

Perspective

Essential Features in a Theory of Context for Enabling Artificial General Intelligence

Mayank Kejriwal 

Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Ste. 1001, Marina del Rey, CA 90292, USA; kejriwal@isi.edu

Abstract: Despite recent Artificial Intelligence (AI) advances in narrow task areas such as face recognition and natural language processing, the emergence of general machine intelligence continues to be elusive. Such an AI must overcome several challenges, one of which is the ability to be aware of, and appropriately handle, *context*. In this article, we argue that context needs to be rigorously treated as a first-class citizen in AI research and discourse for achieving true general machine intelligence. Unfortunately, context is only loosely defined, if at all, within AI research. This article aims to synthesize the myriad pragmatic ways in which context has been used, or implicitly assumed, as a core concept in multiple AI sub-areas, such as representation learning and commonsense reasoning. While not all definitions are equivalent, we systematically identify a set of seven features associated with context in these sub-areas. We argue that such features are necessary for a sufficiently rich *theory of context*, as applicable to practical domains and applications in AI.

Keywords: context; artificial general intelligence; context-rich AI; knowledge graphs; commonsense reasoning; semantic web; explainable AI; representation learning



Citation: Kejriwal, M. Essential Features in a Theory of Context for Enabling Artificial General Intelligence. *Appl. Sci.* **2021**, *11*, 11991. <https://doi.org/10.3390/app112411991>

Academic Editor: Luis Javier Garcia Villalba

Received: 2 November 2021

Accepted: 14 December 2021

Published: 16 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial Intelligence (AI) has made enormous strides in the last several years both due to the advent of technologies such as deep learning [1], and the cost-effectiveness and pervasiveness of supporting infrastructures, such as the evolution of GPUs [2], cloud computing [3], and open datasets and software [4]. In a number of well-defined and significant tasks, including face recognition [5], machine translation [6], and question answering [7], achieving or even crossing human-level performance is considered inevitable by mainstream researchers. It has even been argued that due to the success of AI in these individual applications, the achievement of Artificial General Intelligence (AGI) may be closer than expected. The challenges in truly accomplishing AGI are formidable, and has always been recognized by the experts. For example, in interviews conducted in 2018 with 23 prominent researchers and stakeholders working in AI today (including Google AI chief Jeff Dean and Stanford AI Director Fei-Fei Li), Martin Ford reports in their book that, among the 18 who answered, the average estimate was that AGI was still at least 80 years away [8]. A primary reason cited was that we were still missing the fundamental breakthroughs needed for machines to achieve a general and powerful model of intelligence [9].

Before AGI or any intelligent system that interfaces with humans in complex problem settings can be developed and deployed, the system must be capable of understanding *context*. Imbuing AI with context is an important milestone in realizing the broader goal of building adaptive architectures that can interact with humans naturally and seamlessly. Unfortunately, while many 'dictionary' (and in a few cases, academic) definitions of context exist (as further detailed in Section 2), and there is commonsense understanding of the word, there is lack of clarity between how such definitions map to current research and practice in the AI community.

With this motivation in mind, we argue that context needs to be conceptualized as a *novel class of information* in intelligent systems. Only through a direct discussion of context

can we start to distinguish between the different kinds of context, and disambiguate the varying, conceptually overloaded ways in which context is used as a term in professional practice. To this end, we discuss how context is understood and applied in various influential sub-fields in AI, including representation learning, Semantic Web, and explainable AI (Section 3). Based on this discussion, we distill seven key domain- and application-independent features that we argue to be important aspects of any theory of context that is sufficiently general (Section 4).

Context-rich AI is starting to witness a growing body of attention in industry, government and academia. Nevertheless, context is still not recognized as a fundamental kind of information. Most discussions on context, including the use of the word in AI publications and books, has been informal or even implicit. We argue that, because there is increasing interest from the community in context-rich AI, this is a productive time for researchers to be making progress on the theory and practice of context (Section 5). The article concludes in Section 6.

This article is *not* an attempt to present an actual theory of context. Indeed, it is not evident that a single theory of context can even be applied to all AI systems. A preliminary and informal version of some of the arguments in this article was also published earlier by the author in a blog post: <https://mkejriwal1.medium.com/artificial-intelligence-needs-to-be-more-context-aware-ecc3097ee2ea> (accessed on 1 November 2021). Rather, our more modest goal is to *systematically* discuss context, as a first-class citizen and novel class of information, in AI systems and applications that are expected to interface and interact with humans in complex problem domains, and that continue to be active subjects of research.

