The AI Methods, Capabilities and Criticality Grid

A Three-Dimensional Classification Scheme for Artificial Intelligence Applications

Thomas Schmid^{1,2} · Wolfgang Hildesheim³ · Taras Holoyad⁴ · Kinga Schumacher⁵

Received: 15 October 2020 / Accepted: 14 June 2021 / Published online: 2 August 2021
© The Author(s) 2021

Abstract

Many artificial intelligence (AI) technologies developed over the past decades have reached market maturity and are now being commercially distributed in digital products and services. Therefore, national and international AI standards are currently being developed in order to achieve technical interoperability as well as reliability and transparency. To this end, we propose to classify AI applications in terms of the algorithmic methods used, the capabilities to be achieved and the level of criticality. The resulting three-dimensional classification scheme, termed the AI Methods, Capabilities and Criticality (AI-MC²) Grid, combines current recommendations of the EU Commission with an ethical dimension proposed by the Data Ethics Commission of the German Federal Government (Datenethikkommission der Bundesregierung: Gutachten. Berlin, 2019). As a whole, the AI-MC² Grid allows not only to gain an overview of the implications of a given AI application as well as to compare efficiently different AI applications within a given market or implemented by different AI technologies. It is designed as a core tool to define and manage norms, standards and compliance of AI applications, but helps to manage AI solutions in general as well.

Keywords Artificial intelligence · AI applications · AI methods · AI capabilities · AI criticality · AI classification scheme

1 The Need for a Practical Classification Scheme

Recently, the European Union has initialized a regulation proposal on harmonized rules for artificial intelligence (AI) [2, 3]. In particular, an obligatory use of European Standards

and Technical Specifications for conformity assessments of so-called high-risk AI Systems is proposed. Similar efforts are being undertaken in Germany, where the national institute for norms and standards has set up a roadmap [4] in order to put reliable industry standards into practice. Both endeavors require robust ways to classify and assess AI products and services. Such standards need to fulfill several practical needs of AI users, developers, leaders and regulators by providing, for instance

- a better market overview for customers and policy makers;
- an easier identification of potential re-use of established AI components;
- a framework to orient investors and strategic decision makers;
- a sound foundation for risk and compliance assessments;
- help to implement norms, standards and compliance of AI applications;

This standardization roadmap is a process long awaited by many AI stakeholders. After decades of basic research, the

✉ Thomas Schmid
schmid@informatik.uni-leipzig.de

Wolfgang Hildesheim
hildeshe@de.ibm.com

Taras Holoyad
taras.holoyad@bnetza.de

Kinga Schumacher
kinga.schumacher@dfki.de

¹ Lancaster University in Leipzig, Leipzig, Germany

² Universität Leipzig, Leipzig, Germany

³ IBM Deutschland, Hamburg, Germany

⁴ Bundesnetzagentur, Mainz, Germany

⁵ Deutsches Forschungszentrum für Künstliche Intelligenz, Berlin, Germany

Fig. 1 Schematic display of the AI Methods, Capabilities and Criticality (AI-MC²) Grid

trend to use and develop AI for commercial products has accelerated considerably since 2010 [5, 6]. This is primarily driven by both the huge increases in the volume of available digital data, as well as the increased in processing power that can now run improved methods such as neural networks faster and more cost-effectively [7, 8]. This opens up new possibilities for human-machine interfaces based on AI applications that can either carry out physical movements enabling surveillance robots, or simulate chats and speech, thereby enabling automation of call-centre dialogs and the filling out forms.

Today, a large variety of such AI applications can be purchased or rented as products or services by companies and users to increase the productivity of business processes or to enable innovations in business models [9]. As of 2018, business analysts counted more than 1000 AI start-ups for the US market alone [10]. Typical AI services offered “out-of-the-box” are image recognition, video analysis, language-to-text conversion, text-to-speech conversion, machine translation, text analysis or automation of chats and email. To facilitate the use of such applications, commercial services and products are typically offered from public or private cloud environments. This allows users to start immediately with the adaptation of the service or product to their own needs without spending time and effort on building hardware and software.

Together with such commercial AI applications, specialized software markets have emerged over the last decades [11], which are uniformly designated worldwide and regularly monitored by independent market analysts (e.g. IDC, Gartner, Forrester, etc.). Potential users, projects and investors can thus be well informed about the status of the capabilities. These markets of AI products and services can

be roughly divided into the main areas of business intelligence [12] and decision making [13], AI-based customer interaction [14], AI-based services [15] and AI development environments and tools.

2 Three Dimensions of AI Applications

Due to the growing diversity of the field, it is becoming increasingly important to maintain an efficient overview of application scenarios and their characteristics. To this end, we introduce the Methods, Capabilities and Criticality (AI-MC²) Grid (Fig. 1), which reflects not only a matrix of AI methods and capabilities, but also the degree of criticality of an AI application. This scheme allows top-level characterization and comparison of AI applications along three dimensions, which are described in more detail in the following.

Based on a position paper of the AI Expert Group of the European Union [16], we first distinguish between AI methods and AI capabilities. By representing these aspects on separate axes, a two-dimensional classification matrix is created that allows to display which core methods are used to achieve which capabilities. In the following, we summarize the currently most important methods and capabilities based not only on standard text books [17, 20–22], but also in reflection of other related work, such as classification schemes of research institutes [23, 24], AI organizations (ACM,¹ Plattform Lernende Systeme²) and recognized community work from AAAI and IEEE conferences.

¹ <https://dl.acm.org/ccs> (2020.09.29)

² www.plattform-lernende-systeme.de (2020.09.29)

With the large-scale use in a large variety of applications, however, AI products and services do not only face technical aspects. Public debate has recently put particular emphasis on the aspects information rights and obligations and liability by human decision-makers [1]. By introducing an ethical and legal framework, legislative actors in Europe plan to provide regulatory “guidelines for transparency, explainability and traceability” of algorithmic systems. Real-world AI applications should therefore also be characterized regarding their impact on their users. Classifying AI applications based on a combination of ethical, legal and technical criteria allows to avoid shortcomings during development and deployment as well as efficient assessments of conformity and quality. Moreover, use case-related evaluations of AI products and services will play a key role in the social acceptance of AI as such.

3 Dimension 1: AI Methods

Today’s AI is built on a variety of methods. Due to historical developments, they are often roughly distinguished between symbolic or sub-symbolic—sometimes even numerical—AI methods. Both paradigms indeed form the foundation of many AI approaches [18]. On the side of symbolic methods, techniques of knowledge representation and logical reasoning are prominent, while the side of sub-symbolic methods is primarily represented by neural networks and machine learning techniques. Yet, this traditional distinction is not comprehensive. It neglects methods of problem solving, optimizing, planning and decision making, of which many combine symbolic and sub-symbolic techniques. In addition, the developments of the last decades have further blurred traditional boundaries, and more and more combined or hybrid approaches are coming to the fore, e.g. the entire field of hybrid learning [18, 19].

