, autonomous driving is the arguably most prominent pertained technology discussed in the literature [2], [3], [4]. The questions of cause and responsibility after (possibly fatal) accidents involving self-driving cars will have to be answered from the perspectives of AgT, AgA and RtE. Observe that the relevant results established in [2], [3], and [4] are directly related to our framework, which we will further discuss in Section V.

In order for a group of individuals to maintain CyS, the group must understand the cause of and responsibility for algorithm-based decisions in the context of AgA and RtE. However, not AgA nor RtE directly apply to hard- and software design. Instead, they are part of the group's societal and judicial discourse. The primary responsibility of engineers consists of providing systems that operate according to the relevant formal specifications. In this context, AgT is the relevant principle. As indicated before, it is possible to formalize AgT to an extent that makes it applicable to algorithms. When implemented, it ensures the group can exercise AgA and RtE. Thus, we will base our framework primarily on the said formalization of AgT. As part of this framework, we will characterize the exact relation between AgT and integrity, c.f. Section III-C.

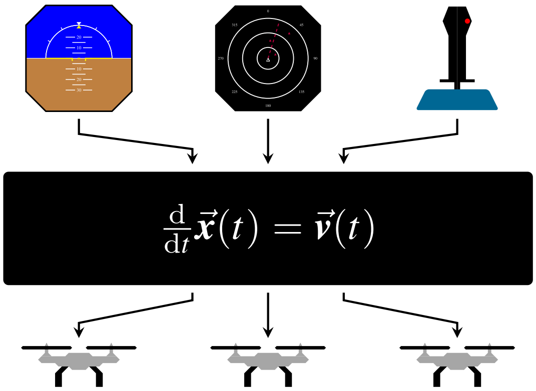

FIGURE 1. **Example of a *Physical Problem.* The flight-control computer of a drone network receives input data from different sources. Based on a mathematical model for flight dynamics, it processes the input into control outputs for the drones' engines such that the drones attain the desired flight path.**

B. PHYSICAL PROBLEMS: FORMALIZING THE ANALOG WORLD

Our formalization of AgT is best explained by a didactic example, as it facilitates defining the relevant terminology of *physical problems*. Having highlighted its relevance in the context of autonomous driving in Sections I and III-A already, we resort to another instance of trajectory planning for this purpose. Consider a scenario in which a network of drones is governed by a joint flight-control computer (from Section I, recall that low-altitude air traffic control is one of the envisioned services provided by 6G communication networks). The flight-control computer receives various inputs, ranging from attitude and navigational data to possible steering commands from a human operator. The computer processes these inputs into control outputs for the drones' engines, as visualized in Figure 1. For the remainder of the article, we refer to a scenario of this kind, i.e., one that encompasses the control of physical components, as a *physical problem*.

For now, we consider the flight-control computer a black box, ignoring any details regarding hard- or software. It is essential to note that the purpose of the flight-control computer *can* be specified in an entirely agnostic manner to these details. That is, the underlying physical problem exists as a mathematical model, regardless of the actual implementation within the computer. In the simplest case, a set of differential equations models the drones' flight dynamics. These differential equations characterize a relationship between the input of the flight-control computer and the desired control output for the drones' engines. In abstract terms, we summarize the nature of physical problems as follows.

Axiom 1: *A* physical problem *is characterized by an* input-output relation *that links elements of an input space to elements of an* output space.

Axiom 2: The input space *of a* physical problem *is characterized by a set of formalized* input attributes. *Each element of the* input space *consists of an individual expression of these attributes.*

Axiom 3: The output space *of a* physical problem *is characterized by a set of formalized* output attributes. *Each element of the* output space *consists of an individual expression of these attributes.*

In our example, the attitude and navigational data as well as the steering commands from a human operator form the input attributes of the physical problem. They are commonly expressed in terms of a real-valued tuple that specifies, e.g., points in 3D space, angles of alignment, etc. Hence, an expression of input attributes consists of a list of numerical values. Likewise, the control outputs for the drones' engines form the physical problem's output attributes, specifying voltage levels, for example. Again, an expression of these attributes consists of a list of numerical values.

Axioms 1, 2, and 3 apply directly to any engineering problem characterized by a mathematical model. The desired behavior of the technical system later implemented in practice is already entirely specified by the abstract physical problem. The action of a drone in the face of a collision with some other object, e.g., a building, a crewed aircraft, another drone, or a bird, is a result of mathematical modeling and criteria built thereupon. These criteria are visible, and any cyber-sovereign group can agree upon them. However, the subsequent process of hard- and software design gives rise to another nontrivial facet: Any real-world hardware platform is necessarily subject to mathematical modeling itself. In particular, the relevant mathematical model determines the class of algorithms that the platform can implement. In Section III-C, we discuss the relationship between physical problems and algorithmic implementations thereof in abstract terms, leading to the 4th and last axiom of our theory. In Section IV, we apply the established axioms to the mathematical model of Turing machines and deduce a necessary and sufficient condition for the existence of transparent algorithmic implementations in digital computing.