2. Background on Context: Definitions and Example Usage

While AI research has now been maturing for almost 60 years since the famous Dartmouth conference [10], interestingly, task-specific development of AI (or what is now popularly referred to as Artificial Narrow Intelligence (ANI), in contrast with AGI) was not the most compelling research area at the time. Instead, the goal was to build machines that exhibited true intelligence, and had a degree of generalizability that could potentially make them indistinguishable from humans. There was an optimistic belief that such breakthroughs were on the horizon, but an analysis of AI history shows that the field has tended to go through ebbs and flows, with several AI winters having come and gone over the decades, many involving neural networks and their promise [11]. To make meaningful progress, and advance the research agenda, the last three decades have witnessed a particularly increased development in building algorithms and systems that could do well in specific tasks including face recognition [12], machine translation [13], sentiment analysis [14] and information extraction [15], to only name a few.

However, now that excellent progress has been achieved on some of those tasks, researchers are again starting to come around to the development of more ‘general’ and less ‘task-centric’ AI, particularly in the NLP and computer vision communities. In Section 3, we describe research in specific sub-areas in AI to bolster the claim that context is an important part of this discussion. However, despite this interest, a standard definition of context has proven to be surprisingly complex. Oxford Languages defines it as the “circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed.”

More academic treatments, such as in areas within the social sciences and humanities (e.g., anthropology and sociology), generally understand context to refer to objects or entities surrounding a “focal event” e.g., that it is a “frame that surrounds the event and provides resources for its appropriate interpretation” [16]. Table 1 provides a sample of definitions and studies from both common and academic sources. We are not aware of context ever having been defined in an AI paper, although it is in common use today in many applications and sub-communities (as subsequently shown).

Table 1. Example definitions and studies of context from six independent sources. We quote from the sources directly in the table, except where indicated in italics.

Definition / Study	Source
The circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed.	Oxford Languages
An early survey of context in AI in the late 1990s, including context in knowledge acquisition, context in communication, and context in ontologies, to only name a few.	[17]
Any information that can be used to characterize the situation of an entity (whereby) an entity is a user, a place, or a physical or computational object that is considered relevant to the interaction between a user and an application, including the user and application themselves.	[18]
(i) The parts of a discourse that surround a word or passage and can throw light on its meaning; (ii) The interrelated conditions in which something exists or occurs.	Merriam-Webster
(i) Applying context to a situation, task, or system state provides meaning and advances understanding that can affect future decisions or actions; (ii) Integration of context-driven AI is important for future robotic capabilities to support the development of situation awareness, calibrate appropriate trust, and improve team performance in collaborative human–robot teams.	[19]
<i>Domain-specific usages include: “a probability distribution over the concepts in an environment [20], a set of relationships between objects [21], logical statements that represent cause and effect [22], or a function to select relevant features for object recognition [23].”</i>	Quote is from [19], with cited sources including [20–23]

No matter the definition adopted, we could well argue that, at an intuitive level, these definitions share many common threads, including the relativistic nature of context. Beyond a definition, it is important to explicitly address the question: why is context important for AI? We believe that the most effective way to answer this question is to present some representative examples below where context is shown to be vital for human thinking, reasoning and navigation of everyday situations. For an AI to successfully interact with humans, or replicate (or even supersede) these abilities, it must be both aware of context, and be able to handle it appropriately, which may or may not mirror human cognition. We designate such an AI as a ‘context-rich AI’. Subsequently, in Section 3, we study various sub-areas of current AI research to understand how context is drawn upon in such areas, directly or indirectly. Based on this study, Section 4 enumerates and makes a case for essential features that a theory of context must incorporate and address, as applicable to AI research.

2.1. Example Usage: Context in Conversation

Philosophers of language, and cognitive scientists, are familiar with the conversational maxims proposed by [24] as a model for why (and how) most people are able to make sense of conversation even when so much is left unsaid. Conversations rarely happen in a vacuum, but are situated in a context. The most immediate context might be the previous ‘sentence’ that was spoken, but it is only a small part of the conversational context. The people who are conversing, their roles in the conversation (e.g., an individual talking to another as a friend versus as a colleague), their individual goals in the conversation and the circumstances of the conversation (e.g., planned versus serendipitous) as well as the history of past conversations and shared assumptions, all constitute the context around a particular conversation, although some aspects may dominate over others. One kind of AI that attempts to maintain and remember such long histories is a never-ending learning system such as NELL [25], but it has not been shown to be proficient at navigating in, or designed for, such personalized contexts. Chatbots that are currently being developed in

industry would be the greatest beneficiaries of context-rich AI research that is able to take Grice's maxims and conversational contexts more fully into account [26,27].

