Review

A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions

Ali Mohammad Saghiri [1,*] **, S. Mehdi Vahidipour** [2] **, Mohammad Reza Jabbarpour** [3] **, Mehdi Sookhak** [4]
and Agostino Forestiero [5]

[1] Soft Computing Lab, Computer Engineering and Information Technology Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran 1591634311, Iran

[2] Computer Engineering Department, University of Kashan, Kashan 8731753153, Iran; vahidipour@kashanu.ac.ir

[3] Information and Communications Technology Research Department, Niroo Research Institute, Tehran 1468613113, Iran; mrjabbarpour@nri.ac.ir

[4] Department of Computer Science, Texas A&M University, Corpus Christ, TX 78412, USA; mehdi.sookhak@tamucc.edu

[5] Institute for High Performance Computing and Networking (ICAR), National Research Council of Italy (CNR), 00185 Rende, CS, Italy; agostino.forestiero@icar.cnr.it

* Correspondence: saghiri@aut.ac.ir

Abstract: In recent years, artificial intelligence has had a tremendous impact on every field, and several definitions of its different types have been provided. In the literature, most articles focus on the extraordinary capabilities of artificial intelligence. Recently, some challenges such as security, safety, fairness, robustness, and energy consumption have been reported during the development of intelligent systems. As the usage of intelligent systems increases, the number of new challenges increases. Obviously, during the evolution of artificial narrow intelligence to artificial super intelligence, the viewpoint on the challenges such as security will be changed. In addition, the recent development of human-level intelligence cannot appropriately happen without considering whole challenges in designing intelligent systems. Considering the mentioned situation, no study in the literature summarizes the challenges in designing artificial intelligence. In this paper, a review of the challenges is presented. Then, some important research questions about the future dynamism of challenges and their relationships are answered.

Keywords: artificial intelligence; artificial superintelligence; artificial general intelligence; artificial narrow intelligence; human-level intelligence; challenges

1. Introduction

Artificial intelligence (AI) has been widely used in recent years. In the literature, numerous articles, such as [1,2], only focus on AI's extraordinary capabilities. Few papers, such as those reported in [3–7], focused on the challenges of AI. During analyses of the challenges of AI, numerous problems have been reported in the literature, some of which are security [8], safety [9], fairness [10], energy consumption [11], and ethics [12], to mention a few. The widespread usage of AI leads to arising new challenges. This issue becomes more complicated when the definitions of challenges are modified based on new dimensions explained in the next paragraph.

From a historical point of view, the evolution of AI-based systems starts with artificial narrow intelligence (ANI), then continues with artificial general intelligence (AGI), and finally meets artificial super intelligence (ASI), which will surpass human capabilities in all dimensions [13,14]. All of the mentioned terms will be explained in the rest of this paper. Defining the boundaries between different definitions of AI and new concepts is not an easy task. In addition, we are faced with other definitions in this area, such as human-level

intelligence (HLI), which defines an agent that is equivalent to a human agent in terms of thinking and acting capabilities. The main problem is that there is no comprehensive study in the literature for discussing AI challenges that address some research questions about the evolution of challenges, and also their relationships considering different classes of AI.

In this paper, the challenges of AI are discussed with a particular focus on HLI. We gathered 28 challenges from the literature, which makes this paper unique because of the high coverage of challenges. For each challenge, the definition and its changes considering some dimensions in ANI, AGI, ASI, and HLI are discussed. After summarizing the challenges, some research questions about AI development considering challenges are answered to give a better understanding of the future development of AI. To the best of our knowledge, there is no comprehensive study on challenges that focus on the relation among challenges and their evolution considering different classes of AI such as HLI in the literature.

The rest of the paper is organized as follows. Section 2 is dedicated to preliminaries. In Section 2, different classes of AI and related terms are explained. Section 3 focuses on challenges. Section 4 is dedicated to a detailed discussion on challenges based on some research questions. The last section is dedicated to conclusions.

2. Preliminaries

Since this paper focuses on the challenges of AI considering different terms such as ANI, AGI, ASI, and HLI, this section is dedicated to the definition of mentioned terms and their relationships. This section starts with a classification of AI to state the position of existing studies on AI, and then focuses on HLI because of its priority. According to [15], intelligent systems are classified into three classes, as explained below (Figure 1):

- **ANI:** This type of intelligence refers to intelligent systems that do specific tasks. For example, an agent with capabilities such as face recognition and games playing. These agents are programmed to do tasks and cannot detect and formulate unknown tasks in a self-organized manner. We do not expect to see self-awareness in these agents.
- **AGI:** The concept of this type of intelligence does not refer to a unique thing in the mind of all leading scientists of AI. Most researchers use AGI for those agents whose intelligence is equivalent to human agents. AGI can be equivalent to HLI [14].
- **ASI:** In [16], Bostrom introduced three types of super intelligence: Speed ASI, collective ASI, and quality ASI. Speed ASI refers to an agent faster than a human, collective ASI refers to decision-making capabilities similar to a group of humans, and quality ASI refers to an agent that can do work that humans cannot.
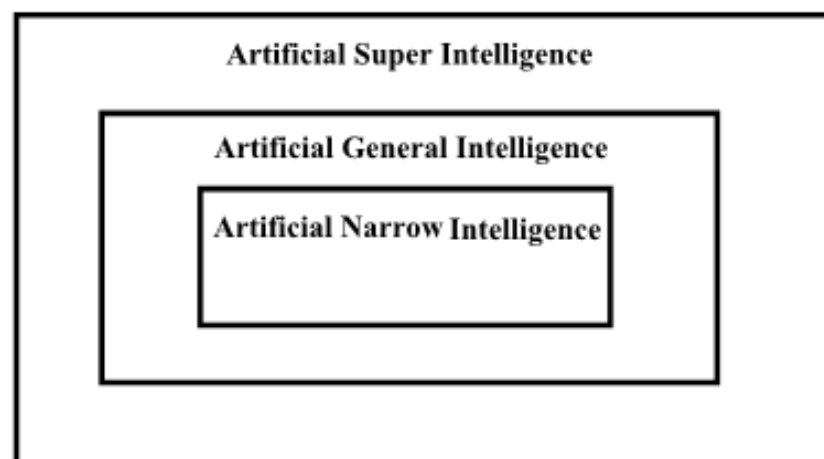

Figure 1. Venn diagram for definitions of artificial intelligence (AI).

Recently, some changes have been made in the above classification. In [17,18], the authors argue that HLI is different from AGI because humans may put some assumptions

and limitations in the machine's computations. These assumptions and limitations come from the nature of humans which inherited by the machines. Therefore, AGI may not solve a wide range of problems that humans cannot solve. In other words, humans determine some implicit upper bounds in the machines and downgrade the generalization capabilities.

On the other hand, authors of some papers, such as [18], argue that there is no difference between AGI and ASI when the definitions of AGI are not limited to HLI. According to [16,18], if AGI-based agents' capabilities are further than human intelligence and there is no exact definition for ASI capabilities, there is no need to differentiate AGI from ASI. Accordingly, in [19], Searle divided AI into two classes: weak AI and strong AI. Since the challenges related to HLI will be vital in the near future, the rest of this part is dedicated to HLI.

HLI is developed based on knowledge about humans, as illustrated in Figure 2. This figure is divided into: (1) human intelligence, (2) sciences related to HLI development, and (3) challenges. The human intelligence part divides the domain of our knowledge into *Known* (for example, accounting is a known capability of humans), *Semi-Known* (for example, mind computation is a semi-known capability of humans), and finally *Unknown* (for example, goals of human creation is an unknown concept for humans) sections. Many sciences were organized based on observing human intelligence, including mathematics and philosophy. According to [20], several challenges have been raised during the development of these sciences at the heart of AI-based systems. Thus, the third part of Figure 2 summarizes the challenges of AI during the development of HLI. It should be noted that providing connections between challenges and their related sciences is based on articles reported in the literature. The number of related sciences may increase as the knowledge of humans about AI increases.
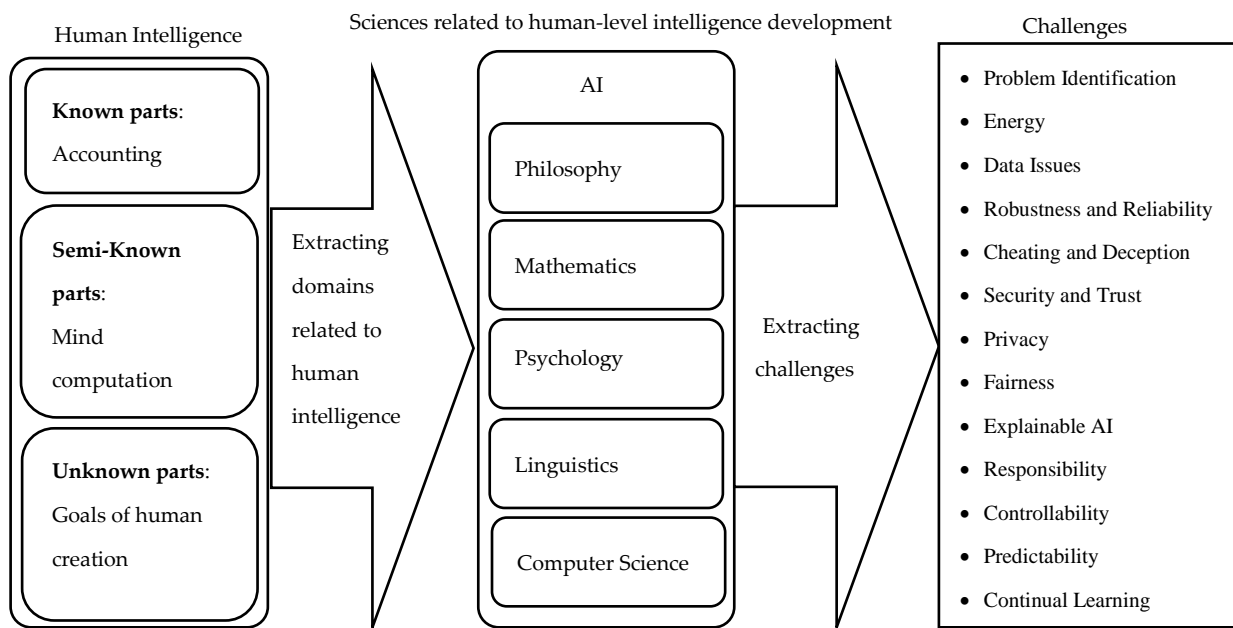
Figure 2. Challenges and their origins.