C. THE INVISIBLE LAYER: ALGORITHMIC TRANSPARENCY AND INTEGRITY

The characteristics of the relevant physical problem comprise the *visible layer* of a technological system. The details and specifications on this level are visible to the outside and thus provide the basis for AgT. Nevertheless, after fixing the visible specifications, the abstract physical problem has to be implemented through an algorithm and a suitable hardware platform. The details of this implementation form a technological system's *invisible layer*. If AgT is satisfied, the technological system operates exclusively according to the visible specifications, in which case the invisible layer is irrelevant to the user.

On the invisible layer, the algorithmic implementation must capture the characteristics of the *physical problem* through a suitable *machine-readable language*. As used in our context, the term "machine-readable language" refers to the specifications of how expressions of input and output attributes are represented on the hardware. It is not to be confused with an actual programming language. Notably, each expression of input and output attributes must possess adequate machine-readable *descriptions*. In order to illustrate this principle, consider Euler's number e. In mathematical terms, e exists as an abstract entity. Since e is an irrational number, we cannot store all of its digits in the memory of a (real-world) digital computer. However, we *can* store a finite source code that, when executed, accepts a natural number and returns as many decimal digits of e. The corresponding code uniquely determines the abstract object e. There exists an infinite variety of codes that determine e in this sense, each of which is a machine-readable *description* of e. The description details, e.g., the specific programming language or method we use to approximate e, are irrelevant, provided the description ultimately determines e in the above sense. In that case, we call the description a *feasible translation* of e.

Returning to the general principle of physical problems, each possible visible-layer expression of input attributes must possess feasible translations, i.e., corresponding machine-readable descriptions on the invisible layer. We will refer to these as *input descriptions* in the following. If the expressions of input attributes are real-valued tuples, an input description might consist of a source code that determines each of the tuple's components according to the abovementioned scheme. Whenever we present the algorithm on the invisible layer with an input description, it computes a description of some expression of output attributes. Analogously, we will refer to these as *output descriptions*.

In the context of CyS, relevant characteristics of the (physical) engineering problem need to be addressed at the visible layer. That is, all factors that influence the behavior of the implemented system need to be formalized in terms of one of the physical problem's input attributes. Given an input description of an expression of these attributes, the computed output description must unambiguously determine an expression of output attributes such that the physical problem's input-output relation is satisfied. Only if this is the case regardless of which specific input description we present to the algorithm, we can accept the physical problem's algorithmic implementation to be agnostic to any invisible-layer details. Accordingly, we abstract the principle of AgT as follows.

Axiom 4: Given a physical problem, we call an *algorithmic implementation* thereof *transparent* if it meets the following conditions:

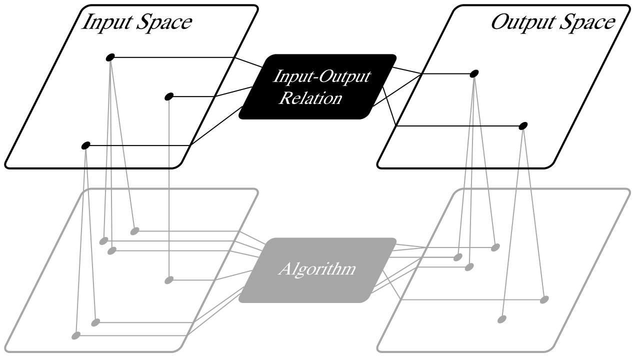

FIGURE 2. Layers of Technical-Systems Design. The *visible layer* (top) consists of the physical problem itself, while the *invisible layer* (bottom) consists of an algorithmic implementation thereof. Each expression of input and output attributes possesses corresponding machine-readable descriptions on the invisible layer. The algorithm transforms these descriptions according to the physical problem's input-output relation. Thus, if two points are connected on the visible layer, they must also be connected via a detour through the invisible layer.

(1) It is possible to translate any expression of input attributes into a corresponding machine-readable description.

(2) It is possible to retranslate any expression of output attributes unambiguously from each of its corresponding machine-readable descriptions.

(3) Based on the correspondence between visible-layer expressions and invisible-layer descriptions, the algorithm preserves the physical problem's input-output relation. Any feasible translation of an expression of input attributes is mapped to a feasible translation of the related expression of output attributes, regardless of the specific translation the algorithm receives.

Figure 2 visualizes the substance of Axiom 4. Observe that the question of what machine-readable language is deemed suitable to represent a physical problem is not determined a priori, and cannot be formalized in terms of a mathematical model. In any case, however, AgT requires that that invisible-layer details must not have an influence on the expression of output attributes corresponding to the computed output description. Such details include, the specific hardware model, the choice of (actual) programming language, the specific training data for machine-learning algorithms, and, most notably, the stopping criteria for iterative algorithms in numerical computing. Often, the latter are of the form that the computation is halted once the computed output value stops to change significantly for a prescribed number of successive iterations. These criteria are of heuristic nature and, in the context of our framework, make the computed expression of output attributes depend on the exact implementation of the numerical algorithm, down to the level of machine instructions. If AgT is required, such a stopping criterion is inadmissible. In the context of trajectory planning for autonomous driving, [3] highlighted this exact problem, which we will return to for a more in-depth discussion in Section V.