2.2. Example Usage: Context as Background Knowledge

Background knowledge helps us navigate everyday life in more ways than we consciously realize [28]. The example above also relied on background knowledge as a contextual aspect of conversation, at least to an extent. However, beyond conversation, background knowledge plays an important role in more abstract tasks such as planning [29]. For example, when setting our calendars, we choose not to make business calls on Sunday because we know (or suspect) that the business might be closed on that day. Similarly, we plan ahead for summer vacations because experience has taught us that the summer is a heavy-demand period for vacations, and that we will not get good deals (or any deals) if we book closer to the anticipated vacation.

The influence of personal experience on background knowledge is an interesting one that is not completely settled, and is not deeply relevant to the topic at hand. There is ample evidence that some background knowledge is coded at the level of instincts and may have genetic or evolutionary origins [30] (e.g., fear of snakes, but not electricity). Others are learned from experiences, while yet others are a combination of nature and nurture [31]. This is generally true for many aspects of human cognition and is unlikely to be settled anytime soon. For AI research, the question is more pragmatic. Namely, for an AI to have a relatively complete 'model' of background knowledge and use it in a context-rich fashion, should it learn the model from large quantities of data, as language representation learning models, such as GPT-3 and BERT [32], in the NLP community, try to do, or should it instead apply powerful reasoning techniques on a core set of principles, common entities and relations? Proponents of deep learning believe in the former, but more recently, some researchers have started to explore the benefits of combining or encoding top-down AI principles in neural architectures [33]. In the foreseeable future, such hybrid approaches, properly founded on a rigorous theory of context, may yield many insights that could lead in smaller, more powerful models that are more explainable (see Section 3.4) and context-rich, but without having billions of parameters, such as GPT-3 [34].

2.3. Example Usage: Expression and Behavior Adaptation in Social and Emotional Contexts

Human beings are able to adapt their behavior in varying social and emotional contexts [35]. For example, an ordinary human being would not crack a joke at a funeral, and would be more conservative with humor in a work party than in a party with close friends. While humor is just one aspect of outward social interaction [36], many other examples can be construed that are similar, including formality of conversation, use of sarcasm [37], emotional intensity, choice of verbiage (e.g., explaining a concept to a layperson versus a scientific audience of experts) and even modality of expression (e.g., text message versus email) [38]. Currently, even AIs that are trained to produce conversation or text (including chatbots, and automated email and sentence completion) only work properly in a very specific context, and are not able to adapt seamlessly with changing social and emotional context. An AI guided by an appropriate theory of context may be more human-like in its assessment of, and adaptation to, complex real-world settings.

3. Research in Context-Rich Artificial Intelligence

While context-rich AI continues to be actively researched in the AI community, it is already beginning to achieve a degree of maturity in several sub-communities [19]. In this section, we describe four representative areas of AI research that significantly rely on context, as defined in Section 2, to achieve their goals, some more directly than others. While we do not claim that these are the only areas pertinent to the research of context-rich AI, we use them as our primary basis for enumerating the requirements that a potential theory of context must fulfill to be generalizable to several AI areas that rely on context.

3.1. Representation Learning

Representation learning is intimately connected with the growth and success of deep learning [39,40]. In the early days of machine learning, and even as recently as a decade ago, the primary machine learning workflow involved ‘feature engineering’ as a core step [41]. The importance of feature engineering was so well recognized that it was, at times, difficult to gauge whether an improvement in state-of-the-art performance on some task, such as sentiment analysis [42], was due to an algorithmic or model innovation, or due to better feature engineering [43].

Today, representation learning has largely, although not completely, eliminated the need to painstakingly engineer and evaluate sets of features. The core idea behind learning a representation of a “data item” (whether it is an image, a document, or even a character in a long string of text) is to “embed” it in a real-valued, low-dimensional vector space using a neural architecture. The argument is that such “embeddings” capture key properties of the data using a small set of real numbers, without requiring manual feature engineering [40]. Beyond improvements in performance that have been achieved using these methods, an impressive early finding, especially in the NLP community [44], was that such methods could capture analogies and other abstract relations that are evident in human cognition, e.g., the word2vec model [45] could automatically complete analogies such as *King is to man as Queen is to woman*, despite never having been trained explicitly to do well on such tasks.