3. Analyzing Challenges

In this part, we focus on the challenges of AI, as given in Table 1. For each challenge, the required definitions are summarized. In each challenge, we mainly focused on the research and development related to HLI. Since there are few papers on the relationship between challenges and the classes of AI, we tried to find possible connections to ANI, AGI, and ASI for some challenges. In the rest of this section, each challenge is explained independently. The connections between challenges and also their evolution will be discussed in Section 4.

Table 1. List of challenges.

No.	Challenges	No.	Challenges
1	Problem Identification and Formulation	2	Energy Consumption
3	Data Issues	4	Robustness and Reliability
5	Cheating and Deception	6	Security
7	Privacy	8	Fairness
9	Explainable AI	10	Responsibility
11	Controllability	12	Predictability
13	Continual Learning	14	Storage (limited memory)
15	Semantic and Communication	16	Morality and Ethics
17	Rationality	18	Mind
19	Accountability	20	Transparency
21	Reproducibility	22	Evolution
23	Beneficial	24	Exploration and Exploitation Balance
25	Verifiability	26	Safety
27	Complexity	28	Trustworthy

3.1. Problem Identification and Formulation

Problem identification and formulation are essential processes that should be implemented in AI-based agents. This issue plays an essential role in organizing HLI-based agents designated in a self-organized manner. Implementing the mentioned processes is very difficult in practice because of some issues explained as follows. As was previously mentioned, we know that human knowledge about its environment can be found in different domains, including known parts, semi-known parts, and unknown parts. There is a set of problems that cannot be formulated in a well-defined format for humans, and therefore there is uncertainty as to how we can organize HLI-based agents to face these problems. For example, in [21], it was proved that there are some types of problems with no algorithm for solving. In addition, some problems for humans, such as the goals of human creation, are not clear, and therefore HLI-based agents will not be able to solve these types of problems because they follow the thinking process of humans. This problem becomes more complicated when the environment of an HLI-based agent is dynamic and unknown. Implementing a problem detection and formulation system at the heart of an HLI-based agent will be required. The role of problem identification is discussed in the following paragraph.

According to [20], a method for solving a problem is to map the problem space into a search space. In this method, detecting the problem, which can be single-state, multiple-state, or contingency, and then formulating it in a well-defined manner are the first steps of AI-based problem-solving. In other words, gathering knowledge and insides from the nature of the problem and converting it to a systematic solving procedure has high priority in this regard. In existing intelligent agents, the process of formulating problems considering security, safety, performance, etc., is considered as the responsibility of designers of AI-based agents. To solve this problem, knowledge-based agents and inference engines such as the Cyc project [22] can be utilized to present some solutions.

3.2. Energy Consumption

Some learning algorithms, including deep learning, utilize iterative learning processes [23]. This approach results in high energy consumption. Nowadays, the deep learning method is used to design HLI-based agents because of its high accuracy and its similarity to the human brain in the decision-making process. Deep learning models require a high computational power of GPUs. In [11], it was shown that these models are costly to train and develop from financial and energy consumption perspectives. During the operation of an HLI-based agent, the agent may use a predefined plan for learning multiple models concurrently to support self-awareness. Therefore, high computational power is required to support more abilities of cognition. This means that enough energy

should be supplied for executing the HLI-based agent. In order to mitigate this problem, four solutions were given in the literature, as given below:

- Investing in new paradigms with low energy consumption for HLI, such as quantum computing [24].
- Finding modern mathematical frameworks to find learning models with lower calculations, which leads to lower energy consumption [25].
- Sharing models to prevent energy consumption. A researcher can share a model with other researchers around the world.
- In an AI-based system, if there is no way to decrease the load of computations of learning processes, energy harvesting techniques can be used to return the wasted energy [26,27].

3.3. Data Issues

A type of AI-based agent invests in data-driven methods to construct learning models. In these algorithms, data issues cause various problems. Some of these problems are explained below [28–32]:

- Cost is one of the main issues of data. Major sources of cost are gathering, preparing, and cleaning the data [28].
- The size of collected data in a wide range of systems such as IoT (Internet of Things) is another data-related challenge. This huge amount of data leads to a new concept, called big data. Analyzing big data in an online fashion via machine learning algorithms is a very challenging task [28–30].
- Data incompleteness (or incomplete data) is another challenging problem in machine learning algorithms which leads to inappropriate learning of algorithms and uncertainties during data analysis. This issue should be handled during the pre-processing phase. Various approaches can be used for mitigating this problem. Filling missed (incomplete) data via most frequently observed values or developing learning algorithms to predict missed values are some examples of these approaches [32].
- Data heterogeneity, data insufficiency, imbalanced data, untrusted data, biased data, and data uncertainty are other data issues that may cause various difficulties in data-driven machine learning algorithms [28,29].
- Bias is a human feature that may affect data gathering and labeling. Sometimes, bias is present in historical, cultural, or geographical data. Consequently, bias may lead to biased models which can provide inappropriate analysis. Despite being aware of the existence of bias, avoiding biased models is a challenging task. For more information, please refer to [28].

Not only the mentioned problem, but also other problems such as imbalanced data, synthetic (fake) data, and noisy data can be classified in this part. Data issues become more critical challenges in designing HLI-based agents. Nowadays, finding datasets that reflect the correct human behaviors might not be possible in many of the domains that HLI-based agents can be utilized.

3.4. Robustness and Reliability

The robustness of an AI-based model refers to the stability of the model performance after abnormal changes in the input data. The cause of this change may be a malicious attacker, environmental noise, or a crash of other components of an AI-based system [8,31,32]. For example, in telesurgery, an HLI-based agent may detect a patient's kidney as a bean because of an unknown crash in the machine vision component. Among the several models with similar performance, the robust model has a higher priority in deployment. The traditional mechanisms such as replication and multi-version programming might not work in intelligent systems, and hence this field is in its early stage. Some works, such as [33], discuss the difference between the accuracy of a learning model and the robustness of that model. It seems that theory and concepts of robustness and reliability are in infancy, and new things would appear in this regard. This problem may be challenging in HLI-based

agents because weak robustness may have appeared in unreliable machine learning models, and hence an HLI with this drawback is error-prone in practice.

3.5. Cheating and Deception

The mentioned behaviors refer to two well-known behaviors of humans, but they may appear from intelligent agents such as HLI-based agents. Although many papers such as those papers reported in [34–36] are for detecting this behavior in humans using AI-based systems, few articles, such as [37,38], are conducted on analyzing cheating and deception in AI-based systems. In [37,38], the role of deception in multi-agent systems and its effects are studied. Since HLI-based agents are going to mimic the behavior of humans, they may learn these behaviors accidentally from human-generated data. It should be noted that deception and cheating maybe appear in the behavior of every computer agent because the agent only focuses on optimizing some predefined objective functions, and the mentioned behavior may lead to optimizing the objective functions without any intention. In order to design a HLI-based agent, the mentioned problems must be considered. For the sake of more clarification, some explanations are given as follows. After the invention of some modes, such as generative adversarial networks (GANs) [39], machines can generate data samples. This capability may be used to create fake news, pictures, and videos [40]. This capability may be used to design many cognitive abilities for HLI. Some of these abilities may not be controlled by humans when the HLI is used to construct a self-organized agent. For example, an HLI-based agent may be able to tell lie and generate real evidence for it. In other words, the inference engine of an intelligent agent might be able to generate any type of facts about data (real/synthetic), and humans would not be able to differentiate between real and synthetic data. This problem might be very challenging when machines try to replace real data with synthetic ones, leading to vital decision-making in human societies.

3.6. Security

AI-based systems are widely used to design secure methods, such as those reported in [41–43], but from another point of view, it is obvious that every piece of software, including learning systems, may be hacked by malicious users [44–46]. In designing intelligent systems, the security challenge is a critical issue that has received much attention. For example, consider an ant-based path planning in which the pheromone update function is hacked to manipulate the pathfinding process. Security challenges in AI may bring several new challenges that cannot be covered in this paper, and for more information, please refer to [46]. A security dimension is covered in the next paragraph, focusing on securing data-driven machine learning.

In data-driven machine learning, AI system developers might want to reverse-engineer the training data or learn how to develop a model that creates the desired output [49]. For instance, a neural network that utilizes untrusted (synthetic) data can be considered as an untrusted learning model in data-driven machine learning methods. Most AI researchers and developers neglect this issue. Adversarial machine learning can be considered as the first attempt to solve some security problems in machine learning [47]. It is worth mentioning that at the same time, with the evolution of attacks, security mechanisms should also be evolved. Recently, AI algorithms have been utilized by attackers to organize attacks. Hence, AI-based defense mechanisms must be applied to enhance the security of AI-based systems [48]. This problem becomes more challenging with the creation of HLI-based agents as malicious agents, because humans will not be able to organize defense mechanisms faster than those agents.