Finally, we can now identify the relationship between AgT and integrity. From Section III-A, recall that integrity refers to a state where a technological system resides within its design-based margin of operation. Regarding the automated control of a physical agent, integrity requires the relevant algorithm to capture the agent's (physical) dynamics correctly. Employing the nomenclature established within this article, the agent's dynamics and prescribed operation margin form the relevant physical problem, i.e., the visible-layer characteristics. The abovementioned requirement refers to the algorithm preserving the physical problem's input-output relation in the sense of Axiom 4. If satisfied, it ensures that the technological system does not exceed its operation margin due to uncontrollable variations in the invisible-layer details. Consequently, while stemming from two different contexts, AgT and integrity are ultimately paraphrasings of the same principle. The underlying formal criterion, as characterized by Axiom 4, provides an intersection of CyS and technological trustworthiness.

IV. IMPLICATIONS: MODELS OF COMPUTABILITY AND HARDWARE COMPATIBILITY ASSESSMENT

The principle of AgT characterizes a relationship between the visible and the invisible layer of a technological system. This relationship results from mathematical modeling and yields a formal property of algorithms. As indicated in Section III-B, the details of the invisible layer are subject to mathematical modeling themselves. Accordingly, the interplay of two mathematical models provides the basis of any assessment of AgT: The physical problem on the one side and the relevant computing model on the other. In particular, the computing model determines the class of algorithms the employed hardware platform can implement. This interplay ultimately determines whether the conditions of Axiom 4 can or cannot be satisfied.

Today, digital hardware provides the basis for most ICT systems. In theoretical computer science, the mathematical model of *Turing machines* [55], [56] is arguably the most well-established mathematical formalization of digital computing. The widely accepted *Church-Turing Thesis* concludes that Turing machines are a definitive model of digital computers, describing their (theoretical) capabilities perfectly. In mathematics, the study of the applications of Turing's theory to what we defined as *physical problems* is called *computable analysis*. For a comprehensive introduction, we refer to [57]. In this domain, the requirements of Axiom 4 are known and referred to as *Turing computability*. Accordingly, we obtain a necessary and sufficient condition for AgT in digital computing.

Thesis 1 (AgT in Digital Computing): A physical problem exhibits a transparent digital algorithmic implementation if and only if the relevant input-output relation is a Turing computable function.

Remark: Regarding Turing-computability, the nomenclature employed in the literature is partially inconsistent.

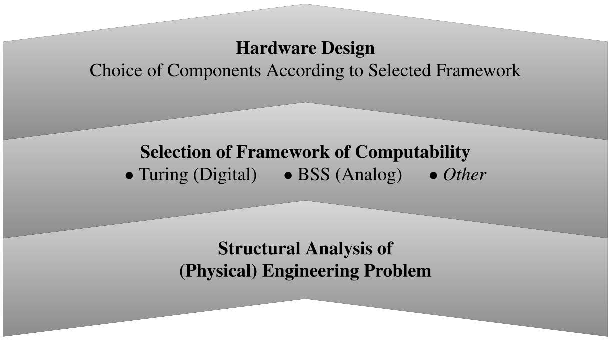
Hardware Design
Choice of Components According to Selected Framework
Selection of Framework of Computability
• Turing (Digital)
• BSS (Analog)
• Other
Structural Analysis of
(Physical) Engineering Problem

FIGURE 3. Paradigm of hard- and software design for ICT under transparency constraints. The engineer selects a framework of computability in which the relevant physical problem exhibits a transparent algorithmic implementation. The manufacturer then chooses a hardware platform that fits the selected framework of computability.

For example, [57] uses the term *effectively determined operator* for Turing-computable operators on Banach spaces.

Aside from Turing machines, the model of *Blum-Shub-Smale* (BSS) *machines* is, albeit rarely mentioned explicitly, one of the most common heuristic formalizations of digital computing (in contrast, Turing machines are considered a precise model according to the Church-Turing thesis) [58]. In a different context, BSS machines have recently gained interest due to their relation to analog hardware. It has, for example, been argued that they provide a suitable formalization of biocomputing [59]. The differences between both models of computability is present in theoretical applications relevant to 6G communications.