We note that the embedding of data in low-dimensional spaces is itself not novel: previously, topic models such as Latent Dirichlet Allocation (for text) [46], and other classic dimension-reduction methods (for data that was represented more abstractly or was otherwise vectorized already) had already achieved some breakthrough innovations in this area. Today, however, representation learning is almost exclusively associated with neural representation learning. The topic receives wide coverage in the top machine learning conferences, including NeurIPS (<https://nips.cc/>, accessed on 1 November 2021) and the International Conference on Learning Representations (ICLR) (<https://iclr.cc/>, accessed on 1 November 2021).

In practice, representation learning is a broad area of research, and different communities approach it in different ways. In the computer vision community, the neural network itself learns the representations (given just raw inputs of pixels) that is task-optimal. The last layer of a convolutional network trained in a supervised setting, for example, could be used for getting learned representations of an image, including an unlabeled image [47].

In the NLP community, the earliest representation learning algorithms following the advent of neural networks have assumed a local definition of context. For example, models based on word2vec (skip-gram and continuous-bag-of-words) typically slide a “window” of a pre-specified size over the text, and use the words in the window flanking the target word as co-context [48]. Similarly, the GloVe model uses a matrix of co-occurring words [49]. In the data mining community, researchers have shown that the premise behind word2vec can even be applied to graphs, and have used it to learn representations for the nodes in a graph, such as a social network [50,51], or even a geospatial network [52]. When this was first proposed in the mid-2010s, state-of-the-art performance was achieved for important data mining tasks such as link prediction, using considerably simpler systems than had been previously proposed.

Gradually, the definition of context has broadened, along with (not coincidentally) the advent of more powerful and novel neural network models such as transformers, which have led to new-age embedding models such as BERT, RoBERTa and, more recently, GPT-3 [32,34]. Among other capabilities, such models have the capability of retrieving different embeddings of the same word based on the sentence in which the word is situated. Context has, therefore, started to play a more direct role, even after the training is finished. Transformers are inherently (and explicitly) more context-friendly than some of the other models that came before them. Research on transformers continues at a frenetic pace, and more recently, they have even been applied in computer vision, with promising results [53]. Locality, but also selective activation of data items or elements that are salient

to the target for which an embedding is being learned (sometimes achieved through “attention” mechanisms, such as in [54]), are both important features in the definition of context employed in the representation learning literature. In Section 4, we discuss these features further.

3.2. Commonsense Reasoning and Knowledge

Commonsense reasoning involves processing information about a scenario in the world, and making inferences and decisions by using not only the ‘explicit’ information available to the senses and conscious mind, but also context and implicit information that is based on our “commonsense knowledge” [55,56]. Commonsense knowledge is difficult to define precisely but it may be assumed to be a broad body of knowledge of how the world works. While interesting also to cognitive scientists and philosophers, commonsense reasoning is intriguing to specialists in AI because it has also revealed potential pitfalls in the manner in which such reasoning systems are traditionally evaluated, i.e., using one or more multiple choice question answering or QA benchmarks [57]. Two assumptions have come under attack in the AI community in the last two years: first, that multiple choice questions are adequate for assessing commonsense reasoning capabilities, and second, that simple ‘one-off’ question answering itself is enough to evaluate commonsense reasoning, as opposed to other tasks such as storytelling and abstractive summarization [58] that are more context-rich, and also seem to rely heavily on commonsense knowledge in the real world.

To address some of these concerns, generative QA tasks have already been proposed and are slated to grow in popularity [59]. A critical issue that the community is looking at is a reliable mode of evaluating system performance on such tasks. Because generative QA can result in open-ended answers, more innovation is needed on how to evaluate the “goodness” of a system’s answers without always relying on human annotation that is available on-demand, such as in a crowdsourcing framework.

True commonsense reasoning in machines requires a theory of context, especially for enabling some of the example usages we described in Section 2. Recent work has attempted to build a theory of commonsense [60], which could potentially be used for deriving a theory of context. However, recent work has shown that the theory may have to be supplemented with additional terms, since it does not always definitively map to human annotations [61]. Because commonsense includes both intuitive physics and intuitive psychology, advances in machine commonsense will likely have to draw on a broad and holistic definition of context, including understanding how events relate to one another, relational dependencies between objects and events, and appropriate framing of goals and assumptions given a task as input. We discuss some these features in more detail in Section 4.