Tables 1 and 2 give an overview of these methods over three levels of granularity and name well-known representatives of each technology. This scheme is a momentary, non-exhaustive snapshot, and will be subject to additions with new AI technologies arising in the future. Also, it is often not possible to draw a clear line between categories, and methods may belong to several categories. In such cases, methods were assigned according to Russell and Norvig [17] or according to the category for which it was originally proposed. To this end, an online version of the full AI-MC² Grid is considered to foster both discussion and refinement, and to allow for integration of new developments.

3.1 Problem Solving, Optimizing, Planning and Decision Making

From a historical point of view, approaches for problem solving, optimizing, planning and decision making were among the earliest AI methods to be developed. Problem solving describes goal-based search strategies and intelligent agents that solve problems by formulating a goal, searching for the right sequence of actions with a defined problem as input in order to execute the solution and reach the goal. In competitive multi-agent environments, where goals conflict with each other, adversarial and constraint-based search are used to solve complex problems [25, 26].

In contrast to problem solving methods, which explore search spaces systematically, optimization algorithms do not worry about the path towards reaching the goal but focus on the optimum solution. They may be divided in statistical optimization, using local search algorithms, search in unknown and continuous spaces, partial observation and dynamic programming, and bio-inspired optimization [17]. Bio-inspired optimization can be classified into three classes: evolutionary, swarm-based, and genetic or ecology-based [27].

Planning methods may be autonomous or semi-autonomous techniques like steady state search, planning graphs, hierarchical planning, non-deterministic planning, time and resource planning and generation of plans [17]. In contrast to planning, plan recognition models or methods such as deductive and synthesis plan recognition, library-based plan recognition and planning by abductive reasoning need to represent actual events or actions that have happened, as well as proposing hypothetical explanations [28]. Methods of planning play a role in robotics, dialog systems and human-machine-interaction.

Decision making or decision analysis is an engineering discipline that addresses the pragmatics of applying decision theory to defined problems [29]. There are several approaches towards decision making such as process models, value of information, decision networks, expert systems, sequential decision making and iteration models [17].

3.2 Knowledge Representation and Reasoning

Symbolic AI methods are characterized by a deductive approach, i.e. by the algorithmic application of logical rules or relationships to individual cases. Core concepts of symbolic AI are on the one hand techniques for representing knowledge and on the other hand methods for applying this knowledge onto a given input. Knowledge may

Table 1 An overview of AI methods—symbolic and traditional AI

CLASSIFICATION OF METHODS			EXAMPLES
PROBLEM SOLVING, OPTIMIZATION, PLANNING AND DECISION MAKING	Problem Solving	Problem Solving Agents & Searching Strategies	Uninformed and Informed Search Strategies Adversarial Search (Game Theory) Constraint based Search, Constraint Solving
	Optimization	Statistical Optimization	Local Search Optimization Search in Continuous Spaces Partial Observation Search Search in Unknown Spaces Dynamic Programming
			Evolutionary Algorithms Genetic Algorithms, Genetic Programming Swarm Intelligence
	Planning and Plan Recognition	Autonomous and Semi-autonomous Planning	Steady State Search Planning Graphs Hierarchical Planning Non-deterministic Planning Time and Resource Planning Generation of Plans
			Abductive Plan Recognition Deductive Plan Recognition Library based Plan Recognition Synthesis Plan Recognition
		Plan Recognition Methods	Process Models Value of Information Decision Networks Expert Systems Sequential Decision Making Iteration Models
			Ontology (RDF, RDFS and OWL) Taxonomy Interpretation Calculus Deduction Abduction Ontology Mapping Knowledge Graphs and Networks Existence Graphs Graph Traversal Algorithms Mapping Semantic Web
			Propositional Logics and Predicate Logics Higher Order Logics, Non-monotonic Logics Temporal and Modal Logics
			Resolution and Connectivity Verification SAT and SMT Solver Model Checking Tactical Theorem Verification
			Bayesian Rule
KNOWLEDGE REPRESENTATION AND REASONING	Knowledge Representation	Ontological Engineering	Ontology (RDF, RDFS and OWL) Taxonomy Interpretation Calculus Deduction Abduction Ontology Mapping Knowledge Graphs and Networks Existence Graphs Graph Traversal Algorithms Mapping Semantic Web
			Propositional Logics and Predicate Logics Higher Order Logics, Non-monotonic Logics Temporal and Modal Logics
		Knowledge Graphs and Semantic Networks	Resolution and Connectivity Verification SAT and SMT Solver Model Checking Tactical Theorem Verification
			Bayesian Rule
			Bayesian Networks
	Reasoning	Formal Verification	Precis Inference Approximate Inference Markov-Chains
			Relational Probability Models in Closed And Open Universes Hidden Markov Models Kalman Filter Dynamic Bayesian Networks
			Reasoning with Default Information Truth Management System (TMS) Rule-based Reasoning with Security Fuzzy Sets and Logic Dempster-Shafer Theory
	Uncertain Knowledge	Interactive Verification	Spatial Reasoning Case-based Reasoning (CBR) Qualitative Physics Psychological Reasoning
	Probabilistic Reasoning	Uncertainty Quantification	
	Non-probabilistic Reasoning	Representation of Uncertain Knowledge	
	Other Approaches for Uncertain Reasoning	Bayesian Inference	

be represented either as safe or uncertain knowledge. By means of reasoning, conclusions may be drawn from such knowledge.

Formal knowledge representation includes concepts like ontologies, semantic networks, knowledge graphs and knowledge maps, which accumulate and systematize information into structures, syntax, semantics and semiotics [30]. The focus of standardized description languages like Resource Description Framework and Web Ontology Language lays on the creation of unique specifications for objects, characteristics and general concepts by logical relationships [31]. Using these and further semantic web standards and technologies, logically related data can be shared across domains of different applications, facilitating semantic interoperability [30]. In general, fundamental concepts of ontological engineering are based on taxonomy, calculus, deduction, abduction and processing and modelling of ontologies. Furthermore, logical relationships and abstractions of domains may be established by knowledge graphs, semantic networks and knowledge mapping. In case of a graph-based abstraction of knowledge, graph traversal algorithms provide common solutions of problems with regard to searching, verifying and updating vertices. Moreover, logically related data can be modelled by propositional, predicate, high order, non-monotonic, temporal and modal logics [17].

With certain knowledge, the application of formal knowledge is often operationalized by means of the classical methods of logical reasoning. In particular, satisfiability and other techniques of formal verification may be applied for this [32]. For reasoning on the basis of uncertain knowledge, probabilistic approaches are widespread, but non-probabilistic approaches have been proposed, too [33]. In the course of probabilistic reasoning, information can be derived by sampling a knowledge base and processed by relational probabilistic models or the concept of Bayesian Inference [34]. Regarding uncertain knowledge, the Bayesian Rule dominates the AI field in uncertainty quantification [17]. Non-probabilistic reasoning may be applied for ambiguous information in case of vagueness under consideration of evidence. In these situations, a truth management system and reasoning with default information can be used for qualitative approaches [35]. Furthermore, methods of non-probabilistic reasoning may be implemented using rule-based approaches or fuzzy sets. Moreover, the Dempster-Shafer Theory (reasoning with believe function) represents a common approach for non-probabilistic reasoning, where all available evidence is combined for the calculation of a degree of believe [36]. Other approaches for uncertain reasoning involve spatial reasoning, case-based reasoning [37], qualitative reasoning and psychological reasoning [17].