In general, every mathematical model of computing hardware induces a class of implementable algorithms. The authors conjecture that Thesis 1 extends to models other than the Turing machine upon proper formalization, adding a novel facet to hard- and software design for ICT in the context of societal and judicial constraints: Given a physical problem, specific hardware platforms may be required to achieve AgT. Whether a particular platform is suitable for this task depends on the structural properties of the physical problem, leading to the following design paradigm. Whenever a technological system needs to satisfy AgT, the design process must start with a structural mathematical analysis of the relevant physical problem, incorporating its computability within different hardware models. This analysis leads to choosing a hardware model in which the physical problem exhibits a transparent algorithmic implementation. Subsequently, the manufacturer must choose the physical hardware so that it sufficiently meets the characteristics of the selected hardware model. Figure 3 summarizes the established paradigm. Following the equivalence we derived in Section III-C, it may equally be applied to ensure integrity in the context of technological trustworthiness.

The established paradigm essentially utilizes the degrees of freedom provided by considering different models of computing hardware. Given the recent advances in the science of unconventional computing, we can expect to encounter various novel types of hardware in the near future of ICT. In this context, recall that BSS machines may serve as a suitable formalization of certain types of analog computing hardware. From a theoretical perspective, several potential application cases exist in which BSS machines exhibit supremacy of computing capabilities over Turing machines. The recent concept study [60] has extensively discussed this observation in the context of *virtual-twinning*, a newly emerging technique in control and systems engineering. Particularly, the study discussed two well-known engineering problems relevant to the field of ICT: Detecting unobservability in *remote state estimation* (RSE) – this problem is closely related to trajectory planning – detecting susceptibility to *denial-of-service* (DoS) attacks in network-resilience planning. Both physical problems consist of a mathematical model of wireless communication links and a go/no-go decision based on particular formal channel-quality measures. The resulting classification functions are computable within the theory of BSS machines but *not* within the theory of Turing machines [61], [62], [63]. Accordingly, given any corresponding algorithmic implementation on digital hardware, there always exists a set of channels for which the resulting system's behavior depends on details of the invisible layer. On the other hand, if Thesis 1 extends to BSS computability, transparent algorithmic implementations exist for both physical problems on hardware that fits the model of BSS machines.

An analogous type of computability supremacy occurs in the context of *inverse problems*, which form the mathematical basis for a broad range of applications in, among others, signal processing and machine learning [64], [65]. Finally, [66] highlighted that the conclusions on AgT drawn in [26] might analogously apply to near-future quantum hardware. The present work provides a generalized framework that allows for a coherent analysis of all application scenarios discussed above.

V. CONCLUSION AND PROSPECTS

The limitations of digital computing are evident regarding the societal and judicial requirements for near-future ICT systems. Concerning novel types of computing hardware, the theory established in the present article predicts substantial benefits. The authors conclude that the research on different frameworks of computability and different forms of computing hardware will develop to be significant for engineering in cyber-sovereign societies.

As indicated all throughout this article, specifications for trajectory planning under societal constraints – so far, primarily in the context of autonomous driving – already forms a major contemporary topic within the relevant engineering literature. In particular, [2], [4] discussed different societal and ethical theories and incorporated them into models for vehicular motion. The results in [2] were furthermore supported by numerical experiments. The problem of *technical feasibility* – that is, the question of how the decision rules derived from abstract ethical and societal principles are mapped into software – was discussed explicitly in [4]. However, both [2], [4] established their results exclusively in terms of real-valued physical problems, i.e., concerning the relevant technology's visible layer. Particularly, the supporting virtual experiments relied on heuristic numerics.

The ethical concerns emerging from heuristic numerics were highlighted in [3], indicating the expectable differences in system behavior per manufacturer, as well as the problem of how the software should determine if the accuracy of its trajectory calculation is sufficient to act upon it. In Section III-C, we have pointed to this issue in the context of stopping criteria for iterative methods. Often, these are of the form that the computation halts whenever the computed output value does not change significantly for a prescribed number of successive iterations. Expressed in the terminology of our theory, the computed expression of output attributes then depends on the exact details of the invisible layer, down to the level of machine instructions. The approach presented in this article avoids such issues entirely.

Finally, the authors conclude that the understanding of computability frameworks in the context of technology and sensitive human goods critically needs to be extended. Aside from the problem of remote state estimation discussed in Section IV, there exists a variety of other relevant (physical) engineering problems that have been shown to *not* be Borel-Turing computable. While Turing's theory is widely accepted as the definitive formalization of digital computing, equivalent theories do not exist for other types of hardware. In this regard, the BSS framework makes a promising starting point for developing a suitable model of several forms of analog computing. On the other hand, theories for neuromorphic hardware, for example, feature aspects that potentially go beyond the capabilities of BSS machines. In view of the contemporary advances in the manufacturing of unconventional computing hardware, the exploration of new computing theories for analog physical problems yields a promising field of further research. In addition, further research is necessary to formalize the societal and judicial constraints for ICT in the context of AI, such that a complete and comprehensive framework that integrates these constraints with the computing theory of different hardware platforms may be developed. A first step toward this goal was taken in [26].

REFERENCES

[1] P. Schwenteck, G. T. Nguyen, H. Boche, W. Kellerer, and F. H. P. Fitzek, "6G perspective of mobile network operators, manufacturers, and verticals," *IEEE Netw. Lett.*, vol. 5, no. 3, pp. 169–172, Mar. 2023.