3.3. Knowledge Graphs and Semantic Web

Graphs have been ubiquitous in AI since its founding, but have tended to be more associated with the planning community [62,63]. With the publishing of Google’s influential post on “things, not strings” in 2011 [64], knowledge graphs (KGs) have come to be recognized as a flexible and data-driven way of modeling a domain using entities, relations, and events, rather than as collections of primitives (e.g., strings and numbers) that are lacking in semantics. Figure 1 above illustrates a simple domain-specific knowledge graph over the Books domain. Companies such as Google and Microsoft have taken this concept much further [65], and the Google Knowledge Graph is known to be a web-scale graph constructed in a data-driven fashion from numerous sources and over broad domains. The Google Knowledge Graph underlies at least some of the search artifacts (such as a knowledge panel) that are produced in response to queries such as “Leonardo da Vinci”. Knowledge graphs have led to enormous advanced in both generic and domain-specific (e.g., e-commerce) search [66], and are being constructed in some form in almost all the big technology companies. COVID-19 knowledge graphs were even constructed in the earliest

days of the pandemic [67]. Typically, KGs are constructed over text and Web data, though more recently, multi-modal and domain-specific KG construction has started to become popular [68,69].

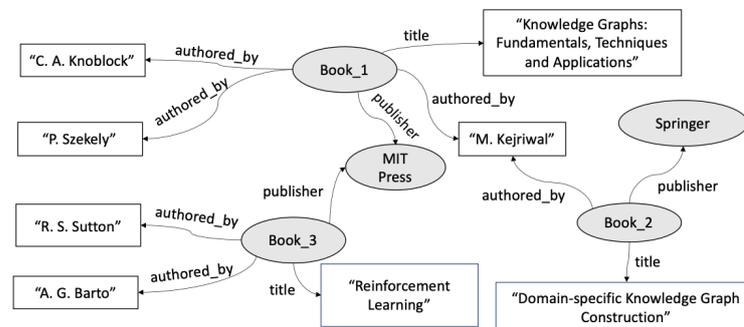


Figure 1. An example knowledge graph (KG) fragment from the Books domain.

KG research has a closer connection to context-rich AI than may initially meet the eye. First, KGs are (typically) populated according to a domain model, called an ontology, that contains the concepts, constraints and relations that serve to define both the scope and vocabulary of a domain. It would not be incorrect to say that a KG exists in the context of a domain and a task (or set of tasks), with the former defined using frameworks such as the Web Ontology Language (OWL) [70]. For example, the ontology underlying the KG in the figure may contain concepts such as “book”, “person” and “publisher” (among others), and relations such as “author_of” and “publisher_of”. Concepts and relations may have sub-concepts and sub-relations; for example, an “author” concept that is a sub-class of ‘person’ could be declared in the ontology. These relations connect the entities, which are formally instances of concepts, in the KG. Constraints, which are not visible on the surface, prevent unwanted declarations in the KG, e.g., one reasonable constraint might be that the range of the “author_of” relation must be an instance of a “person” concept.

The Semantic Web community has conducted much research over the years on rigorous ways to formally express ontologies and knowledge graphs [71], and has even developed reasoning engines to infer more facts than are explicitly declared in the KG. More recently, KG research has intersected with both representation learning and commonsense reasoning [72]. Within the Semantic Web community, context is therefore often associated with correctly inferring knowledge that is implicit and possibly, only approximately correct. Recent advances in neuro-symbolic AI aim to combine the benefits of symbolic and logic-based AI, and machine learning methods such as deep neural networks [73].

3.4. Explainable AI

Explainable AI has witnessed a resurgence in the last 5–7 years [74,75], both due to interest from the community and well-funded programs such as the Explainable AI (XAI) program that was instituted by the United States Defense Advanced Research Projects Agency (DARPA) in 2016 [76]. According to the program’s website, its central goal is to ‘produce more explainable models, while maintaining a high level of learning performance.’ The definition of explainable AI is more straightforward and less controversial than commonsense reasoning, where agreement on a single definition and set of tasks has been lacking, although there has been convergence over the last two years.

While the direct connection between explainable AI, and some of the other research areas such as knowledge graphs, representation learning and commonsense reasoning, is not fully evident, the argument is that any AI system that exhibits the kind of robustness that would be expected of human intelligence should have the capacity to explain itself [77]. For example, an explanation-capability ought to be a defining feature of a commonsense reasoner that is asked to provide a justification for an answer or action. Current language models that have achieved state-of-the-art performance in commonsense QA tasks do not seem to have this capability (at least in their current state), although recent work has made

progress on this issue [78]. An advantage of an explanation-capability is that it makes the AI more trustworthy to non-developers and domain scientists who can probe the AI for the reasoning behind its decisions. Such a capability would be obviously critical in an HMI system deployed in the real world. The DARPA XAI program mentioned above also mentions as its the goal the development of a suite of machine learning techniques that ‘enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.’