3.3 Machine Learning

In contrast to symbolic AI, subsymbolic AI methods are characterized by an inductive procedure, i.e. by the algorithmic derivation of general rules or relationships from individual cases. To this end, two major machine learning approaches are typically distinguished: supervised learning which has given target parameters and unsupervised learning, which does not.

Supervised learning is typically used to implement regression or classification tasks. If an estimation of the probability distribution for input and output variables is generated, such as in Naive Bayes [38] or Hidden Markov Models [39], one speaks of generative procedures. Practical applications, however, are dominated by discriminative procedures like logistic regression [40], decision trees [41] or neural networks [42]. Neural networks are considered particularly flexible, since they may theoretically learn any mathematical function without any prior knowledge [43]. Also widely used are support vector machines [44], but their use requires the assumption and definition of a kernel function.

Unsupervised learning methods, on the other hands, are typically implemented for clustering or dimension reduction tasks. One of the oldest and most widely used algorithms is k-means [45]. In addition to other statistically motivated methods such as hierarchical clustering [46], biologically inspired algorithms like Kohonen's self-organizing map [47] or Grossberg's adaptive resonance theory [48] have been suggested. Also for regression-tasks some unsupervised procedures have been proposed (e.g. [49, 50]), which are mainly used for dimensional reduction [51].

Not all learning algorithms may be clearly identified as supervised or unsupervised. For example, a multilayer perceptron, i.e. a supervised procedure, may be used to map a given data set to itself [52]. If the output layer of such an autoencoder is subsequently removed, a net remains, which maps the original data set to a new data set of lower dimension according to the number of hidden neurons. This principle is, for example, an important basis of Deep Learning [53]. Another example of intermediate forms of machine learning, are algorithms of so-called semi-supervised learning, in which supervised and unsupervised learning are combined and thus a given target value is required for only a part of the data used [54]. This not only allows the analysis of incomplete data sets, but in some cases even achieves better results than classical supervised learning methods (e.g. [55–57]). For semi-supervised learning algorithms, however, assumptions about distribution densities must be made in advance. In case of unfavorable assumptions, the results can be significantly worse than with a supervised learning procedure [58]. Furthermore, there is little evidence so far that

Table 2 An overview of AI methods—subsymbolic and hybrid AI

CLASSIFICATION OF METHODS			EXAMPLES
MACHINE LEARNING	Supervised Learning	Neural Networks	Multi-layer Perceptron
			Learning Vector Quantization (LVQ)
			Radial Basis Networks (RBF)
			Adaptive Resonance Theory (ART)
			Convolutional Neural Networks (CNN)
			Recurrent Neural Networks (RNN)
			Time-Delay Networks (TDNN)
			Long-Short Term Memory (LSTM)
			Hopfield Networks
			Boltzmann Machine
		Statistical Learning	Decision Trees
			Random Forests
			Support Vector Machines (SVM)
		Probabilistic Methods	Naïve-Bayes
			Fuzzy Classifier
	Unsupervised Learning	Clustering	k-Means
			Hierarchical Clustering
			DBSCAN
			Fuzzy Clustering
			Self-Organizing Map
		Dimensionality Reduction	Autoencoder
			Principal Component Analysis
		Probabilistic Methods	Fuzzy k-Means
	Semi-supervised Learning	Statistical Approaches	Expectation Maximization with Generative Models
			Transductive Support Vector Machines
		Modified Learning Concepts	Self-training
			Co-training
	Reinforcement Learning	Graph-based Learning	Graph-based Approaches
		Temporal Difference Learning	Q-Learning
			State-action-reward-state-action (SARSA)
		Monte Carlo Methods	Markov Chain Monte Carlo
HYBRID LEARNING	Hybrid Neural Systems	Unified Neural Architectures	Constructivist Machine Learning
		Transformation Architectures	Rule Extraction for Neural Networks
		Hybrid Modular Architectures	Neuro-Fuzzy Expert Systems
	Learning with Knowledge	Learning by Logic and Deduction	Current-Best-Learning
		Inductive Logical Programming	Sequential Covering-Algorithm
			Constructive Induction Algorithms
		Explainable Artificial Intelligence	Local Interpretable Model-agnostic Explanations (LIME)
		Relevance-based Learning	
	Conversational Learning	Active Dialog Learning	Supervised Conversational Learning
			Reinforcement Conversational Learning

supervised and unsupervised learning in the human brain take place in similarly close association [59].

An alternative procedure is followed by reinforcement learning [60], which requires feedback for predictions, but not the exact target value. Analogously to behavioristic learning, the method only takes into account whether the learning goal has been reached or not. Reinforcement learning has been proven particularly helpful in the areas of robotics [61] and adaptive control [62].

3.4 Hybrid Learning

Hybrid learning methods are characterized by combining concepts from the previously presented methods, e.g. when training neural networks by applied evolutionary algorithms to adapt network weights [63]. Such combinations have been proposed for a wide variety of areas like finance applications [64], oceanographic forecast [65], outpatient visits forecasting [66], compounding meta-atoms into metamolecules [67] and classification of tea specimens [68].

Due to the scientific creativity in this area, it is challenging to display a comprehensive overview. A large portion of hybrid AI, however, focuses on combining symbolic and subsymbolic AI in order to operate both inductively and deductively. Recent research activities investigate, e.g., combining machine learning and knowledge engineering [69]. A prominent subfield are hybrid neural systems, which may be further separated into unified neural architectures, transformation architectures and hybrid modular architectures [70]. In contrast to classical subsymbolic techniques, such techniques either allow extraction of rules or use an additional form of knowledge representation. In contrast to classical symbolic methods, however, such knowledge representations are often algorithmically modified on the basis of given data.

In a wider sense, hybrid learning may also be thought of as learning with knowledge. To this end, further approaches have been proposed over the last decades, e.g. learning by logic and deduction [71], inductive logical programming [72], explainable artificial intelligence and relevance-based learning [73]. A novel approach of hybrid AI is conversational learning or active dialog learning, which aims to improve the performance of machine learning by incorporating human knowledge collected in dialog [74].

4 Dimension 2: AI Capabilities

AI as a scientific discipline is inspired by human cognitive capabilities [17], which have been investigated widely by psychology, educational and social sciences. In educational contexts, human capabilities have been defined and assessed since the 1950s based on so-called learning outcomes [103]. Taxonomies like the works of Benjamin S.

Bloom and Robert M. Gagné reflect human capabilities and are the basis for European educational systems [104]. Gagné distinguishes five essential categories [78]: verbal information, intellectual abilities, cognitive strategies, attitudes, and motor skills. Bloom [77], on the other hand, distinguishes a cognitive, an affective and a psychomotor domain, which are further subdivided into more specific capabilities. Based on Bloom's ideas, taxonomies for the affective domain referring to the processing of sensory impressions [97], the psychomotor domain referring to the control and coordination of muscles [85] and the cognitive domain [77] were postulated. This three-fold structure is reflected in the following with the three general capabilities named sense, process and understand, and act (cf. Table 3).