[2] S. M. Thornton, S. Pan, S. M. Erlien, and J. C. Gerdes, "Incorporating ethical considerations into automated vehicle control," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 6, pp. 1429–1439, Jun. 2017.

[3] S. Karnouskos, "Self-driving car acceptance and the role of ethics," *IEEE Trans. Eng. Manag.*, vol. 67, no. 2, pp. 252–265, May 2020.

[4] M. Geisslinger, F. Poszler, J. Betz, C. Lüge, and M. Lienkamp, "Autonomous driving ethics: From trolley problem to ethics of risk," *Philosophy Technol.*, vol. 34, no. 4, pp. 2210–5441, Dec. 2021.

[5] G. P. Fettweis and H. Boche, "On 6G and trustworthiness," *Commun. ACM*, vol. 65, no. 4, pp. 48–49, Apr. 2022.

[6] *G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI*, OECD Publishing, Paris, France, 2023.

[7] Eur. Commission. (2024). *AI Act—Shaping Europe's Digital Future*. Accessed: Aug. 19, 2024. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[8] B. Goertzel, "Artificial general intelligence: Concept, state of the art, and future prospects," *J. Artif. Gen. Intell.*, vol. 5, no. 1, pp. 1–48, Dec. 2014.

[9] M. Roser. (2023). *AI Timelines: What Do Experts in Artificial Intelligence Expect for the Future?*. Our World Data. [Online]. Available: https://ourworldindata.org/ai-timelines

[10] G. Ras, N. Xie, M. Van Gerven, and D. Doran, "Explainable deep learning: A field guide for the uninitiated," *J. Artif. Intell. Res.*, vol. 73, pp. 329–397, Jan. 2022.

[11] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," in *Proc. ICLR*, Jan. 2014, pp. 1–11.

[12] Y. Zhang, Y. Li, L. Cui, D. Cai, L. Liu, T. Fu, X. Huang, E. Zhao, Y. Zhang, Y. Chen, L. Wang, A. T. Luu, W. Bi, F. Shi, and S. Shi, "Siren's song in the AI ocean: A survey on hallucination in large language models," 2023, *arXiv:2309.01219*.

[13] A. Bastounis, A. C. Hansen, and V. Vlačić, "The mathematics of adversarial attacks in AI—Why deep learning is unstable despite the existence of stable neural networks," 2021, *arXiv:2109.06098*.

[14] B. Adcock and N. Dexter, "The gap between theory and practice in function approximation with deep neural networks," *SIAM J. Math. Data Sci.*, vol. 3, no. 2, pp. 624–655, Jan. 2021.

[15] A. Ilyas, S. Santurkar, D. Tsipras, L. Engstrom, B. Tran, and A. Madry, "Adversarial examples are not bugs, they are features," in *Proc. Adv. Neural Inf. Process. Syst* Red Hook, NY, USA: Curran Associates, Jan. 2019, pp. 1–16.

[16] N. Carlini and D. Wagner, "Audio adversarial examples: Targeted attacks on speech-to-text," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2018, pp. 1–7.

[17] D. Tsipras, S. Santurkar, L. Engstrom, A. Turner, and A. Madry, "Robustness may be at odds with accuracy," in *Proc. ICLR*, Jan. 2018, pp. 1–9.

[18] V. Antun, F. Renna, C. Poon, B. Adcock, and A. C. Hansen, "On instabilities of deep learning in image reconstruction and the potential costs of AI," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 48, pp. 30088–30095, Dec. 2020.

[19] S.-M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "DeepFool: A simple and accurate method to fool deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2574–2582.

[20] A. Boulemtafes, A. Derhab, and Y. Challal, "A review of privacy-preserving techniques for deep learning," *Neurocomputing*, vol. 384, pp. 21–45, Apr. 2020.

[21] Y. He, G. Meng, K. Chen, X. Hu, and J. He, "Towards security threats of deep learning systems: A survey," *IEEE Trans. Softw. Eng.*, vol. 48, no. 5, pp. 1743–1770, May 2022.

[22] X. Liu, L. Xie, Y. Wang, J. Zou, J. Xiong, Z. Ying, and A. V. Vasilakos, "Privacy and security issues in deep learning: A survey," *IEEE Access*, vol. 9, pp. 4566–4593, 2021.

[23] F. Mireshghallah, M. Taram, P. Vepakomma, A. Singh, R. Raskar, and H. Esmaeilzadeh, "Privacy in deep learning: A survey," 2020, *arXiv:2004.12254*.

[24] O. Willers, S. Sudholt, S. Raafatnia, and S. Abrecht, "Safety concerns and mitigation approaches regarding the use of deep learning in safety-critical perception tasks," in *Proc. SAFECOMP Workshops*, A. Casimiro, F. Ortmeier, E. Schoitsch, F. Bitsch, and P. Ferreira, Eds., Cham, Switzerland: Springer, Jan. 2020, pp. 336–350.