3.7. Privacy

Users' data, including location, personal information, and navigation trajectory, are considered as input for most data-driven machine learning methods. Data will be the most important thing in different types of AI in the next century. As explained in the

next paragraph, this challenge can be viewed from two perspectives: data factors and human factors.

Each fragment of data carries a different form of information which should be protected considering specific factors. In [49], three factors, persistence, repurposing, and spillovers related to data, are explained to highlight the characteristics of appropriate protection mechanisms in the AI era. On the other hand, data owner, data manipulator, and data visualizer are three typical roles that should be determined during the execution of machine learning methods. Preserving privacy considering different roles requires more effort and consideration by researchers [50]. Federated learning is one of these efforts [50]. This type of learning algorithm trains an algorithm across multiple decentralized computational resources, thus addressing critical issues such as data privacy, data security, and data access rights. During the design of an HLI-based agent, the definition of privacy issues may be somehow complex and complicated, whether humans accept privacy for other intelligent entities or not. How an HLI-based agent can be trained without the private data of humans is an important question.

3.8. Fairness

This challenge appears when the learning model leads to a decision that is biased to some sensitive attributes such as race, skin color, gender, religion, national origin, citizenship, age, pregnancy, familial status, disability status, veteran status, and genetic information [51]. The literature review on fairness in AI can be divided into three different approaches, as explained below [52]:

- In the first approach, data itself could be biased, which results in unfair decisions. Therefore, this problem should be solved on the data level and as a preprocessing step [53–55]. In [56], the problems of datasets and their issues that lead to unfair results are discussed. Some clues for preparing appropriate versions of existing datasets that were created during the last decade are given.
- In the second approach, the fairness criteria can be satisfied by some manipulation of the model after learning to attain a fair model [57–60].
- In the third approach, a process along with a straining procedure is conducted to satisfy fairness constraints by imposing them as a constraint to the main learning objective [61–64].

There are no exact solutions for implementing fair behavior in human societies, and we know the history of humans is full of unfair activities. Therefore, a massive amount of data for training machines may lead to the design of learning systems that are not fair; when there is no appropriate dataset and theory, how can an HLI-based agent be trained? This capability may be vital in some situations during the operation of an HLI-based agent when the agent has a responsibility to consider sensitive attributes.

3.9. Explainable AI

Explainable AI is an emerging field with many applications in different domains, including healthcare, transportation, and military services [65,66]. In this field, a set of tools and processes may be used to bring explainability to a learning model. With such capability, humans may trust the decisions made by the models from different points of view, including bias and fairness challenges to mention a few [67]. This means that explainability may determine some solutions for other challenges, such as fairness and trustworthiness, which will be discussed in Section 4.

Many learning methods, such as neural networks and deep learning, have extraordinary capabilities, but they invest in non-explainable symbols to do tasks. In many situations, including mission-critical tasks, we need to know the rationale behind decision-making in the intelligent system, and hence explainable AI can be useful [67]. In [68,69], recent developments and applications in explainable learning algorithms in practice have been summarized. This challenge becomes more critical when a human agent is replaced with

an HLI-based agent to do a critical task in healthcare, military, or other mission-critical situations. In these situations, any sort of decision must be explainable.

3.10. Responsibility

In the last century, humans have acted as operators of machines. In this situation, morality and legality issues associated with the responsibility of human agents are clear to all of us. However, this situation is going to be changed considering fully autonomous machines. Autonomous machines based on neural networks, genetic algorithms, and learning automata bring new problems where humans can no longer predict future machine activities [70]. As the next paragraph explains, this issue becomes more critical in designing HLI.

HLI-based systems such as self-driving drones and vehicles will act autonomously in our world. In these systems, a challenging question is "who is liable when a self-driving system is involved in a crash or failure?". Various aspects and domains should be considered in the answer to this question. Dual use is another challenge that comes from the abuse of the HLI-based system. Like other software, these intelligent systems can be used for good or evil purposes. For instance, publicly released sound imitation software can be used by malicious agents to mimic the voice of bank clients and execute a phone call on their behalf. Further discussions on responsibility issues can be found in [71].

3.11. Controllability

In computability theory, the halting problem is the problem of determining, from a description of an arbitrary computer program and an input, whether the program will finish running or continue to run forever. In 1936, Alan Turing proved that a general algorithm to solve the halting problem for all possible program-input pairs cannot exist [72,73]. For more clarification, consider an AI-based agent, called x, that controls the input, computation, and output of another agent called y. Therefore, there is no general algorithm to program agent x. Informally, some parts of AI control problems that can be reduced to halting problems are not considered as solvable problems [73]. If we accept this problem is solvable, we must accept that a wide range of problems, such as halting problems, are solvable.

In the era of superintelligence, the agents will be difficult to control for humans [74,75]. Note that controllability according to its context has many dimensions, and few papers, such as [76], focus on this problem in a specific domain such as safety. In [76], it is shown that controllability has four types: explicit, implicit, delegated, and aligned. It is shown that this problem is not solvable considering safety issues, and will be more severe by increasing the autonomy of AI-based agents. Therefore, because of the assumed properties of HLI-based agents, we might be prepared for machines that are definitely possible to be uncontrollable in some situations.

3.12. Predictability

One important challenge which may not be solved is whether the decision of an AI-based agent can be predicted in every situation or not [77–79]. A direct consequence of this challenge is that we can find out that whether intelligent agents can be controllable and future smart bots can be safe (or trusted) or not. Resolving this challenge is not an easy task because of the nature of the predictability challenge. In what follows, we highlight some parts of it. It should be noted that unpredictability can be seen in the behavior of an agent with reinforcement learning algorithms because of the nature of these types. In [80], some issues that affect the predictability of AI-based agents are summarized. It seems that chaotic behavior in mathematical and physical systems plays a critical role in this issue. In other words, there are some analogies between chaos in mathematics, physics, and AI, and hence AI decisions cannot be predictable in some cases. Other issues, including ambiguity and paradox, may lead to an unpredictable AI-based system. In an HLI-based agent, unpredictability may lead to many subproblems such as safety, trust, accountability, and fairness, to mention a few.

3.13. Continual Learning

During the lifecycle of AI-based systems, the accuracy of the learning model goes down because of changes in the data and environment of the model. Therefore, the learning process should be changed using new methods to support continual and lifelong learning. On the other hand, in real-world applications, a massive amount of data is produced via the Internet of Things (IoT) and embedded sensors in a real-time manner. Hence, new systems' algorithms should be considered a stream of data and continuous learning. In [81], a comprehensive survey on continual learning is given. According to this reference, continual learning methods can be classified into three classes, as explained below:

- Reply methods: These algorithms store samples in raw format or generate pseudo-samples using a generative model.
- Regularization-based methods: These algorithms provide an additional regularization term in the loss function while learning new data.
- Parameter isolation methods: This type of algorithm provides different model parameters for each task to prevent any possible forgetting.

In all types of AI-based agents, such as HLI-based agents, continual learning methods enable agents to adapt to a continually changing and uncertain environment [82]. No study in the literature solves problems related to continual learning considering HLI.

3.14. Storage (Memory)

Memory is an important part of all AI-based systems. A limited memory AI-based system is one of the most widely and commonly used types of intelligent systems [83]. In this type, historical observations are used to predict some parameters about the trend of changes in data. In this approach, some data-driven and also statistical analyses are used to extract knowledge from data. This approach is not new to the AI field and fed by data, storage capabilities, computational power, and learning capabilities, as explained in the following. In many situations, learning abilities can be improved via more data. Since the size of data collected by AI-based systems increases, efficient algorithms are essential for data analysis and decision-making. With the drastic increase in data, technologies behind storage and computation might be revolutionized in the near future. Information may be stored in short-term or long-term memory units, leading to several problems in different domains, including reading, computing, and writing. In [84], a method based on cognitive computation has been proposed to tackle this problem. It is worth noting that executing cognitive engines considering big data in the memory through online learning algorithms may cause challenging tasks [30]. For more information about data management in machine learning, please refer to [85].

During designing HLI-based agents, storage (memory) issues arise other sub-challenges such as real-time decision-making capabilities, emulating functions of different forms of memories (short term and long term) similar to humans, and supporting human thinking styles for computation, such as those identified in cognitive architectures [86].

3.15. Semantic and Communication

Semantic refers to a well-known issue in AI-based systems. From semantic web techniques to linguistic analysis and natural language processing may be related to semantic computations in AI-based systems [87–89]. On the other hand, communication among intelligent agents leads to flowing information in a population of agents resulting in increasing knowledge and intelligence in that population. For more than one century, humans have tried to communicate with other intelligent entities in the universe. Therefore, semantics and communication may refer to two different concepts, although there is an intersection between them which we focus on in the next paragraph. We know that defining or determining a shared ontology among intelligent entities in an AI-based system is possible because of maturing some parts of knowledge in ontology manipulations and defining some tools in semantic web techniques [88]. In other words, determining some standards for semantics and communication among intelligent agents is possible, leading to

organization of an intelligent population. In the next paragraph, we focus on the importance of the mentioned issue in HLI-based agents.

HLI-based agents are utilized in a wide range of areas such as computer networks, Internet of Things (IoT), and robotics. Communication is the cornerstone of these areas, which is performed using a standard language [90], and also ontology [91]. Therefore, learning algorithms embedded in all devices can use a massive amount of data to evolve their cognitive abilities. Peer-to-peer file sharing and blockchain can be considered as potential solutions to provide infrastructure for communication and knowledge (ontology) sharing [92].