Compared against such taxonomies, all current AI-based systems implement only a subset of human cognition. At the same time, many existing AI systems and applications aim at implementing additional functionality like enhanced sensory input and intelligent interaction with the environment. Based on these observations, both existing and potentially achievable AI capabilities may be roughly divided into the areas of sensing, processing and understanding, acting and communicating. While this four-fold distinction of capabilities takes common knowledge from psychological and educational research into account, it is primarily intended to structure capabilities that can be implemented by AI systems today. In some cases, additional details might be added to discuss a specific AI solution and the underlying process steps. Table 3 gives an overview of the suggested classification of AI capabilities.

4.1 Sense

Traditionally, sensing is related to the human sensory organs. Ancient Greek philosophers, e.g., distinguished the five senses vision, hearing, smell, taste and contact [86]. More recently, a pronounced distinction is made between sensory organs, which transduce sensory stimuli and act as a form of perceptual preprocessors, and sensory modalities, which basically describe the output of sensory organs directed at subsequent cognitive processing. Regarding stimuli, it should be noted that the number of stimuli directly perceivable by humans is smaller than for stimuli perceivable by specialized technical sensors, which are available today for a wide range of acoustic, biological, chemical, electric, magnetic, optical, mechanical, radiation and thermal stimuli [87].

Describing human perception, many researchers today focus on sensory modalities. The term modality is often used to describe the encoding or "mode of presentation" resulting from transduction of a sensory input [88]. Olfaction, e.g., may be represented by the four primary modes fragrant, acid, burnt, and caprylic [89]; other classification systems

assume up to nine primary olfactory modes [90]. Apart from sensory transduction and representation, some researchers also see, e.g., the utilization of the modality by the organism “to guide intentional action” as a key characteristic of a modality [91].

With a modality-based view of perception, the classical five-senses scheme has been questioned and several alternative schemes have been proposed [92]. While the number of senses in such classification schemes vary between 8 and 17, consent appears to grow that in addition to the classical five senses direct at external input, also internal senses like body awareness and balance contribute to human perception by detecting disturbances and anomalies [91]. Also, touching is often considered a multi-faceted sense that includes sensing of temperature, pressure and pain [93].

With respect to these findings, we suggest to align AI capabilities in the area of perception with human sensory modalities. This allows, e.g., to reflect that AI has recently made significant progress in the ability to transduce images, auditory and haptical signals into processable information. AI applications for sensing scent and taste, on the other hand, have been investigated, but are currently relatively rare in practice.

4.2 Process and Understand

The ability to process and understand information is an essential ability for intelligent behavior. From a practical point of view, adopting a learning outcomes-based classification scheme appears to be a good choice to specify this ability. To this end, Bloom’s taxonomy for cognitive learning outcomes is particularly well known. Having been criticized from various disciplines [94–96], it was later modified by Anderson and Krathwohl [75]. Bloom’s revised taxonomy distinguishes human abilities in a primary dimension by the four domains of factual, conceptual, procedural and meta-cognitive cognition [80].

Factual cognition, as the least complex domain, includes capabilities to process and understand knowledge of terminology and knowledge of specific details and elements. Conceptual cognition includes capabilities to process and understand knowledge of classifications and categories, principles and generalizations, and knowledge of theories, models, and structures. Procedural cognition refers to processing knowledge of subject-specific skills and algorithms, of subject-specific techniques and methods, and of criteria for determining when to use appropriate procedures. Meta-cognitive cognition includes capabilities for processing and understanding strategic knowledge, knowledge about cognitive tasks (including contextual and conditional knowledge) and self-knowledge.

In a second level, these domains are further distinguished by six cognitive levels. For conceptual knowledge, e.g., the basic capabilities to recognize or classify are separated from medium-complex capabilities to provide or differentiate information and from the advance capabilities to determine or assemble information. Using both cognitive dimensions, Bloom’s revised taxonomy allows to distinguish up to 24 human cognitive capabilities.

Capabilities included the category of processing and understanding cover a wide range of human cognition. Regarding currently available AI capabilities, this scheme allows in particular to reflect the ability to evaluate, remember, decide or predict. They include, e.g., the abilities or perceptual fusion, memory and models, explanation and self-regulation. These rational capabilities represent the core of many advanced AI systems and are often applied in combination with capabilities to sense, act or communicate.

4.3 Act

For humans, the ability to act is a constitutive ability. In a more general sense, action may be related to a human or non-human agent. It may be described as something that an agent does and that was “intentional under some description” [84]. AI action may be distinguished into physical and non-physical action. Often, it is implemented by combining mechatronic and software components in robots or software robots. In particular, the field of robotics describes mechanically or physically executed activities like robot perception, motion planning, sensor technology and manipulators, kinematics and dynamics as well as the field of human-robot interaction, since this form of interaction focuses on physical human-machine interaction. These capabilities are roughly inspired by human abilities of the control and coordination of muscles [85]. The methods used for software agents are dependent on the particular goal or task of the agent itself. For example, such autonomous agents are essential in the area of process automation.

4.4 Communicate

Although communication is a well-known and well-researched human capability, communication researchers traditionally found it hard to consent to a common definition or taxonomy of it [98]. In fact, the number of concepts defining communication was large enough for researchers to suggest a three-dimensional distinction scheme for them based on level of observation, intentionality and judgment [99]. One of the simplest technologically motivated definitions understands communication as the transmission of information between given subjects [100]. In a complementary approach, communication is defined via the ability to communicate, which is described

Table 3 An overview of AI capabilities

CLASSIFICATION OF CAPABILITIES			EXAMPLES		
SENSE	External	See	Optical Character Recognition Object Recognition Gesture Recognition		
		Hear	Speech Recognition Sound Recognition		
		Smell	Fragrant Detection Acid Detection Burnt Detection Caprylic Detection		
		Taste	Sweetness Detection Sourness Detection Saltiness Detection Bitterness Detection Umami Detection		
		Touch	Temperature Detection Pressure Detection Pain Detection		
		Internal	Body Awareness	Detection of Self Movement Detection of Body Position	
		Balance	Balance Detection		
		PROCESS AND UNDERSTAND	Factual	List	Context-specific Terminology
				Summarize	Automated Report Generation
				Respond	Data Association
Select	Semantic Search				
Check	Parsing (Syntactic Analysis)				
Generate	Deductive Knowledge Extraction				
Conceptual	Recognize		Named Entities Recognition		
	Classify		Semantic Domain Recognition		
	Provide		Explanation		
	Differentiate		Disambiguation		
	Determine		Semantic Interpretation		
	Assemble		Interlingual Translation		
Procedural	Recall		Process Memory		
	Clarify		Environmental Modeling		
	Carry Out		Discourse Modeling		
	Integrate		Sensory Data Fusion		
	Judge		User Modeling		
	Design		Human Processor Models		
Metacognitive	Identify		Strategy Selection		
	Predict		State Estimation		
	Use	Transfer Approach			
	Deconstruct	Code-Switching Strategies			
	Reflect	Self-optimization Methods			
	Create	Generation of Narratives			
ACT	Physical	Motion Planning	Motion Planning with Uncertainty		
		Sensors and Manipulators	Passive und Active Sensors		
		Kinematics and Dynamics	Dynamical Movement		
		Human-robot Interaction	Multimodal Teleoperation		
Non-physical	Software Agents	Process Automation Transactional Systems Chatbots or Customer Service			
	COMMUNICATE	Natural Language Processing	Text Generation	Paraphrasing Tools	
Machine Translation			Interlingual Translation		
Text Analysis			Parsing (Syntactic Analysis)		
Information and Knowledge Extraction			Named Entities Recognition		
Information Retrieval			Semantic Search		
Document Analysis			Semantic Domain Recognition		
Spoken Dialog Systems			Clarification Dialogue Generation of Narratives		
Human-Machine Interaction		Cognitive Systems	User Modelling		
Interaction Paradigms and Modalities	Multimodal Interaction				