[25] H. Tan, "A brief history and technical review of the expert system research," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 242, Sep. 2017, Art. no. 012111.

[26] H. Boche, A. Fono, and G. Kutyniok, "A mathematical framework for computability aspects of algorithmic transparency," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024, pp. 3089–3094.

[27] A. Alsajri and A. Steiti, "Intrusion detection system based on machine learning algorithms: (SVM and genetic algorithm)," *Babylonian J. Mach. Learn.*, vol. 2024, pp. 15–29, Jan. 2023.

[28] M. A. Khalaf and A. Steiti, "Artificial intelligence predictions in cyber security: Analysis and early detection of cyber attacks," *Babylonian J. Mach. Learn.*, vol. 2024, pp. 63–68, May 2024.

[29] E. M. Clarke, O. Grumberg, D. Peled, and D. A. Peled, *Model Checking* (The Cyber-Physical Systems Series). Cambridge, MA, USA: MIT Press, 1999.

[30] A. Bauer and D. S. Scott, "The realizability approach to computable analysis and topology," Ph.D. thesis, School Computer Science, Carnegie Mellon Univ., Pittsburgh, PA, USA, 2000.

[31] A. Mądry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *Proc. ICLR*, Jan. 2017, pp. 1–14.

[32] N. Carlini and D. Wagner, "Towards evaluating the robustness of neural networks," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2017, pp. 39–57.

[33] Z. C. Lipton, "The mythos of model interpretability," *Commun. ACM*, vol. 61, no. 10, pp. 36–43, 2018.

[34] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning," in *Proc. IEEE 5th Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Oct. 2018, pp. 80–89.

[35] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," 2017, *arXiv:1708.08296*.

[36] M. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2016, pp. 1135–1144.

[37] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," 2017, *arXiv:1702.08608*.

[38] A. Zou, Z. Wang, N. Carlini, M. Nasr, J. Zico Kolter, and M. Fredrikson, "Universal and transferable adversarial attacks on aligned language models," 2023, *arXiv:2307.15043*.

[39] J. Wang, X. Hu, W. Hou, H. Chen, R. Zheng, Y. Wang, L. Yang, H. Huang, W. Ye, X. Geng, B. Jiao, Y. Zhang, and X. Xie, "On the robustness of ChatGPT: An adversarial and out-of-distribution perspective," 2023, *arXiv:2302.12095*.

[40] J. A. Kroll, J. Huey, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu, "Accountable algorithms," *Univ. PA Law Rev.*, vol. 165, pp. 633–706, Jul. 2017.

[41] C. Olah. (2022). *Mechanistic Interpretability, Variables, and the Importance of Interpretable Bases*. Accessed: Oct. 19, 2024. [Online]. Available: https://www.transformer-circuits.pub/2022/mech-interp-essay

[42] L. Kästner and B. Crook, "Explaining AI through mechanistic interpretability," *Eur. J. Philosophy Sci.*, vol. 14, no. 4, p. 52, Dec. 2024.

[43] J. Angwin, J. Larson, S. Mattu, and L. Kirchner. (2016). *Machine Bias*. Accessed: Oct, 19, 2024. [Online]. Available: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[44] M. Veale and R. Binns, "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data," *Big Data Soc.*, vol. 4, no. 2, Dec. 2017, Art. no. 205395171774353.

[45] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," *ACM Comput. Surv.*, vol. 54, no. 6, pp. 1–35, Jul. 2022.

[46] S. Vadhan, "The complexity of differential privacy," in *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, Y. Lindell, Ed., Cham, Switzerland: Springer, 2017, pp. 347–450.

[47] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning: Limitations and Opportunities*. Cambridge, MA, USA: MIT Press, 2023.

[48] S. Wachter, B. Mittelstadt, and L. Floridi, "Why a right to explanation of automated decision-making does not exist in the general data protection regulation," *Int. Data Privacy Law*, vol. 7, no. 2, pp. 76–99, May 2017.

[49] D. K. Citron and F. Pasquale, "The scored society: Due process for automated predictions," *Wash. L. Rev.*, vol. 89, pp. 1–34, Jan. 2014.

[50] S. Fazelpour and Z. C. Lipton, "Algorithmic fairness from a non-ideal perspective," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, Feb. 2020, pp. 57–63.

[51] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena, "AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," *Minds Mach.*, vol. 28, no. 4, pp. 689–707, Dec. 2018.

[52] A. Biondi, F. Nesti, G. Cicero, D. Casini, and G. Buttazzo, "A safe, secure, and predictable software architecture for deep learning in safety-critical systems," *IEEE Embedded Syst. Lett.*, vol. 12, no. 3, pp. 78–82, Sep. 2020.

[53] H. Zhang, H. Chen, C. Xiao, S. Gowal, R. Stanforth, B. Li, D. S. Boning, and C.-J. Hsieh, "Towards stable and efficient training of verifiably robust neural networks," in *Proc. ICLR*, 2020, pp. 1–11.