3.16. Morality and Ethical

Informally, both ethics and morals terms relate to "right" and "wrong". Sometimes, these terms are used interchangeably, but they may refer to different things. Ethics refers to rules provided by external entities, such as rules of workplaces or principles in religions. Morals refer to personal principles regarding right and wrong. Because of the wide intersection between the mentioned concepts, we talk about the shared attributes in the rest of this section, focusing on AI and HLI.

Morality is a complicated concept of AI due to its various dimensions and different actors. Ethics are considered as the set of moral principles that guide a person's behavior. From a perspective of morality issue, it is preserving the privacy of data within learning processes [93]. In this perspective, the engineers and social interactions of humans are the subjects of morality. From another perspective, implementing the concepts related to morality in a cognitive engine can be seen as a goal of AI designers. This is because we expect to see morality in an agent designated based on AGI and also HLI. Since governments try to determine some ethics of AI-based systems, relevant concepts such as accountability, explainability, and trustworthiness have appeared in this area.

3.17. Rationality

An important goal of AI research is understanding intelligence. The concept of rational agency has long been considered as a critical role in defining intelligent agents. Rationality computation plays a key role in distributed machine learning, multi-agent systems, game theory, and also AGI [94–96]. In [97], the definition of rationality is categorized into four categories, as explained below:

- **Perfect rationality:** An agent with this type of rationality can generate maximal successful behavior based on its available information.
- **Calculative rationality:** An agent with this type of rationality can compute a perfectly rational decision given the initially available information.
- **Metalevel rationality:** An agent with this type of rationality can select the optimal combination of computation-sequence-plus-action. During this process, the constraint is that the action must be selected under which is computed.
- **Bounded rationality:** An agent with this type of rationality can behave successfully based on available information and computational resources.

Because of the direct relationship between the natural definition of rationality and HLI-based intelligence, the mentioned definitions all belong to the basic principles of HLI. Each definition leads to the design of a specific type of intelligence. Unfortunately, a lack of required information prevents the creation of an agent with perfect rationality. Other types required heavy design and computation.

In cognitive science, which plays an essential role in HLI, rationality means how a human agent behaves based on a particular normative theory. Considering the abilities of humans, human deficiencies, and unknown parts of humans, intelligent systems modeling considering rationality leads to challenging problems [94]. A surprising fact in this area is the rare research in this field. It seems that considerable effort should be made into this field.

3.18. Mind

Theory of mind is one of the active research fields in artificial intelligence, robotics, and cognitive science. According to [98], a wide range of scientists, including psychiatrists, psychologists, and neuroscientists, are involved in constructing some algorithms and machines that can implement mind computations and also mental states. In [98], the theory of mind and machine were studied, and an interesting proposal based on deep learning and reinforcement learning is also suggested. This is because the behavior of a person and interaction of a person with others in a society can be learnt by the mentioned technologies. In [99], another example based on reinforcement learning and neural networks and Markov theory are suggested in this regard. All in all, after a long time of effort in this field, many unknown questions are left, and supervising few scientists works in this era. According to [98–100], the theory for mind computation will substantially affect HLI-based agents.

3.19. Accountability

An essential feature of decision-making in humans, AI, and also HLI-based agents is accountability. Implementing this feature in machines is a difficult task because many challenges should be considered to organize an AI-based model that is accountable. It should be noted that this issue in human decision-making is not ideal, and many factors such as bias, diversity, fairness, paradox, and ambiguity may affect it. In addition, the human decision-making process is based on personal flexibility, context-sensitive paradigms, empathy, and complex moral judgments. Therefore, all of these challenges are inherent to designing algorithms for AI and also HLI models that consider accountability. In the next paragraph, some related works are studied.

In [101], it was highlighted that accountability has a very close relationship with regulating AI-based systems. Without this feature, thinking about the law and regulation may not lead to concrete rules. In [102], the problem of bias and its origin in human decision-making is studied. This work also discusses the diversity issue, which may be used to define some algorithms to mitigate the negative effects of bias in decision-making. Since the decision-making of humans suffers from bias, this problem may affect understanding of the accountability capabilities of humans, which may appear in AI-based systems. In [103], a recent case in the USA known as "State v Loomis" was used to show how the ultimate and unrestrained dedication of public power to machines may ruin human rights and the rule of law. This paper focuses on further problems that may appear in this regard, and studies the balance between the benefits and costs of algorithmizing.

All in all, accountability has a close relationship with explainability, trust, security, responsibility, and fairness. In addition, in [104], the relationship between accountability and safety is reported.

3.20. Transparency

According to [105], transparency may involve input data and the underlying computation of the learning model. From a computational perspective, this issue implicitly arises with accountability and explainability in the decision-making process of the learning model. From a data consumption perspective, an external entity of an AI-based ecosystem may want to know which parts of data affect the final decision in a learning model. In [106], a novel form of transparency, such as translational transparency, and also different perspectives on this issue from different roles, such as engineers, users, and legal experts, are explained. In what follows, some highlights of this issue are explained.

- Similar to accountability, humans like to see transparent behavior from AI-based models. By increasing the use of AI for making decisions in public affairs, this challenge becomes more complicated. In addition, with the creation of self-organized learning systems that may be hurtful to people, this issue grabs more attention. In some applications, such as military and healthcare systems, a learning system as a black box may not be appropriate, and some features such as explainability and

accountability may not be able to bring transparency abilities. This feature will be vital in HLI-based agents.

- Recently, we are faced with many papers that report the results of accuracy of data-driven machine learning models. It is obvious that the reported accuracies are obtained from best practices; thus, with increasing parameter space, finding the exact accuracy that is equal to the reported results may not be possible. This issue may lead to a confusing process in reproducing models when the model generator codes are not reported [107].

A summary of related works is given as follows. In [106], a study is given that identifies what transparency means for technical, legislative, and public realities and stakeholders. The study reported in [108] provides lawyers with a practical approach to the functions of potential AI transparency legislation in addition to a set of legal instruments which can be employed to this end. In [109], the issues of transparency from socio-legal and computer science aspects are studied. In [110], transparency is addressed by integrating legal, social, and ethical aspects considering contextual and performative factors which affect transparency computation. Such studies which consider general data protection regulation and its ethical groundwork will be essential for policymakers.

3.21. Reproduceability

This issue can be interpreted as two concepts, as explained below:

- How a learning model can be reproduced when it is obtained based on various sets of data and a large space of parameters. This problem becomes more challenging in data-driven learning procedures without transparent instructions [107]. Many papers on the scope of applied machine learning are not well documented according to the report given in [111]. This problem becomes more serious when privacy-preserving concerns about data should be considered.
- How a learning model can reproduce itself when self-reproducibility is considered as a final destination of AI-based models [112].

3.22. Evolution

To define AI-based models, Darwinian evolution theory including "survival of the fittest" was reported in the literature. In this approach, AI models can be improved during the evolution of generations without human aid. This approach is not perfect, and many problems must be solved to provide an AI-based model that can improve itself.

- From a programming perspective, existing structures for genes and chromosomes, and the evolution process are not naturally similar to evolutionary processes that happen in nature. It seems that computer worms and viruses that utilize some programming such as quine codes and polymorphic structure are more pioneers than existing evolutionary programs that may be found in repositories such as GitHub. In [113], a self-replicating neural network is presented that may be used in an evolutionary strategy.
- From the theory of evolutionary computation, the concept behind genetic and memetic computation is changing over time to better model the phenomena that happen in nature. These algorithms have many variations such as variable chromosomes, diverse crossover, and complex mutations. It should be noted that many of the existing evolutionary algorithms are based on a very simplified search process that happens in nature [114].

It seems that existing evolutionary computation may not reflect the power of evolutionary strategies that can evolve simple AI models to an ASI-based agent. In HLI, this challenge should be solved because AI models must be able to adapt to their environment through evolution strategies.

3.23. Beneficial

A beneficial AI system is designated to behave in such a way that humans are satisfied with the results. Designing these systems will be required, but extending their theory is still an ongoing process. In these systems, the agent is initially uncertain about the preferences of humans, and human behavior will be used to extract information about human preferences [115]. This problem becomes a challenging issue in the development of HLI-based agent because information about the reactions of humans in many situations is required to organize a beneficial learning model for HLI.

3.24. Exploration and Exploitation Balance

The problem of balancing exploration and exploitation refers to a challenging problem that arises in different domains of machine learning, such as evolutionary computations, active learning, and reinforcement learning to mention a few. Exploration and exploitation decisions refer to trading off the benefits of exploring unknown opportunities to learn more about them, by exploiting known opportunities. Such decisions are usually in practice, but from a computational perspective, they are hard. In [116], a new active learning algorithm that balances exploration with the refinement of the decision boundary by dynamically adjusting the probability is reported. In [117], several methods reported in the literature are summarized. In [118], an auto-tuning strategy by using fuzzy logic control was reported. In this method, for taking the balance among the stochastic search and local search probabilities, an adaptively regulation algorithm based on the change of the average fitness of parents and offspring at each generation is designated. In [119], an information-theoretic approach for solving the exploration-exploitation dilemma in reinforcement learning was reported. In this method, a criterion that highlights the optimal trade-off between the expected returns and a policy's degrees of freedom was deployed, which considers the value of information. In different domains and algorithms, keeping the balance between exploration and exploitation lead to very challenging and hard problems. This problem in HLI-based agents becomes harder because of the self-organized capabilities of HLI in general domains.