as the ability to process, understand and to make a distinction between utterance and information and consequently between “the information value of its content” and “the reasons for which the content was uttered” [101]; to this end, the ability to communicate is— similarly to the ability to act—considered a higher level capability that requires not only the capability to perceive or sense as a prerequisite, but also to process and understand. While types of communication are often distinguished by the medium (e.g. oral, written, etc) in popular literature, in communication research the distinction by the number of participants (intrapersonal, interpersonal, transpersonal) is a more widely recognized criterion. Also the influence of feedback on the communication process (one-way, bidirectional, omnidirectional) is often considered [102].

5 Dimension 3: AI Criticality

As AI is increasingly used in widely available products and services, the question of its damage potential has been raised [1]. Acknowledging the arguments put forward in the public discussion, the European Commission has proposed to design specific requirements for conformity assessments with regard to high-risk AI in its April 2020 regulatory approach for AI-based systems [3]. Such requirements should be laid down in standardization documents and focus on safety and fundamental rights with respect to development, market access and use of AI in Europe. Building on that, the proposal suggests to prohibit AI practices, that contravene EU values, for instance involvement of real-time remote biometric identification systems and scoring by public authorities.

One of the aspects that may influence the damage potential of algorithmic systems is the degree to which human supervision over such systems is possible or implemented. A distinction between three classes of autonomy is recommended [1]: algorithm-based applications, algorithm-driven applications and algorithm-determined applications. While algorithm-based applications basically act as assistance tools, the class of algorithm-determined applications may in some cases be operating without human supervision.

One of the concepts that have been suggested to analyze and judge potential risks is the so-called criticality pyramid, which introduces different levels of criticality [1]. It distinguishes five levels of criticality, where the damage potential of a socio-technical system grows with increasing criticality. The goal is to provide clear guidelines for reliably identifying AI products and services that need to be regulated or prohibited. And vice versa to avoid unnecessary restrictions on AI applications with little or no damage potential.

The criticality defines that two aspects for a “socio-technical” system are to be assessed: a possible occurrence of damage, e.g. caused by human beings and/or algorithm-terminated systems, and its extent, e.g. “right to privacy, fundamental right to life and physical integrity” and “non-discrimination”. Other authors emphasize that a potential damage for societies and countries may also be of economic nature, e.g., a heavy dependance on a single AI solution or AI provider, a monopoly so to speak [106]. As a consequence, all technical components (hardware, software and training data), human actors (developers, manufacturers, testers and users) as well as life cycle phases (development, implementation, conformity assessment and application) need to be reflected. This criticality of a system should be assessable by developers and testers and transparent to users and the legislative. In the following, the pyramid is used to assess the damage potential originating from AI applications. In particular, its criticality levels are used to classify AI applications.

AI products and services that belong to the criticality level 1 are “applications with no or little potential for damage” [1]. This would be assumed, e.g. for systems implementing automatic purchase recommendations or anomaly detection in industrial production. Level 1 systems should be checked for correctness, but are not subject to risk-adaptive regulation. The underlying assumption of the concept of the criticality pyramid is that a large portion of AI applications will fall into this category (cf. depiction in Fig. 2).

Level 2 marks the beginning of specific regulation. AI applications that belong to this level are “applications with a certain potential to cause damage” [1]. Examples would be systems that implement non-personalised dynamic pricing or automatic settlement of claims. Level 2 systems should have disclosure requirements on transparency. Also, misconduct audits are necessary, e.g., by analysing input-output behaviour.

For systems at level 3 like automated credit allocation or fully automated logistics (“applications with regular or significant damage potential” [1]), approval procedures, public audits and controls by third parties should be used in addition to level 2 measures.

Level 4 systems like AI-based medical diagnostics or autonomous driving (“applications with significant potential for damage” [1]) should, in addition to level 2 and 3 measures, fulfill further obligations regarding control and transparency, such as publication of algorithms and parameters and creation of an interface for direct influence on the system. Extended controls and certifications by third parties are needed.

Systems at level 5 like autonomous weapon systems and real time biometric identification systems in public (“applications with unacceptable damage potential” [1]) should be

Fig. 2 The 5-Level Criticality Pyramid as recommended by the German National Commission on Data Ethics [1]

partially or completely prohibited. Such prohibition been requested repeatedly for AI-enhanced arms and weapons [107].

6 Case Study 1: Traffic Sign Recognition

In the following, we demonstrate usage and benefits of the AI-MC² Grid with a practical example related to the field of self-driving cars. There, AI-based recognition of traffic signs is employed to detect speed limits and other information from camera images. Such systems have to deal with a wide range of different signs and variations of images to be classified accordingly. This task has been investigated in basic research for more than two decades, but has become of particular interests to the automotive industry in recent years.

To implement traffic sign recognition, the AI capability of image understanding (sense) is required. To this end, the task is similar in a way to tasks like video analysis, perceptual reasoning, scene interpretation or photometry. Capabilities to understand, act or communicate are not required for traffic sign recognition tasks, but may be combined by additional AI components connected to the recognition system. In particular, this is the typical case in self-driving cars, where the speed or direction of the car may adjust automatically according to the output of the traffic sign recognition system.

The main methodology to implement traffic sign recognition involves convolutional neural networks [108]. This supervised technique from the field of machine learning is specialized on image processing and able to cope with image variations regarding color, shape or the presence of

pictograms or text, illumination, occlusions or others. As machine learning applications are highly data-driven, the quality of the resulting system will highly depend on the data employed for training. Due to this fact, such applications are sometimes regarded as black box systems. Formal methods from the areas of knowledge representation and reasoning are usually not employed to implement this AI capability. The same holds true for the other main fields of AI methods.

While traffic sign recognition systems offer economic chances for both companies developing and companies employing such systems in automotive environments, their use also inherits substantial risks, e.g., when failing to recognize signs correctly. The criticality of this task, however, will depend on the use case. Damage potential and regulatory measurements will obviously not be the same for usage as assistance systems (e.g. in garbage trucks) as for usage as a controlling component in autonomous cars. The criticality of AI-based traffic sign recognition may reach from no or a low damage risk (level 1) to an unacceptable damage risk (level 5). In road traffic, e.g., a substantial damage risk (level 4) can be assumed for self-driving cars and additional measures like approval procedures, audits and controls by third parties should be undertaken (e.g. by using the “German Traffic Sign Recognition Benchmark” [109]).