[54] M. Mirman, A. Hägele, P. Bielik, T. Gehr, and M. Vechev, "Robustness certification with generative models," in *Proc. 42nd ACM SIGPLAN Int. Conf. Program. Lang. Design Implement.*, New York, NY, USA, Jun. 2021, pp. 1141–1154.

[55] A. M. Turing, "On computable numbers, with an application to the entscheidungsproblem," *Proc. London Math. Soc.*, vol. 42, no. 1, pp. 230–265, 1937.

[56] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem. A correction," *Proc. London Math. Soc.*, vol. 43, no. 1, pp. 544–546, 1938.

[57] M. B. Pour-El and J. I. Richards, *Computability in Analysis and Physics*. Berlin, Germany: Springer, 1989.

[58] L. Blum, "Computing over the reals: Where Turing meets Newton," *Notices Amer. Math. Soc.*, vol. 51, no. 9, pp. 1024–1034, Jan. 2004.

[59] L. Grozinger, M. Amos, T. E. Gorochowski, P. Carbonell, D. A. Oyarzún, R. Stoof, H. Fellermann, P. Zuliani, H. Tas, and A. Goñi-Moreno, "Pathways to cellular supremacy in biocomputing," *Nature Commun.*, vol. 10, no. 1, p. 5250, Nov. 2019.

[60] Y. N. Böck, H. Boche, R. F. Schaefer, F. H. P. Fitzek, and H. V. Poor, "Virtual-twin technologies in networking," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 136–141, Nov. 2023.

[61] H. Boche, Y. N. Böck, and C. Deppe, "On the semidecidability of the remote state estimation problem," *IEEE Trans. Autom. Control*, vol. 68, no. 3, pp. 1708–1714, Mar. 2023.

[62] H. Boche, Y. Böck, C. Deppe, and F. H. P. Fitzek, "Remote state estimation and Blum-Shub-Smale machines—A computability analysis with applications to virtual-twinning," *IEEE Trans. Autom. Control*, early access, Nov. 19, 2024, doi: 10.1109/TAC.2024.3502314.

[63] H. Boche, R. F. Schaefer, H. V. Poor, and F. H. P. Fitzek, "On the need of neuromorphic twins to detect denial-of-service attacks on communication networks," *IEEE/ACM Trans. Netw.*, vol. 32, no. 4, pp. 2875–2887, Aug. 2024.

[64] H. Boche, A. Fono, and G. Kutyniok, "Limitations of deep learning for inverse problems on digital hardware," *IEEE Trans. Inf. Theory*, vol. 69, no. 12, pp. 7887–7908, Dec. 2023.

[65] H. Boche, A. Fono, and G. Kutyniok, "Inverse problems are solvable on real number signal processing hardware," *Appl. Comput. Harmon. Anal.*, vol. 74, Jan. 2025, Art. no. 101719.

[66] H. Boche, Y. N. Böck, Z. G. del Toro, and F. H. P. Fitzek, "Feynman meets Turing: The uncomputability of quantum gate-circuit emulation and concatenation," *IEEE Trans. Comput.*, early access, Nov. 27, 2024, doi: 10.1109/TC.2024.3506861.

YANNIK N. BÖCK (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electronics engineering and information technology from Technische Universität München (TUM), Munich, Germany, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Chair of Theoretical Information Technology, TUM.

Since 2019, he has been a Member of the Research and Teaching Staff of the Chair of Theoretical Information Technology, TUM. His research interests include quantum information theory, discrete mathematics, and applications of computability theory in engineering.

HOLGER BOCHE (Fellow, IEEE) received the Dipl.-Ing. degree in electrical engineering from Technische Universität Dresden in 1990, the Dipl.-Math. degree in mathematics from Technische Universität Dresden in 1992, the Dr.-Ing. degree in electrical engineering from Technische Universität Dresden in 1994, the Postgraduate Studies from Friedrich-Schiller Universität Jena in 1997, and the Dr. rer. nat. degree in pure mathematics from Technische Universität Berlin in 1998.