3.25. Verifiability

Verification is one method used by software developers to gain trust. In [120], this problem was addressed in AI. In many applications of AI-based systems such as medical healthcare and military services, the lack of verification of code may not be tolerable. Therefore, verifying and validating AI-based models have received considerable attention recently [121,122]. It should be noted that, due to some characteristics such as the non-linear and complex structure of AI-based solutions, existing solutions have been generally considered "black boxes", not providing any information about what exactly makes them appear in their predictions and decision-making processes. In this regard, the development of algorithms for visualizing, explaining, and interpreting learning models has recently attracted increasing attention. For example, the aim of [123] was to evaluate the use of learning algorithms for the detection of breath sounds in a real clinical environment among children with pulmonary diseases. Because of rare attention regarding verification challenges, it seems that many things, including theories and applications, should be presented in this field in the near future. Since the HLI-based agents can exhibit self-organized capability, verifying them will not be easy task.

3.26. Safety

The actions of a learning model may easily hurt humans in both explicit and implicit manners. This challenge is not new to the AI community, but its dimensions have evolved along with the drastic usage of intelligent systems. In [124–126], several algorithms based on Asimov's laws have been proposed that try to judge the output actions of an agent considering the safety of humans. Asimov's laws belong to reactive solutions, but some

may prefer to organize proactive solutions, such as [9]. In this solution, a computational model that considers the cost and weight of harmful actions is proposed.

As a pioneer in the scope of AI and HLI, we can refer to safety engineering principles given in [127]. This challenge becomes more critical in mission-critical and real-world environments where an AI-based agent should be used, because there is no tolerance for failure that can damage humans and devices.

3.27. Complexity

The primary algorithms of AI-based systems were designated to solve specific and simple problems. Therefore, their complexity was low. This assumption is not correct any longer. Nowadays, we are faced with systems that utilize numerous learning models in their modules for their perception and decision-making processes. This means that the complexity of intelligent systems is increasing day by day. One aspect of an AI-based system that leads to increasing the complexity of the system is the parameter space that may result from multiplications of parameters of the internal parts of the system. To address the mentioned issue, some efforts such as AutoML [128] and parameter optimization procedures [129] are reported in the literature. Complexity issues in HLI-based agents may be harder issues compared to other types of intelligent systems because of the many unknown parts of human intelligence that should be considered during mitigation of the complexity of the system.

3.28. Trustworthiness

According to [130], trustworthiness in AI will feed societies, economies, and sustainable development to bring the ultimate benefits of AI to individuals, organizations, and societies. In [130], five foundational principles; beneficence, non-maleficence, autonomy, justice, and explicability are explained in this regard. This field is highly interdisciplinary and also dynamic, and covers many issues including psychology, sociology, economics, management, and computer science. From a social perspective, trustworthiness has a close relationship with ethics and morality as explained in [12]. In [131], fairness, explainability, accountability, and reliability are reported as related concepts to trustworthiness. Designing an HLI-based agent that considers trustworthiness is the ultimate goal of AI research.

4. Discussion

The following questions will be answered in the rest of this section to analyze all aspects of the challenges.

- Q1: What are fundamental problems, theoretical and machines limitations during the development of AI? Section 4.1 is dedicated to answering this question.
- Q2: What are the potential consequences of the evolution of challenges during the development of AI? Section 4.2 is dedicated to answering this question.
- Q3: What are combinations of challenges? Section 4.3 is dedicated to answering this question.
- Q4: What is the importance of the context of challenges during the development of AI? Section 4.4 is dedicated to answering this question.
- Q5: What is the role of analyzing only human intelligence during the development of AI? Section 4.5 is dedicated to answering this question.

4.1. Fundamental Problems, Theoretical and Machines Limitations

As mentioned earlier, some problems related to the halting problem [73] and controllability [76] might not be solved in the future. Therefore, many AI-based systems that inherit these problems might not be manageable. On the other hand, some phenomena, such as strange loops explained in [79], cannot be managed by machines with existing frameworks for mathematics. In addition, most parts of mathematics required by learning algorithms are still in progress [132]. It seems that mathematics behind cognitive sciences should be mature enough to be used in the theory of mind [133–135]. All the mentioned

problems become more challenging when we discover new things about humans and our surrounding environment. It is obvious that our knowledge about humans is not complete. Therefore, some challenges such as controllability, predictability, security, and safety, which might have a close relationship with the mentioned problems, will become more severe during the development of AI, especially for AGI, HLI, and also ASI.

4.2. The Evolution and Transformation of Challenges

A supervising fact about development of AI-based systems is the evolution of challenges over time. This issue becomes more critical when challenges and their associated problems are detected after displaying harmful consequences at the industrial level. The evolutions of some challenges are explained below:

- **Evolution of data challenges:** With the creation of data-driven machine learning algorithms, we had rare data. Now, because of the use of cloud computing, we have big data. In the near future, because of the use of AI-based generative data, such as generative adversarial networks (GANs), we would need some technology to differentiate real-world and appropriate data from other types of data in a vast amount of data [136,137].
- **Evolution of robustness and security challenges:** If the noise of input data is synthetic and designated for changing the model's behavior abnormally, we will face another challenge, known as security against adversarial attacks, studied in security challenges. Since the number of adversarial attacks is increasing, robustness will be connected to security in most cases soon [31].
- **Evolution of energy consumption challenge:** It was previously mentioned that energy consumption is a substantial challenge in the lifecycle of AI-based systems. This challenge may lead to other challenges, such as carbon emissions, environmental pollution, and global warming [138].
- **Evolution of complexity challenges:** A system with high complexity might lead to non-verifiable, unreliable, unaccountable, and nonreproducible AI-based systems. For example, a model with a large set of parameters and also a complex process for tunning them may lead to reproducibility problems, as explained in [107].

All of the challenges mentioned in Section 3 might evolve in the future because of evolving human needs and changes that may occur in human environment.

4.3. Combinations of Challenges

Some of the challenges, such as security and trustworthiness, are the result of combinations among others. We briefly describe some of them as below.

- According to [8], the security challenge implicitly arises from other challenges such as trust, confidentiality, and privacy.
- According to [109], the transparency challenge implicitly leads to other challenges such as accountability and explainability in the decision-making process of the learning model.
- According to [131], the trustworthiness challenge implicitly leads to other challenges such as fairness, explainability, accountability, and reliability.
- According to [122], the verification challenge implicitly results in other challenges such as safety and robustness.
- According to [46,139], the safety challenge has a solid relationship with security and ethical challenges.

In combined challenges, we may face paradox and contradiction. For example, improving the security of a system might not be possible without using users' private information [140].

4.4. Importance of Context of Challenge: Deep Dive into Dark Sides

Many AI researchers only focus on solving a specific problem with a fixed set of assumptions as context, and usually in a stationary environment. These assumptions may lead to terrible situations in practice because intelligent systems will be used in many products and industries with variable situations. For example, the primary concepts about the fairness challenge were considered for static data. Nowadays, this assumption is totally changed, and we need to adapt the learning model to some dynamic environments. Hence, a fixed algorithm for fair decision-making may not be applicable [51]. In [51], an approach for fairness-aware stream classification has been supported, which can manage predictive performance with low discrimination scores for the stream. The mentioned problem for fairness can appear in all challenges studied in Section 3. From another perspective, according to the scope of intelligence, such as ASI, AGI, HLI, and ANI, the interpretation of challenges might be changed.

4.5. Human Intelligence Is Partially Known to Design HLI

The definitions of intelligence utilized in this paper and also a wide range of papers in the literature focus on humans as partially known intelligent entities [20,141]. Mind and intelligence are not homogenous among humans [142]. Sometimes, an action in one society is good (or rational) and in another is terrible [97]. With such diversity, how we can define beneficial computations for the machines is an important question [115]. Utilizing all concepts arising in human intelligence in developing a machine with HLI may not lead to an appropriate situation. In this regard, some of the problems are explained below:

- Some psychological concepts such as classification over self-awareness might be applicable to machines, such as the one reported in [143], but there is no literature on the evolution of this concept in machines. In other words, the context that leads to self-awareness in humans is based on limited information and limited sensors, but these assumptions might not be correct for machines. This means that there is little information about the type of self-awareness in machines that meet HLI, AGI, or ASI criteria. To the best of our knowledge, the connections among mind, cognitive science, learning capabilities, and self-awareness in humans and also animals are not totally understood in the literature [144–146].
- Pain and death are two key factors in analyzing cognitive concepts such as the theory of meaning management [147], psychology of souls [148], and religious understanding [149]. Since, machines might not face with death and pain, how can we expect to use computation behind the human mind to control machines?

5. Conclusions

In this paper, the challenges of AI were analyzed with a particular focus on HLI. It should be noted that the popularities of challenges were not similar to each other. In other words, some of the challenges, such as security and fairness, are more popular than other challenges, such as energy and complexity. Therefore, a brief description was given for each challenge to understand the nature of the challenge. In addition, the connections and combinations among challenges, and also their evolutions, were studied. We left some well-known challenges such as the curse of dimensionality because of its popularity in the AI domain. It is obvious that during the evolution from ANI to ASI, all of the challenges may get new dimensions as we discover more information about our environment during interaction with intelligent systems. We tried to address this issue in the discussion section. Some challenges such as the monopoly of corporations during the AI era, as well as widespread job loss, were not in the scope of this paper because the focus of this paper is mainly on computer science. These challenges can be considered in future works.