By using the AI-MC² Grid to analyze traffic sign recognition (Fig. 3), regulators will see at a glance that while this field is a niche from a technological point of view, the implications for society are significant. Further, it becomes obvious that regulation of this area will require differentiation and a variety of regulatory actions (instead of a single measure) depending on context and criticality. A conventional

Fig. 3 Localization of use cases for traffic sign recognition within the AI-MC² Grid

Fig. 4 Localization of the main steps of automated claim settlement within the AI-MC² Grid

garbage truck with the same AI system for traffic sign recognition as a pure assistance system, e.g., does not inherit the same damage risk as an autonomus vehicle and should be regulated differently.

7 Case Study 2: Automated Settlement of Insurance Claims

We demonstrate usage and benefits of the AI-MC² Grid with a practical example from the insurance industry. For insurance companies, the settlement of customer claims is a core challenge. It may require staff to obtain and assess information from multiple sources (e.g. personal data, reviews, geographic information, coverage details or medical data) in order to decide and settle claims. The incorporation of AI techniques allows an increase in quality and

efficiency of claim settlements, in particular in terms of consistent, accurate and quick processing. Furthermore, in case of accidents, data-driven AI systems may facilitate reconstructions of circumstances to estimate the damage and carry out an assessment of legal, financial and clinical implications. Existing AI solutions allow, for example, automated settlement based on analysis of abstracted data like damage patterns from images.

Almost all types of AI capabilities may be involved in implementing automated claim settlement. Detection of damage patterns will require the capability to sense or to process and evaluate sensor data. A typical AI capability here would be image analysis. In order to decide on identified damages in a next step, the capabilities to process and understand as well as to act (software transaction and assistance systems) are required. In advanced systems, even the ability to communicate automatically

with the customer in written or spoken language may be implemented.

These capabilities may be implemented with a variety of AI methods. While image analysis (sense) is typically carried out with neural networks or other learning algorithms, the capability required to evaluate, decide or predict (understand) may be implemented with formal as well as with machine learning methods. While AI methods from hybrid learning and the field of problem solving, optimization planning and decision making may theoretically also be used to implement understanding or acting, their use in this matter is currently not very wide-spread in practice.

Assessing the criticality of a task as complex as automated claim settlement requires the assessment of the main processing steps individually, as risk assessment of each stage may vary by context. Typical main steps are damage recognition, text analysis & information or knowledge “gathering”, case evaluation and transaction processes like payment automation (Fig. 4). In case of a dented fender, a regular or significant damage risk (level 3) may represent the highest level of criticality. In comparison, the risk of damage in a nuclear power plant can be significantly higher and reach the level of an unacceptable damage risk (level 5). During AI-based claim processing of a bent fender, case evaluation and payment automation tend to have a higher risk level (here: regular or significant damage risk, level 3) in comparison to the other processing steps (level 2). This is due to the fact that assessment and settlement are completed decisively in the context of these processes.

Using the AI-MC² Grid to analyze automated settlement of claims (Fig. 4), decision makers see that certain stages like case evaluation display a higher criticality level than others. As both rule-based and “black box” methods are available to implement this stage, decision makers may then, e.g., define appropriate requirements for method-specific audits of such case evaluation stages.

8 Conclusions and Outlook

With the AI Methods, Capabilities and Criticality (AI-MC²) Grid, we have introduced a practical classification scheme for AI applications that consists of three dimensions: AI methods, AI capabilities and the criticality of the AI application. Each discussed AI application can be placed in these three dimensions and compared along this dimensions. Complex AI solutions can be analyzed by placing its individual components within the AI-MC² Grid. In this context, AI methods correspond to AI algorithms and technologies and reflect specific concepts from mathematics and computer science. AI capabilities correspond to typical process steps to build intelligent workflows and cognitive systems and reflect typically broader concepts from psychology,

educational sciences or social sciences. Once the relevant AI methods and capabilities of a given AI application are located, benefits, consequences and possible risks may be discussed in a broader context and potential alternatives may be identified more easily. The level of criticality of a given AI application is thought to correlate with potential damages, and higher levels of criticality imply higher levels of tests, controls and certification. The AI-MC² Grid can be used and displayed as a tool in one, two or three dimensions according to the needs.

With the importance, performance and complexity of AI solutions growing rapidly, the need to keep an overview of relevant AI products and services is growing rapidly as well. Same holds true for the need to manage priorities and risks in this area. The AI-MC² Grid provides a straight-forward way to gain this overview and manage the space of AI solutions not only for managers, developers and users of AI solutions. It may be in particular helpful for investors, politicians and regulators who want to understand, market and control certain AI solutions. With this good cause the AI-MC² Grid is at the heart of the German National Roadmap for Artificial Intelligence Norms [4], placed as a method and tool to support the development and management of future AI standards and norms in Germany. Other platforms, societies and programmes have the same need and could profit from the use of the AI-MC² Grid. In particular, it is intended to facilitate faster and better decisions. These may be decisions by AI researchers to focus on certain areas or decisions by companies and investors to develop and market certain AI products and AI services or the outlined decisions by politicians and regulators to limit the potential of damage by certain AI solutions and to protect our assets and social values. All together, we see the AI-MC² Grid as a powerful method and tool to define and manage all kinds of upcoming AI solutions and therefore as a substantial foundation to create norms and standards of AI applications.