In 1997, he joined the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institute (HHI), Berlin, Germany. From 2002 to 2010, he was a Full Professor of mobile communication networks with the Institute for Communications Systems, Technische Universität Berlin. In 2003, he became the Director of the Fraunhofer German-Sino Laboratory for Mobile Communications, Berlin, and in 2004, he became the Director of the Fraunhofer Institute for Telecommunications, HHI. He was a Visiting Professor with ETH Zurich, Zürich, Switzerland, from 2004 to 2006, and KTH Stockholm, Stockholm, Sweden, in 2005. Since 2010, he has been a Full Professor with the Chair of Theoretical Information Technology, Technische Universität München, Munich, Germany. Since 2014, he has been a member and a Honorary Fellow of the TUM Institute for Advanced Study, Munich, and since 2018, he has been the Founding Director of the Center for Quantum Engineering, Technische Universität München. Since 2021, he has been leading jointly with Frank Fitzek the BMBF Research Hub 6G-Life. Among his publications is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017). He is a member of the IEEE Signal Processing Society SPCOM and SPTM Technical Committees. He was an Elected Member of German Academy of Sciences (Leopoldina) in 2008 and Berlin Brandenburg Academy of Sciences and Humanities in 2009. He was a recipient of the Research Award ‘‘Technische Kommunikation’’ from the Alcatel SEL Foundation in October 2003, the ‘‘Innovation Award’’ from the Vodafone Foundation in June 2006, and the Gottfried Wilhelm Leibniz Prize from the Deutsche Forschungsgemeinschaft (German Research Foundation) in 2008. He was a co-recipient of the 2006 IEEE Signal Processing Society Best Paper Award and a recipient of the 2007 IEEE Signal Processing Society Best Paper Award. He was the General Chair of the Symposium on Information Theoretic Approaches to Security and Privacy at IEEE GlobalSIP 2016.

FRANK H. P. FITZEK (Senior Member, IEEE) received the Diploma (Dipl.-Ing.) degree in electrical engineering from Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen, Germany, in 1997, the Ph.D. (Dr.-Ing.) degree in electrical engineering from Technical University Berlin, Germany, in 2002, and the Honorary (Doctor Honoris Causa) degree from Budapest University of Technology and Economics (BUTE), in 2015.

He became an Adjunct Professor with the University of Ferrara, Italy, in 2002. In 2003, he joined Aalborg University, as an Associate Professor and later became a Professor. He co-founded several start-up companies starting with Acticom GmbH, Berlin, in 1999. He visited various research institutes, including Massachusetts Institute of Technology (MIT), VTT, and Arizona State University. He is currently a Professor and the Head of the Deutsche Telekom Chair of Communication Networks, TU Dresden, coordinating the 5G Laboratory, Germany. Furthermore, he is the spokesman of the DFG Cluster of Excellence CeTI. His research interests include wireless and 5G communication networks, network coding, cloud computing, compressed sensing, cross layer, and energy efficient protocol design and cooperative networking. In 2005, he won the YRP Award for the work on MIMO MDC and received the Young Elite Researcher Award of Denmark. He was selected to receive the NOKIA Champion Award several times in a row from 2007 to 2011. In 2008, he was awarded the Nokia Achievement Award for his work on cooperative networks. In 2011, he received the SAPERE AUDE Research Grant from the Danish Government, and in 2012, he received the Vodafone Innovation Prize.

GITTA KUTYNIOK (Fellow, IEEE) received the Diploma degree in mathematics and computer science and the Ph.D. degree from the Universität Paderborn, Germany, and the Habilitation degree in mathematics from the Justus-Liebig Universität Gießen, in 2006.

From 2001 to 2008, she held visiting positions at several U.S. institutions, including Princeton University, Stanford University, Yale University, Georgia Institute of Technology, and Washington University in St. Louis. In 2008, she became a Full Professor of mathematics with the Universität Osnabrück, and moved to Berlin three years later, where she held the Einstein Chair of the Institute of Mathematics, Technische Universität Berlin, and a courtesy appointment with the Department of Computer Science and Engineering, until 2020. In 2023, together with colleagues, she founded the start-up EcoLogic Computing GmbH. She is currently with the Bavarian AI Chair for Mathematical Foundations of Artificial Intelligence, Ludwig-Maximilians-Universität München. In addition, she is affiliated with German Aerospace Center (DLR) and the University of Tromsø. She also acts as the LMU-Director of the Konrad Zuse School of Excellence in Reliable AI (relAI), Munich, and the spokesperson of the DFG-Priority Program ‘‘Theoretical Foundations of Deep Learning’’ and of the AI-HUB@LMU, which is the interdisciplinary platform for research, teaching, and transfer in AI and data science at LMU. Her research interests include applied and computational harmonic analysis, artificial intelligence, compressed sensing, deep learning, imaging sciences, inverse problems, and applications to life sciences, robotics, and telecommunication.

Dr. Kutyniok received various awards for her research, such as the award from the Universität Paderborn in 2003, the Research Prize of the Justus-Liebig Universität Gießen, and the Heisenberg-Fellowship in 2006, and the von Kaven Prize by DFG in 2007. She was invited as the Noether Lecturer at the ÖMG-DMV Congress in 2013, a Plenary Lecturer at the 8th European Congress of Mathematics (8ECM) in 2021, and the Lecturer of London Mathematical Society (LMS) Invited Lecture Series in 2022. She was also honored by invited lectures at both the International Congress of Mathematicians 2022 (ICM 2022) and the International Congress on Industrial and Applied Mathematics (ICIAM 2023). Moreover, she was elected as a member of Berlin-Brandenburg Academy of Sciences and Humanities in 2017 and European Academy of Sciences in 2022, became a SIAM Fellow in 2019 and an IEEE Fellow in 2024, and served as the Vice President-at-Large for SIAM from 2021 to 2023.