References

1. Binu, D.; Rajakumar, B.R. *Artificial Intelligence in Data Mining: Theories and Applications*; Academic Press: Cambridge, MA, USA, 2021.
2. Ahmadi, A.; Meybodi, M.R.; Saghiri, A.M. Adaptive search in unstructured peer-to-peer networks based on ant colony and Learning Automata. In Proceedings of the 2016 Artificial Intelligence and Robotics, Qazvin, Iran, 9 April 2016.
3. Cheng, X.; Lin, X.; Shen, X.-L.; Zarifis, A.; Mou, J. The dark sides of AI. *Electron. Mark.* **2022**, 1–5. [CrossRef]
4. Jabbarpour, M.R.; Saghiri, A.M.; Sookhak, M. A framework for component selection considering dark sides of artificial intelligence: A case study on autonomous vehicle. *Electronics* **2021**, *10*, 384. [CrossRef]
5. Kumar, G.; Singh, G.; Bhatanagar, V.; Jyoti, K. Scary dark side of artificial intelligence: A perilous contrivance to mankind. *Humanit. Soc. Sci. Rev.* **2019**, *7*, 1097–1103. [CrossRef]
6. Mahmoud, A.B.; Tehseen, S.; Fuxman, L. The dark side of artificial intelligence in retail innovation. In *Retail Futures*; Emerald Publishing Limited: Bingley, UK, 2020.
7. Wirtz, B.W.; Weyerer, J.C.; Sturm, B.J. The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *Int. J. Public Adm.* **2020**, *43*, 818–829. [CrossRef]
8. Hanif, M.A.; Khalid, F.; Putra, R.V.W.; Rehman, S.; Shafique, M. Robust machine learning systems: Reliability and security for deep neural networks. In Proceedings of the 2018 IEEE 24th International Symposium on On-Line Testing and Robust System Design (IOLTS), Platja d'Aro, Spain, 2–4 July 2018; pp. 257–260.
9. Varshney, K.R. Engineering safety in machine learning. In Proceedings of the 2016 Information Theory and Applications Workshop (ITA), La Jolla, CA, USA, 31 January–5 February 2016; pp. 1–5.
10. Bellamy, R.K.; Dey, K.; Hind, M.; Hoffman, S.C.; Houde, S.; Kannan, K.; Lohia, P.; Martino, J.; Mehta, S.; Mojsilović, A. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM J. Res. Dev.* **2019**, *63*, 4:1–4:15. [CrossRef]
11. Strubell, E.; Ganesh, A.; McCallum, A. Energy and policy considerations for deep learning in NLP. *arXiv* **2019**, arXiv:1906.02243.
12. Smuha, N.A. The EU approach to ethics guidelines for trustworthy artificial intelligence. *Comput. Law Rev. Int.* **2019**, *20*, 97–106. [CrossRef]
13. Legg, S.; Hutter, M. A collection of definitions of intelligence. *Front. Artif. Intell. Appl.* **2007**, *157*, 17.
14. Legg, S. Machine Super Intelligence. Ph.D. Thesis, University of Lugano, Lugano, Switzerland, 2008.
15. Saghiri, A.M. A Survey on Challenges in Designing Cognitive Engines. In Proceedings of the 2020 6th International Conference on Web Research (ICWR), Tehran, Iran, 22–23 April 2020; pp. 165–171.
16. Boström, N. *Superintelligence: Paths, Dangers, Strategies*; Oxford University Press: Oxford, UK, 2014.
17. Chollet, F. On the measure of intelligence. *arXiv* **2019**, arXiv:1911.01547.
18. Yampolskiy, R.V. Human is not equal to AGI. *arXiv* **2020**, arXiv:2007.07710.
19. Searle, J.R. Minds, brains, and programs. *Behav. Brain Sci.* **1980**, *3*, 417–424. [CrossRef]
20. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall: Hoboken, NJ, USA, 1994.
21. Linz, P. *An Introduction to Formal Languages and Automata*; Jones & Bartlett Learning: Burlington, MA, USA, 2006.
22. Lenat, D.B.; Guha, R.V.; Pittman, K.; Pratt, D.; Shepherd, M. Cyc: Toward programs with common sense. *Commun. ACM* **1990**, *33*, 30–49. [CrossRef]
23. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; Cambridge University Press: Cambridge, UK, 1998.
24. Steane, A. Quantum computing. *Rep. Prog. Phys.* **1998**, *61*, 117. [CrossRef]
25. Wheeldon, A.; Shafik, R.; Rahman, T.; Lei, J.; Yakovlev, A.; Granmo, O.-C. Learning automata based energy-efficient AI hardware design for IoT applications. *Philos. Trans. R. Soc. A* **2020**, *378*, 20190593. [CrossRef] [PubMed]
26. Priya, S.; Inman, D.J. *Energy Harvesting Technologies*; Springer: Berlin/Heidelberg, Germany, 2009.
27. Kamalinejad, P.; Mahapatra, C.; Sheng, Z.; Mirabbasi, S.; Leung, V.C.; Guan, Y.L. Wireless energy harvesting for the Internet of Things. *IEEE Commun. Mag.* **2015**, *53*, 102–108. [CrossRef]
28. Baig, M.I.; Shuib, L.; Yadegaridehkordi, E. Big Data Tools: Advantages and Disadvantages. *J. Soft Comput. Decis. Support Syst.* **2019**, *6*, 14–20.
29. Sivarajah, U.; Kamal, M.M.; Irani, Z.; Weerakkody, V. Critical analysis of Big Data challenges and analytical methods. *J. Bus. Res.* **2017**, *70*, 263–286. [CrossRef]

30. Qiu, J.; Wu, Q.; Ding, G.; Xu, Y.; Feng, S. A survey of machine learning for big data processing. *EURASIP J. Adv. Signal Process.* **2016**, *2016*, 67. [CrossRef]

31. Qayyum, A.; Qadir, J.; Bilal, M.; Al-Fuqaha, A. Secure and robust machine learning for healthcare: A survey. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 156–180. [CrossRef]

32. Bhagoji, A.N.; Cullina, D.; Sitawarin, C.; Mittal, P. Enhancing robustness of machine learning systems via data transformations. In Proceedings of the 2018 52nd Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 21–23 March 2018; pp. 1–5.

33. Rozsa, A.; Günther, M.; Boult, T.E. Are accuracy and robustness correlated. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 227–232.

34. Pérez-Rosas, V.; Abouelenien, M.; Mihalcea, R.; Burzo, M. Deception detection using real-life trial data. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, Seattle, WA, USA, 9–13 November 2015; pp. 59–66.

35. Krishnamurthy, G.; Majumder, N.; Poria, S.; Cambria, E. A deep learning approach for multimodal deception detection. *arXiv* **2018**, arXiv:1803.00344.

36. Randhavane, T.; Bhattacharya, U.; Kapsaskis, K.; Gray, K.; Bera, A.; Manocha, D. The Liar's Walk: Detecting Deception with Gait and Gesture. *arXiv* **2019**, arXiv:1912.06874.

37. Zhao, S.; Jiang, G.; Huang, T.; Yang, X. The deception detection and restraint in multi-agent system. In Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05), Hong Kong, China, 14–16 November 2005; pp. 44–48.

38. Zlotkin, G.; Rosenschein, J.S. Incomplete Information and Deception in Multi-Agent Negotiation. In Proceedings of the IJCAI, Sydney, Australia, 24–30 August 1991; Volume 91, pp. 225–231.

39. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.

40. Blitz, M.J. Lies, Line Drawing, and Deep Fake News. *Okla. Law Rev.* **2018**, *71*, 59.

41. Tsai, C.-F.; Hsu, Y.-F.; Lin, C.-Y.; Lin, W.-Y. Intrusion detection by machine learning: A review. *Expert Syst. Appl.* **2009**, *36*, 11994–12000. [CrossRef]

42. Pawar, S.N.; Bichkar, R.S. Genetic algorithm with variable length chromosomes for network intrusion detection. *Int. J. Autom. Comput.* **2015**, *12*, 337–342. [CrossRef]

43. Kinsner, W. Towards cognitive security systems. In Proceedings of the 11th International Conference on Cognitive Informatics and Cognitive Computing, Kyoto, Japan, 22–24 August 2012; p. 539.

44. Biggio, B.; Fumera, G.; Roli, F. Security evaluation of pattern classifiers under attack. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 984–996. [CrossRef]

45. Barreno, M.; Nelson, B.; Sears, R.; Joseph, A.D.; Tygar, J.D. Can machine learning be secure? In Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security, Taipei, Taiwan, 21–24 March 2006; pp. 16–25.

46. Yampolskiy, R.V. *Artificial Intelligence Safety and Security*; CRC Press: Boca Raton, FL, USA, 2018.

47. Huang, L.; Joseph, A.D.; Nelson, B.; Rubinstein, B.I.; Tygar, J. Adversarial machine learning. In Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence, Chicago, IL, USA, 21 October 2011; pp. 43–58.

48. Ateniese, G.; Felici, G.; Mancini, L.V.; Spognardi, A.; Villani, A.; Vitali, D. Hacking smart machines with smarter ones: How to extract meaningful data from machine learning classifiers. *arXiv* **2013**, arXiv:1306.4447. [CrossRef]

49. Tucker, C.; Agrawal, A.; Gans, J.; Goldfarb, A. Privacy, algorithms, and artificial intelligence. In *The Economics of Artificial Intelligence: An Agenda*; Oxford University Press: Oxford, UK, 2018; pp. 423–437.

50. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]

51. Zhang, W.; Ntoutsi, E. Faht: An adaptive fairness-aware decision tree classifier. *arXiv* **2019**, arXiv:1907.07237.