Author Contributions This work is the result of a collaborative effort and incorporates ideas and work from all four authors. AI methods were classified and described by KS, TS and TH. Capabilities were classified and described by TS, KS and WH. The concept of AI criticality was described by TH, WH and TS. The examples were chosen and described by KS, TH, WH and TS. Graphics were prepared by KS and TH. Tables were prepared by KS, TH, TS and WH. References were identified and integrated by TS, KS, TH and WH.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Datenethikkommission der Bundesregierung (2019) Gutachten. Berlin, Oktober 2019
2. EU Commission (2020) White paper on artificial intelligence—a European approach to excellence and trust. EU Commission, Brussels
3. European Commission (2021) Proposal for a Regulation of the European Parliament and of the Council—laying down harmonised rules on Artificial Intelligence. European Commission, Brussels
4. Deutsches Institut für Normung (2020) German Standardization Roadmap on Artificial Intelligence. Berlin, November 2020.
5. IDC (2020) IDC forecasts strong 12.3% growth for AI market in 2020 amidst challenging circumstances. Online article <https://www.idc.com/getdoc.jsp?containerId=prUS46757920>. Accessed 30 June 2021
6. Tractica (2020) Artificial intelligence software market to reach \$126.0 billion in annual worldwide revenue by 2025. Online article <https://tractica.omdia.com/newsroom/press-releases/artificial-intelligence-software-market-to-reach-126-0-billion-in-annual-worldwide-revenue-by-2025/>. Accessed 30 June 2021
7. Strigl D, Kofler K, Podlipnig, S (2010) Performance and scalability of GPU-based convolutional neural networks. In: 2010 18th Euromicro Conference on Parallel, Distributed and Network-based Processing, pp 317–324
8. Nurvitadhi E, Sim J, Sheffield D, Mishra A, Krishnan S, Marr D (2016) Accelerating recurrent neural networks in analytics servers: comparison of FPGA, CPU, GPU, and ASIC. In: 2016 26th International Conference on Field Programmable Logic and Applications (FPL), pp 1–4
9. Marr B (2019) Artificial intelligence in practice: how 50 successful companies used AI and machine learning to solve problems. Wiley
10. Simon JP (2019) Artificial intelligence: scope, players, markets and geography. Digit Policy Regul Gov 21(3):208–237
11. Kaplan J (1984) The industrialization of artificial intelligence: from by-line to bottom line. AI Mag 5(2):51–57
12. Aastha J, Shah D, Churi P (2019) A review on business intelligence systems using artificial intelligence. In: Smys S, Tavares J, Balas V, Iliyasu A (eds) International Conference On Computational Vision and Bio Inspired Computing. Springer, Cham, pp 1023–1030
13. Yanqing D, Edwards JS, Dwivedi YK (2019) Artificial intelligence for decision making in the era of Big Data-evolution, challenges and research agenda. Int J Inf Manag 48:63–71
14. Schrottenboer DW (2019) The impact of artificial intelligence along the customer journey: a systematic literature review. Bachelor thesis. University of Twente
15. Huang M-H, Rust RT (2018) Artificial intelligence in service. J Serv Res 21(2):155–172
16. Commission EU (2018) A definition of AI: main capabilities and scientific disciplines. Bericht der High-Level Expert Group on Artificial Intelligence, Brussels
17. Russell SJ, Norvig P (2014) Artificial intelligence: a modern approach, 3rd edn. Pearson, Harlow
18. Goertzel B (2012) Perception processing for general intelligence: bridging the symbolic/subsymbolic gap. In: International Conference on Artificial General Intelligence. Springer, Berlin, Heidelberg, pp 79–88
19. Hammer B, Hitzler P (2007) *Perspectives of neural-symbolic integration*. Springer, Berlin
20. Chowdhary KR (2020) Fundamentals of artificial intelligence. Springer India
21. Jackson PC (2019) Introduction to artificial intelligence. Courier Dover Publications
22. Russell S (2019) Human compatible: artificial intelligence and the problem of control. Penguin
23. Corea F (2019) AI knowledge map: how to classify AI technologies. An introduction to data. Springer, Cham
24. Perrault R et al (2019) The AI Index 2019 Annual Report. AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, California
25. Durfee EH, Rosenschein JS (1994) Distributed problem solving and multi-agent systems: Comparisons and examples. In: *Proceedings of the Thirteenth International Distributed Artificial Intelligence Workshop*. pp. 94–104
26. Hong L, Page SE (2001) Problem solving by heterogeneous agents. J Econ Theory 97(1):123–163
27. Karaboga D, Akay B (2009) A comparative study of artificial bee colony algorithm. Appl Math Comput 214(1):108–132
28. Kautz HA, Allen JF (1986) Generalized plan recognition. AAAI 86:32–37
29. Horvitz EJ, Breese JS, Henrion M (1988) Decision theory in expert systems and artificial intelligence. Int J Approx Reason 2(3):247–302
30. Rebstock M (2008) Ontology-based business integration, 1st edn. Springer-Verlag, Berlin
31. Maniraj V, Sivakumar R (2010) Ontology languages—a review. Int J Comput Theory Eng 2(6):887–891
32. Prasad MR, Biere A, Gupta A (2005) A survey of recent advances in SAT-based formal verification. Int J Softw Tools Technol Transf 7(2):156–173
33. Besold TR et al (2017) Reasoning in non-probabilistic uncertainty: Logic programming and neural-symbolic computing as examples. Minds Mach 27(1):37–77
34. Luger GF (2008) Artificial intelligence: structures and strategies for complex problem solving, 5th edn. Alpha Books
35. Doyle, J (1977) Truth maintenance systems for problem solving. Dissertation, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts
36. Gros Dut XE et al (1997) Dempster-Shafer theory. NDT Data Fusion
37. Dutta S et al (1993) Integrating case- and rule based reasoning. Elsevier, INSEAD, Fontainebleau
38. Lewis DD (1998) Naive (Bayes) at forty: the independence assumption in information retrieval. In: European conference on machine learning. Springer, Berlin, Heidelberg, pp 4–15
39. Eddy SR (2004) What is a hidden Markov model? Nat Biotechnol 22(10):1315–1316
40. Peng C-YJ et al (2002) An introduction to logistic regression analysis and reporting. J Educ Res 96(1):3–14
41. Lim T-S et al (2000) A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. Mach Learn 40(3):203–228
42. Krogh A (2008) What are artificial neural networks? Nat Biotechnol 26(2):195–197
43. Hornik K et al (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2(5):359–366
44. Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, pp 93–124 (**Chap. 6**)