At least in part, an explainable AI can generate convincing explanations by leveraging its task-context and background knowledge, and by activating the elements in its background knowledge that are relevant in the context of the task. For example, when answering a question such as *why will not an elephant fit through my kitchen door?* the salient element that must be activated in the context of the question is *size* since an elephant is too big to fit through the kitchen door. However, along with possessing such an activation capability, the AI must also have the requisite background knowledge about the *typical* sizes of kitchen doors and elephants.

4. Understanding Context in Practical AI Research

Based on the discussion in the previous section, we synthesize seven features that encompass a wide range of context-rich AI applications. These features are neither mutually exhaustive nor completely non-overlapping, but each has distinctive elements. At minimum, a theory of context must accommodate these seven features, although there may be domain-specific exceptions.

1. **Locality:** Locality is usually an important aspect of context, especially in “embedding” or representation learning algorithms. For example, in the word2vec architecture [45], and others inspired by it [48,49], only words within a certain window of the target word are ‘activated’ and considered as the context of that word. Similar notions apply in graph and network embeddings [50,52]. However, more recently, context has become less local due to the use of powerful features such as biased random walks when training graph embeddings [51,79,80].
2. **Selective Activation of Salient Elements:** Whether local (such as in the applications above) or non-local, context involves selective activation of salient elements. In the case of using random walks for embeddings, the nodes and edges in the walk may be considered to be the salient elements, even though they may not be considered local to the target node. In cognitive and agent-based architectures, certain kinds of long and short-term memory retrieval may be used to selectively activate salient elements [81].
3. **Relational Dependencies:** In the definitions in Table 1, we noted indications of relational dependencies between elements or objects when context is invoked. Specifically, depending on the application, context may be defined as (emphasis ours) “a probability *distribution over the concepts* in an environment [20], a set of *relationships* between objects [21], logical statements that represent *cause and effect* [22], or a function to select relevant features for object recognition [23].” Another kind of relational dependency may arise due to the application. On social media, for example, the context for training a tweet embedding may include metadata (such as the user posting the tweet) rather than just the content in the tweet [82,83]. This metadata expresses the relation between a user profile and between the actual tweet content, both of which can exist independently, but which need to be related to each other to learn good embeddings for either.
4. **Implicitness:** Especially in common sense reasoning and explainable AI research, certain pieces of information are considered implicit. Grice’s conversational maxims are good examples of statements that explain some of the implicitness in conversation [24]; however, other similar maxims, or a generalization of the original maxims, may be necessary to categorize and explain implicitness in domains such as computer vision that are non-linguistic.

5. **Open-World Environments:** As AI systems are implemented in organizations and customer-facing applications, they must increasingly adapt to the “open world”. An open world environment is one where either the structure or parameterization (or both) is unknown or unknowable [84]. An example of the latter is a chaotic system with a high degree of uncertainty around the initial conditions [85]. One can probabilistically reason about the outcomes given the laws governing the dynamics of the system, but the outcome at a given point of time may be largely unknown with any reasonable confidence. The real world, and complex systems in the real world, are good examples of open worlds that are (at best) partially unknown, at least in practice, and (at worst) may not be completely knowable, even in theory. Philosophers have long argued about whether complete and provable knowledge is even possible in such systems, due to uncertainty [86], vagueness [87], and the circularity of induction [88].
An important aspect of open worlds is that, due to their (partially) unknown parameterization and structure, unexpected situations and “novelties” might occur [89]. The COVID-19 pandemic is a good example of a global novelty that may have profound long-term consequences, long after the pandemic. If interactions between human or machine agents are occurring in an open world, an instantiated theory of context will likely have to rely on powerful representational techniques, such as open sets and infinite-state Markov processes [90,91]. We hypothesize that early theories of context will likely make the closed-world assumption, with a proper framing of tasks, goals and assumptions (as discussed below). Research on open worlds is still in its infancy in the AI community, although recent progress on open-world learning has been impressive [92]. We posit that a theory of context in open-world environments may be necessary for building a sufficiently powerful AGI.
6. **Event-Driven Triggers:** While real life, and aspects of human–human interaction often seems to proceed seamlessly and “naturally”, there are epistemic and behavioral transitions in most non-trivial interactions. In other cases, a specific event (such as a disagreeable statement) may trigger such a transition explicitly. When two friends are having a casual conversation, and one of the friends receives an urgent phone call from her child’s school, the event triggers stress and causes the other friend to elicit concern, sometimes non-verbally. In most real world and open world environments, unexpected events will occur with non-vanishing probability, as argued earlier. Such events could potentially alter the terms of an interaction, which must be accounted for by a robust theory of context [93,94].
7. **Framing of Tasks, Goals and Assumptions:** Contextualization often occurs in the presence or ‘frame’ of one or more tasks, goals and assumptions. In cognitive science, such framing is considered vital for human interaction [95,96]. Some or all of these may be implicit. When having a written dialog with a human individual, a machine is implicitly assuming that the person can read and understand the language in which the machine is outputting dialog. Moreover, in the context of a customer service dialog, the machine may assume that the customer has a specific need or problem, and the conversation is occurring with the goal of solving the problem. The task is then sequential and multi-step: to first understand the need of the customer, and to then devise a solution for it, without recourse to a human operator, if possible. Within a specific application, therefore, an instantiation of a ‘good’ theory of context must formalize and epistemically represent the tasks, goals and assumptions for all agents (human and machine) interacting in an environment. For example, if these epistemic states are considered continuous, rather than discrete, the theory may rely on Markov processes and differential equations, whereas workflow-like models [97,98] may suffice if there is only a small set of tasks, goals and assumptions.