52. Kamani, M.M.; Haddadpour, F.; Forsati, R.; Mahdavi, M. Efficient fair principal component analysis. In *Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–32.

53. Dwork, C.; Hardt, M.; Pitassi, T.; Reingold, O.; Zemel, R. Fairness through awareness. In Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, Cambridge, MA, USA, 8–10 January 2012; pp. 214–226.

54. Kamiran, F.; Calders, T. Classifying without discriminating. In Proceedings of the 2009 2nd International Conference on Computer, Control and Communication, Karachi, Pakistan, 17–18 February 2009; pp. 1–6.

55. Calders, T.; Kamiran, F.; Pechenizkiy, M. Building classifiers with independency constraints. In Proceedings of the 2009 IEEE International Conference on Data Mining Workshops, Miami, FL, USA, 6 December 2009; pp. 13–18.

56. Quy, T.L.; Roy, A.; Iosifidis, V.; Ntoutsi, E. A survey on datasets for fairness-aware machine learning. *arXiv* **2021**, arXiv:2110.00530.

57. Hardt, M.; Price, E.; Srebro, N. Equality of opportunity in supervised learning. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1–9.

58. Kamishima, T.; Akaho, S.; Sakuma, J. Fairness-aware learning through regularization approach. In Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops, Vancouver, BC, Canada, 11 December 2011; pp. 643–650.

59. Goh, G.; Cotter, A.; Gupta, M.; Friedlander, M.P. Satisfying real-world goals with dataset constraints. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1–9.

60. Calders, T.; Verwer, S. Three naive Bayes approaches for discrimination-free classification. *Data Min. Knowl. Discov.* **2010**, *21*, 277–292. [CrossRef]
61. Donini, M.; Oneto, L.; Ben-David, S.; Shawe-Taylor, J.S.; Pontil, M. Empirical risk minimization under fairness constraints. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 1–11.
62. Morgenstern, J.; Samadi, S.; Singh, M.; Tantipongpipat, U.; Vempala, S. Fair dimensionality reduction and iterative rounding for sdps. *arXiv* **2019**, arXiv:1902.11281.
63. Samadi, S.; Tantipongpipat, U.; Morgenstern, J.H.; Singh, M.; Vempala, S. The price of fair pca: One extra dimension. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 1–12.
64. Pleiss, G.; Raghavan, M.; Wu, F.; Kleinberg, J.; Weinberger, K.Q. On fairness and calibration. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–10.
65. Adadi, A.; Berrada, M. Explainable AI for healthcare: From black box to interpretable models. In *Embedded Systems and Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 327–337.
66. Gade, K.; Geyik, S.C.; Kenthapadi, K.; Mithal, V.; Taly, A. Explainable AI in industry. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 3203–3204.
67. Došilović, F.K.; Brčić, M.; Hlupić, N. Explainable artificial intelligence: A survey. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018; pp. 0210–0215.
68. Samek, W.; Müller, K.-R. Towards explainable artificial intelligence. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 5–22.
69. Sharma, S.; Nag, A.; Cordeiro, L.; Ayoub, O.; Tornatore, M.; Nekovee, M. Towards explainable artificial intelligence for network function virtualization. In Proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies, Barcelona, Spain, 1–4 December 2020; pp. 558–559.
70. Matthias, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics Inf. Technol.* **2004**, *6*, 175–183. [CrossRef]
71. Neri, E.; Coppola, F.; Miele, V.; Bibbolino, C.; Grassi, R. *Artificial Intelligence: Who Is Responsible for the Diagnosis?* Springer: Berlin/Heidelberg, Germany, 2020.
72. Stannett, M. X-machines and the halting problem: Building a super-Turing machine. *Form. Asp. Comput.* **1990**, *2*, 331–341. [CrossRef]
73. Rybalov, A. On the strongly generic undecidability of the Halting Problem. *Theor. Comput. Sci.* **2007**, *377*, 268–270. [CrossRef]
74. Yampolskiy, R.V. On Controllability of AI. *arXiv* **2020**, arXiv:2008.04071.
75. Russell, S. *Human Compatible: Artificial Intelligence and the Problem of Control*; Penguin: London, UK, 2019.
76. Yampolskiy, R. *On Controllability of Artificial Intelligence*; Technical Report; University of Louisville: Louisville, KY, USA, 2020.
77. Dawson, J. *Logical Dilemmas: The Life and Work of Kurt Gödel*; AK Peters: Natick, MA, USA; CRC Press: Boca Raton, FL, USA, 1996.
78. Yampolskiy, R.V. Unpredictability of AI. *arXiv* **2019**, arXiv:1905.13053.
79. Hofstadter, D.R. *I Am a Strange Loop*; Basic Books: New York, NY, USA, 2007.
80. Musiolik, G. Predictability of AI Decisions. In *Analyzing Future Applications of AI, Sensors, and Robotics in Society*; IGI Global: Hershey, PA, USA, 2021; pp. 17–28.
81. Delange, M.; Aljundi, R.; Masana, M.; Parisot, S.; Jia, X.; Leonardis, A.; Slabaugh, G.; Tuytelaars, T. A continual learning survey: Defying forgetting in classification tasks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**. [CrossRef] [PubMed]
82. Shin, H.; Lee, J.K.; Kim, J.; Kim, J. Continual learning with deep generative replay. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 2990–2999.
83. Hassani, H.; Silva, E.S.; Unger, S.; TajMazinani, M.; Mac Feely, S. Artificial intelligence (AI) or intelligence augmentation (IA): What is the future? *AI* **2020**, *1*, 143–155. [CrossRef]
84. Widrow, B.; Aragon, J.C. Cognitive Memory. *Neural Netw.* **2013**, *41*, 3–14. [CrossRef]
85. Kumar, A.; Boehm, M.; Yang, J. Data management in machine learning: Challenges, techniques, and systems. In Proceedings of the 2017 ACM International Conference on Management of Data, Chicago, IL, USA, 14–19 May 2017; pp. 1717–1722.
86. Kotseruba, I.; Tsotsos, J.K. 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artif. Intell. Rev.* **2020**, *53*, 17–94. [CrossRef]
87. Berners-Lee, T.; Hendler, J.; Lassila, O. The semantic web. *Sci. Am.* **2001**, *284*, 28–37. [CrossRef]
88. Feigenbaum, L.; Herman, I.; Hongsermeier, T.; Neumann, E.; Stephens, S. The semantic web in action. *Sci. Am.* **2007**, *297*, 90–97. [CrossRef]
89. Cambria, E.; White, B. Jumping NLP curves: A review of natural language processing research. *IEEE Comput. Intell. Mag.* **2014**, *9*, 48–57. [CrossRef]
90. Chaib-Draa, B.; Dignum, F. Trends in agent communication language. *Comput. Intell.* **2002**, *18*, 89–101. [CrossRef]
91. Maedche, A.; Staab, S. Ontology learning for the semantic web. *IEEE Intell. Syst.* **2001**, *16*, 72–79. [CrossRef]
92. Teslya, N.; Smirnov, A. Blockchain-based framework for ontology-oriented robots' coalition formation in cyberphysical systems. In Proceedings of the MATEC Web of Conferences, Anyer, Indonesia, 4–5 September 2018; Volume 161, pp. 03–18.
93. Luccioni, A.; Bengio, Y. On the Morality of Artificial Intelligence. *IEEE Technol. Soc. Mag.* **2020**, *39*, 16–25. [CrossRef]

94. Abdel-Fattah, A.M.; Besold, T.R.; Gust, H.; Krumnack, U.; Schmidt, M.; Kuhnberger, K.-U.; Wang, P. Rationality-guided AGI as cognitive systems. In Proceedings of the Annual Meeting of the Cognitive Science Society, Sapporo, Japan, 1–4 August 2012; Volume 34.

95. Gigerenzer, G.; Selten, R. *Rethinking rationality. Bounded Rationality: The Adaptive Toolbox*; MIT Press: Cambridge, MA, USA, 2001; Volume 1, p. 12.

96. Halpern, J.Y.; Pass, R. Algorithmic rationality: Game theory with costly computation. *J. Econ. Theory* **2015**, *156*, 246–268. [CrossRef]

97. Russell, S.J. Rationality and intelligence. *Artif. Intell.* **1997**, *94*, 57–77. [CrossRef]

98. Cuzzolin, F.; Morelli, A.; Cirstea, B.; Sahakian, B.J. Knowing me, knowing you: Theory of mind in AI. *Psychol. Med.* **2020**, *50*, 1057–1061. [CrossRef]

99. Rabinowitz, N.; Perbet, F.; Song, F.; Zhang, C.; Eslami, S.A.; Botvinick, M. Machine theory of mind. In Proceedings of the International Conference on Machine Learning, Vienna, Austria, 10–15 July 2018; pp. 4218–4227.

100. Estes, D.; Bartsch, K. Theory of mind: A foundational component of human general intelligence. *Behav. Brain Sci.* **2017**, *40*, 1–3. [CrossRef]

101. Doshi-Velez, F.; Kortz, M.; Budish, R.; Bavitz, C.; Gershman, S.; O'Brien, D.; Scott, K.; Schieber, S.; Waldo, J.; Weinberger, D. Accountability of AI under the law: The role of explanation. *arXiv* **2017**, arXiv:1711.01134. [CrossRef]

102. Porayska-Pomsta, K.; Rajendran, G. Accountability in human and artificial intelligence decision-making as the basis for diversity and educational inclusion. In *Artificial Intelligence and Inclusive Education*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 39–59.