45. Jain AK (2010) Data clustering: 50 years beyond k-means. *Pattern Recognit Lett* 31(8):651–666
46. Murtagh F (1983) A survey of recent advances in hierarchical clustering algorithms. *Comput J* 26(4):354–359
47. Kohonen T (2001) Self-organizing maps, 3rd edn. Springer Series in Information Sciences, Berlin, p 30
48. Carpenter G et al (1988) The ART of adaptive pattern recognition by a self-organizing neural network. *Computer* 21(3):77–88
49. Carreira-Perpinan, MA, Lu Z (2008) Dimensionality reduction by unsupervised regression. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition, pp 1–8
50. Kramer O (2011) Dimensionality reduction by unsupervised k-nearest neighbor regression. In: 2011 10th international conference on machine learning and applications and workshops, pp 275–278
51. Lee J et al (2010) Unsupervised dimensionality reduction: overview and recent advances. In: IEEE International Joint Conference on neural networks (IJCNN), pp 1–8
52. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507
53. Bengio Y et al (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 35(8):1798–1828
54. Seeger M (2006) A Taxonomy for semi-supervised learning methods. In: Chapelle O, Scholkopf B, Zien A (Eds.) *Semi-supervised learning*. MIT Press, Cambridge, pp 17–32
55. Bair E, Tibshirani R (2004) Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2(4):e108
56. Goldberg AB, Zhu X (2006) Seeing stars when there aren't many stars: graph-based semi-supervised learning for sentiment categorization. In: *Proceedings of the First Workshop on graph based methods for natural language processing*, pp 45–52
57. Fergus R et al (2009) Semi-supervised learning in gigantic image collections. In: *Proceedings of the Twenty-Third Annual Conference on Neural Information Processing Systems (NeurIPS)*, pp 522–530
58. Zhu X, Goldberg AB (2009) Overview of semi-supervised learning. In: Brachman RJ, Dietterich T (eds) *Introduction to semi-supervised Learning*, Chap 2. Morgan & Claypool, pp 9–20
59. Zhu X et al (2007) Humans perform semi-supervised classification too. *Proc Natl Conf Artif Intell* 22:864–869
60. Sutton RS, Barto AG (2018) *Reinforcement learning*, 2nd edn. MIT Press
61. Kober J, Bagnell JA, Peters J (2013) Reinforcement learning in robotics: a survey. *Int J Robot Res* 32(11):1238–1274
62. Lewis F et al (2012) Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers. *IEEE Control Syst* 32(6):76–105
63. Leung FHF, Lam HK, Ling SH, Tam PKS (2003) Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Trans Neural Netw* 14(1):79–88
64. Bahrammirzaee A (2010) A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Comput Appl* 19(8):1165–1195
65. Corchado JM, Aiken J (2002) Hybrid artificial intelligence methods in oceanographic forecast models. *IEEE Trans Syst Ma Cybern Part C (Applications and Reviews)* 32(4):307–313
66. Hadavandi E et al (2012) Developing a hybrid artificial intelligence model for outpatient visits forecasting in hospitals. *Appl Soft Comput* 12(2):700–711
67. Liu Z, Zhu D, Lee KT, Kim AS, Raju L, Cai W (2020) Compound-ing meta-atoms into metamolecules with hybrid artificial intelligence techniques. *Adv Mater* 32(6):1904790
68. Pławiak P, Maziarsz W (2014) Classification of tea specimens using novel hybrid artificial intelligence methods. *Sens Actuators B Chem* 192:117–125
69. Martin A, Hinkelmann K, Gerber A, Lenat D, van Harmelen F, Clark P (2019) Preface. In: Martin A, Hinkelmann K, Gerber A, Lenat D, van Harmelen F, Clark P (Eds.) *Proceedings of the AAAI 2019 Spring Symposium on combining machine learning with knowledge engineering*
70. McGarry K, Wermter S, MacIntyre J (1999) Hybrid neural systems: from simple coupling to fully integrated neural networks. *Neural Comput Surv* 2(1):62–93
71. Komendantskaya E (2007) Learning and deduction in neural networks and logic. PhD thesis, Department of Mathematics, University College Cork, Ireland
72. Muggleton S, De Raedt L (1994) Inductive logic programming: theory and methods. *J Logic Program* 19:629–679
73. Schiller D, Huber T, Lingenfelser F, Dietz M, Seiderer A, André E (2019) Relevance-based feature masking: improving neural network based whale classification through explainable artificial intelligence. *Proc Interspeech* 2019:2423–2427
74. Shashank S, Labutov I, Mitchell T (2019) Learning to ask for conversational machine learning. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. pp. 4164–4174
75. Anderson LW, Krathwohl DR (2001) A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives. Longman, New York
76. Krathwohl DR (2002) A revision of Bloom's taxonomy. *Theory Pract* 41(4):212–218
77. Bloom BS (2016) *Taxonomy of educational objectives, Vol. 1: cognitive domain*. McKay, New York
78. Gagné RM (1972) Domains of learning. *Interchange* 3(1):1–8
79. Hartley J (1974) Programmed instruction 1954–1974—a review. *Innov Educ Train Int* 11(6):278–291
80. Krathwohl DR (2002) A revision of Bloom's taxonomy: an overview. *Theory Pract* 41(4):212–218
81. Krathwohl DR et al (1964) *Taxonomy of educational objectives, handbook II. Affective Domain*, David McKay Co
82. Reich K (1996) Systemisch-konstruktivistische Didaktik. Eine allgemeine Zielbestimmung. In: *Die Schule neu erfinden* 3, pp. 70–91
83. Rustemeyer D (2013) Konstruktivismus in der Erziehungswissenschaft. In: *Stichwort: Zeitschrift für Erziehungswissenschaft* 2, pp. 125–144
84. Davidson D (1980) Essay III. In: Davidson D (ed) *Essays on actions and events*. Oxford University Press, Oxford
85. Simpson EJ (1967) The classification of educational objectives, psychomotor domain. Technical Report, University of Illinois, Urbana
86. Johansen TK (1997) *Aristotle on the sense-organs*. Cambridge University Press
87. Fraden J (2010) *Handbook of modern sensors. Physics, design and applications*, 4th edn. Springer, Berlin
88. McGann M (2010) Perceptual modalities: modes of presentation or modes of interaction? *J Consc Stud* 17(1–2):72–94
89. Boring EG (1928) A new system for the classification of odors. *Am J Psychol* 40(2):345–349
90. Zwaardemaker H (1895) *Die Physiologie des Geruch*. Engelmann, Leipzig
91. Ritchie JB, Carruthers P (2015) The bodily senses. In: Matthen M (Ed) *The Oxford handbook of the philosophy of perception*, pp 353–370
92. Biggs S, Matthen M, Stokes D (2015) Sorting the senses. In: Stokes D, Matthen M, Biggs S (eds) *Perception and its modalities*. Oxford University Press

93. Macpherson F (2011) Individuating the senses. In: Macpherson F (ed) *The senses: classic and contemporary philosophical readings*. Oxford University Press
94. Sockett H (1971) Bloom's taxonomy: a philosophical critique (I). *Camb J Educ* 1(1):16–25
95. Pring R (1971) Bloom's taxonomy: a philosophical critique (2). *Camb J Educ* 1(2):83–91
96. Furst EJ (1981) Bloom's taxonomy of educational objectives for the cognitive domain: philosophical and educational issues. *Rev Educ Res* 51(4):441–453
97. Krathwohl DR et al (1964) *Taxonomy of educational objectives, handbook II. Affective Domain*, David McKay C
98. Craig T (1999) Communication theory as a field. *Commun Theory* 9:119–61
99. Dance Frank EX (1970) The 'Concept' of communication. *J Commun* 20:201–210
100. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379–423
101. Luhmann N (1992) What is communication? *Commun Theory* 2(3):251–259
102. Van Ruler B (2018) Communication theory: an underrated pillar on which strategic communication rests. *Int J Strateg Commun* 12(4):367–381
103. Prøitz TS (2010) Learning outcomes: what are they? Who defines them? When and where are they defined? *Educ Assess Eval Account* 22(2):119–137
104. Méhaut P, Winch C (2012) The European qualification framework: Skills, competences or knowledge? *Eur Educ Res J* 11(3):369–381
105. Wahlster W (2017) Künstliche Intelligenz als Grundlage autonomer Systeme. *Inf-Spekt* 40(5):409–418
106. Zweig KA (2019) *Algorithmische Entscheidungen: Transparenz und Kontrolle. Analysen und Argumente* 338, Konrad-Adenauer-Stiftung, Bonn
107. Taddeo M, Floridi L (2018) Regulate artificial intelligence to avert cyber arms race. *Nature* 556:296–298
108. Wali SB et al (2015) Comparative survey on traffic sign detection and recognition: a review. *Przeg Elektrotech* 1(12):40–44
109. Stallkamp J et al (2011) The German traffic sign recognition benchmark: a multi-class classification competition. In: *Proceedings of the 2011 International Joint Conference on neural networks*, pp 1453–1460