We note that this list is intended to be suggestive, not exhaustive. There may very well be other features that a full theory of context could attempt to incorporate and organize. At least one utility of such a theory would be to provide an account (which could be

descriptive or quantitative) of whether one or more contextual assumptions can be used in the system design process, including in such steps as training-data acquisition, to improve key-performance indicators, such as efficiency and generalization. Another explanatory role would be to understand the relative importance of each feature during the design ('offline') phase versus real-time use ('online'). We also suspect that such a theory would have communicative utility, since it would help all stakeholders clearly demarcate (and document) the contexts in and around system usage using specific, consistent terminology defined within the confines of the theory.

Finally, the theory could be used to formulate and assess hypotheses, whether using computational tools, social science methodologies, arguments, or other modalities of hypothesis exploration, as deemed appropriate. For example, we hypothesize that, in an HMI setting involving a pre-trained intelligent architecture, designed for the average working-age adult, features (4) and (6) in the list above are more relevant than (3). We also suspect that feature (5) increases in importance the more the system interacts with someone. In contrast, features (1) and (2) may be particularly important in HMI settings where a physically embodied architecture (such as a robot) is involved. To rigorously test such hypotheses, clear guidelines and metrics may have to be proposed (which may themselves be field- and architecture-specific) to measure "intuitive"-sounding terms such as "importance".

5. Supporting Ecosystems and Social Factors

A range of supporting ecosystems, both in industry and through governmental bodies such as national funding agencies, are currently geared toward supporting some of the areas we covered in Section 3, and a few consider context as a central concept in the future development of AI. Before describing some of these programs, we note the importance of ecosystems and social factors for enabling research in context-rich AI, both theoretically and empirically. In an economy with heavy competition for resources, it is difficult to rally interest in a field unless there are incentives and supporting infrastructures in place. The fact that funding opportunities for developing more context-rich AI already exist (with some having been launched very recently) is an encouraging finding, suggesting that context-rich AI will get the attention it deserves. We focus primarily on supporting ecosystems in the United States, although similar findings apply both in China, where AI research has accelerated immensely over the last few years, and also the European Union.

In the United States, the National Science Foundation (NSF), has instituted multiple programs where context-rich AI is going to play a key role, although the development of such an AI is not the "direct" aim of the effort. One example is the relatively recent call for "AI Institutes", a joint government effort [99] between NSF, the U.S. Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA), the U.S. Department of Homeland Security (DHS) Science and Technology Directorate (S&T), and the U.S. Department of Transportation (DOT) Federal Highway Administration (FHWA). It also involves significant industry participation from the likes of Amazon and Google (among others). Simply put, these institutes are meant to facilitate "new and sustained research in AI to drive science and technology progress". Teams comprising one or more organizations can submit proposals to one of eight themes. One of these themes is "human-AI interaction and collaboration". In both Sections 2 and 3, we argued that, for seamless human-AI interaction and collaboration to occur, an AI must necessarily (although not sufficiently) be context-rich.

Some other themes also have relevance to context-rich AI, including *AI-Augmented Learning* [100]. While this is the most direct and well-funded example of context-rich AI enablement (with an anticipated funding of USD 128 to 160 million over eight institute selections), in recent years, other similar but smaller-scale multi-disciplinary AI programs have also been announced.