103. Liu, H.-W.; Lin, C.-F.; Chen, Y.-J. Beyond State v Loomis: Artificial intelligence, government algorithmization and accountability. *Int. J. Law Inf. Technol.* **2019**, *27*, 122–141. [CrossRef]

104. Habli, I.; Lawton, T.; Porter, Z. Artificial intelligence in health care: Accountability and safety. *Bull. World Health Organ.* **2020**, *98*, 251. [CrossRef]

105. Lepri, B.; Oliver, N.; Letouzé, E.; Pentland, A.; Vinck, P. Fair, transparent, and accountable algorithmic decision-making processes. *Philos. Technol.* **2018**, *31*, 611–627. [CrossRef]

106. van Nuenen, T.; Ferrer, X.; Such, J.M.; Coté, M. Transparency for whom? assessing discriminatory artificial intelligence. *Computer* **2020**, *53*, 36–44. [CrossRef]

107. Haibe-Kains, B.; Adam, G.A.; Hosny, A.; Khodakarami, F.; Waldron, L.; Wang, B.; McIntosh, C.; Goldenberg, A.; Kundaje, A.; Greene, C.S. Transparency and reproducibility in artificial intelligence. *Nature* **2020**, *586*, E14–E16. [CrossRef] [PubMed]

108. Wischmeyer, T. Artificial intelligence and transparency: Opening the black box. In *Regulating Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 75–101.

109. Larsson, S.; Heintz, F. Transparency in artificial intelligence. *Internet Policy Rev.* **2020**, *9*, 1–16. [CrossRef]

110. Felzmann, H.; Villaronga, E.F.; Lutz, C.; Tamò-Larrieux, A. Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data Soc.* **2019**, *6*, 2053951719860542. [CrossRef]

111. Gundersen, O.E.; Kjensmo, S. State of the art: Reproducibility in artificial intelligence. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.

112. Vollmar, R. John von Neumann and Self-Reproducing Cellular Automata. *J. Cell. Autom.* **2006**, *1*, 353–376.

113. Gabor, T.; Illium, S.; Zorn, M.; Linnhoff-Popien, C. Goals for Self-Replicating Neural Networks. In Proceedings of the ALIFE 2021: The 2021 Conference on Artificial Life, Prague, Czech Republic, 19–23 July 2021.

114. Spector, L. Evolution of artificial intelligence. *Artif. Intell.* **2006**, *170*, 1251–1253. [CrossRef]

115. Russell, S.; Dewey, D.; Tegmark, M. Research priorities for robust and beneficial artificial intelligence. *AI Mag.* **2015**, *36*, 105–114. [CrossRef]

116. Osugi, T.; Kim, D.; Scott, S. Balancing exploration and exploitation: A new algorithm for active machine learning. In Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05), Houston, TX, USA, 27–30 November 2005; p. 8.

117. Črepinšek, M.; Liu, S.-H.; Mernik, M. Exploration and exploitation in evolutionary algorithms: A survey. *ACM Comput. Surv.* **2013**, *45*, 1–33. [CrossRef]

118. Lin, L.; Gen, M. Auto-tuning strategy for evolutionary algorithms: Balancing between exploration and exploitation. *Soft Comput.* **2009**, *13*, 157–168. [CrossRef]

119. Sledge, I.J.; Príncipe, J.C. Balancing exploration and exploitation in reinforcement learning using a value of information criterion. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2816–2820.

120. Menzies, T.; Pecheur, C. Verification and validation and artificial intelligence. *Adv. Comput.* **2005**, *65*, 153–201.

121. Xiang, W.; Musau, P.; Wild, A.A.; Lopez, D.M.; Hamilton, N.; Yang, X.; Rosenfeld, J.; Johnson, T.T. Verification for machine learning, autonomy, and neural networks survey. *arXiv* **2018**, arXiv:1810.01989.

122. Wu, T.; Dong, Y.; Dong, Z.; Singa, A.; Chen, X.; Zhang, Y. Testing Artificial Intelligence System Towards Safety and Robustness: State of the Art. *IAENG Int. J. Comput. Sci.* **2020**, *47*, 1–14.

123. Zhang, J.; Wang, H.-S.; Zhou, H.-Y.; Dong, B.; Zhang, L.; Zhang, F.; Liu, S.-J.; Wu, Y.-F.; Yuan, S.-H.; Tang, M.-Y. Real-world verification of artificial intelligence algorithm-assisted auscultation of breath sounds in children. *Front. Pediatr.* **2021**, *9*, 152. [CrossRef] [PubMed]

124. Gordon-Spears, D.F. Asimov's laws: Current progress. In Proceedings of the International Workshop on Formal Approaches to Agent-Based Systems, Greenbelt, MD, USA, 29–31 October 2002; pp. 257–259.

125. Haddadin, S. *Towards Safe Robots: Approaching Asimov's 1st Law*; Springer: Berlin/Heidelberg, Germany, 2013.

126. Murphy, R.; Woods, D.D. Beyond Asimov: The three laws of responsible robotics. *IEEE Intell. Syst.* **2009**, *24*, 14–20. [CrossRef]

127. Yampolskiy, R.; Fox, J. Safety engineering for artificial general intelligence. *Topoi* **2013**, *32*, 217–226. [CrossRef]

128. He, X.; Zhao, K.; Chu, X. AutoML: A survey of the state-of-the-art. *Knowl.-Based Syst.* **2021**, *212*, 106622. [CrossRef]

129. Sinha, A.; Tiwari, S.; Deb, K. A population-based, steady-state procedure for real-parameter optimization. In Proceedings of the 2005 IEEE Congress on Evolutionary Computation, Edinburgh, UK, 2–5 September 2005; Volume 1, pp. 514–521.

130. Thiebes, S.; Lins, S.; Sunyaev, A. Trustworthy artificial intelligence. *Electron. Mark.* **2021**, *31*, 447–464. [CrossRef]

131. Kaur, D.; Uslu, S.; Rittichier, K.J.; Durresi, A. Trustworthy Artificial Intelligence: A Review. *ACM Comput. Surv. (CSUR)* **2022**, *55*, 1–38. [CrossRef]

132. Berner, J.; Grohs, P.; Kutyniok, G.; Petersen, P. The modern mathematics of deep learning. *arXiv* **2021**, arXiv:2105.04026.

133. Wang, Y. A cognitive informatics reference model of autonomous agent systems (AAS). *Int. J. Cogn. Inform. Nat. Intell.* **2009**, *3*, 1–16.

134. Wang, Y. The theoretical framework of cognitive informatics. *Int. J. Cogn. Inform. Nat. Intell.* **2007**, *1*, 1–27. [CrossRef]

135. Wang, Y. Concept algebra: A denotational mathematics for formal knowledge representation and cognitive robot learning. *J. Adv. Math. Appl.* **2015**, *4*, 61–86. [CrossRef]

136. Chen, R.J.; Lu, M.Y.; Chen, T.Y.; Williamson, D.F.; Mahmood, F. Synthetic data in machine learning for medicine and healthcare. *Nat. Biomed. Eng.* **2021**, *5*, 493–497. [CrossRef] [PubMed]

137. El Emam, K.; Mosquera, L.; Hoptroff, R. *Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data*; O'Reilly Media: Sebastopol, CA, USA, 2020.

138. Patterson, D.; Gonzalez, J.; Le, Q.; Liang, C.; Munguia, L.-M.; Rothchild, D.; So, D.; Texier, M.; Dean, J. Carbon emissions and large neural network training. *arXiv* **2021**, arXiv:2104.10350.

139. Yampolskiy, R.V. Artificial intelligence safety engineering: Why machine ethics is a wrong approach. In *Philosophy and Theory of Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 389–396.

140. Papernot, N.; McDaniel, P.; Sinha, A.; Wellman, M.P. Sok: Security and privacy in machine learning. In Proceedings of the 2018 IEEE European Symposium on Security and Privacy (EuroS&P), London, UK, 24–26 April 2018; pp. 399–414.

141. Goertzel, B. Human-level artificial general intelligence and the possibility of a technological singularity: A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil. *Artif. Intell.* **2007**, *171*, 1161–1173. [CrossRef]

142. Yampolskiy, R.V. AI-complete, AI-hard, or AI-easy–classification of problems in AI. In Proceedings of the the 23rd Midwest Artificial Intelligence and Cognitive Science Conference, Cincinnati, OH, USA, 21–22 April 2012.

143. Lewis, P.R.; Chandra, A.; Parsons, S.; Robinson, E.; Glette, K.; Bahsoon, R.; Torresen, J.; Yao, X. A survey of self-awareness and its application in computing systems. In Proceedings of the 2011 Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops, Ann Arbor, MI, USA, 3–7 October 2011; pp. 102–107.

144. Carden, J.; Jones, R.J.; Passmore, J. Defining self-awareness in the context of adult development: A systematic literature review. *J. Manag. Educ.* **2022**, *46*, 140–177. [CrossRef]

145. Cook, S.H. The self in self-awareness. *J. Adv. Nurs.* **1999**, *29*, 1292–1299. [CrossRef]

146. Gallup, G.G., Jr. Self-awareness and the emergence of mind in primates. *Am. J. Primatol.* **1982**, *2*, 237–248. [CrossRef]

147. Wong, P.T. Meaning management theory and death acceptance. In *Existential and Spiritual Issues in Death Attitudes*; Taylor & Francis Group: Milton Park, UK, 2008; pp. 65–87.

148. Bering, J.M. The folk psychology of souls. *Behav. Brain Sci.* **2006**, *29*, 453–462. [CrossRef]

149. Park, C.L. Religion and meaning. In *Handbook of the Psychology of Religion and Spirituality*; The Guilford Press: New York, NY, USA, 2013; pp. 357–379.