The relevance of context-rich AI to military applications and the United States Departments of Defense (DoD) cannot be underestimated. Multiple DARPA programs have

been instituted in the last few years that involve context, either directly or indirectly. Many of these programs are described as “third-wave AI” programs [101] that combine (and improve upon) the best of first-wave AI (logical, rule-based and/or expert systems) and second-wave AI (machine learning and other ‘inductive’ architectures). We mentioned the DARPA XAI program earlier [76], but another one is the DARPA Machine Common Sense (MCS) program [102] that directly funds research and development in commonsense AI, which must be context-rich and context-aware to achieve its full potential.

Yet another example is the DARPA Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON) program [89], which attempts to build AI systems that can ‘act appropriately and effectively in novel situations’ that occur in open worlds, of which the real world is an example. Context and novelty go hand-in-hand, since an event is typically novel with respect to its context. For example, if someone starts shouting in the middle of an examination, that would (generally) be considered more novel than if someone started shouting in a political rally or a protest. We argued in Section 4 that open-world environments are an essential feature that must be accounted for in viable theories of context.

However, despite the interest from national funding agencies and the Department of Defense, we believe that industry will be among the greatest beneficiaries of the coming-of-age of context-rich AI, including AI applications built on rigorous theories of context. Chatbots and autonomous interfaces (e.g., website “assistants”) are becoming more prevalent throughout modern business [26], and not just limited to the “Big Tech” enterprises. Similarly, knowledge graphs and Semantic Web technologies are starting to become increasingly popular in mainstream applications, including finance and medicine [67,103]. Industry-academic cooperation and funding opportunities have been increasing throughout AI, and industry researchers are fairly prominent in the key data mining and machine learning conferences. Another beneficiary of this research will be non-profit and government, especially in topical areas such as disaster relief and crisis management [104]. Building trustworthy AI is an especially important goal if the technology is to witness greater uptake in government agencies and the non-profit sector. We noted earlier the close connection between explainable AI and trustworthy AI, and the belief that context-rich AI will be necessary for achieving both of these goals. A complete and self-consistent theory of context may help us to better understand these issues, especially when it comes to building sophisticated AI systems that need to communicate and collaborate with humans in natural and intuitive ways.

6. Conclusions

In this article, we explored the different definitions of context, and argued that, despite the prevalence and maturity of context-rich AI research, context is still treated implicitly in much of this research. In addition to presenting multiple definitions of context, we distilled seven key features of context by studying how context is viewed in several sub-areas of AI. These features include simple constructs such as relational dependencies (dependencies between elements or objects when context is invoked), to higher-order environmental constructs such as open-world environments, of which either the structure or parameterization is unknown (or unknowable).

We argued that a theory explaining context in human–machine interactions in a structured, domain-independent manner must, at minimum, instantiate and incorporate these features in addition to other features that may be more domain-specific. We did not claim that these seven features were complete or even mutually exhaustive. Determining and quantifying these relationships more formally is a valuable area for future research into theories of context as understood in AI and Human–Machine Interaction (HMI) research.

Finally, we suggested that the current time is just right for investigating context in depth, due to a confluence of favorable factors, including sufficient maturity of underlying technologies, interest from various communities (especially those coming off of recent successes in deep learning and applied AI), and supporting ecosystems that see advanced

AI research as geo-politically necessary. We provided examples of currently funded efforts that are seeking to conduct such advanced research, and where context plays an important role. This confluence of factors is an important practical step toward realizing the vision of a more contextualized AI that can be deployed with ease in many real-world applications.

Without a full exploration of context as a first-class epistemological citizen, it will be difficult to realize such an AI, especially in complex applications where humans and AI must work, interact and make decisions together. A more complete, empirically falsifiable theory of context within AI research is a promising future direction for investigation.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
2. Wang, Y.E.; Wei, G.Y.; Brooks, D. Benchmarking tpu, gpu, and cpu platforms for deep learning. *arXiv* **2019**, arXiv:1907.10701.
3. Rimal, B.P.; Choi, E.; Lumb, I. A taxonomy and survey of cloud computing systems. In Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC, Seoul, Korea, 25–27 August 2009; pp. 44–51.
4. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]
5. Guo, G.; Zhang, N. A survey on deep learning based face recognition. *Comput. Vis. Image Underst.* **2019**, *189*, 102805. [CrossRef]
6. Singh, S.P.; Kumar, A.; Darbari, H.; Singh, L.; Rastogi, A.; Jain, S. Machine translation using deep learning: An overview. In Proceedings of the 2017 International Conference on Computer, Communications and Electronics (Comptelix), Jaipur, India, 1–2 July 2017; pp. 162–167.
7. Sharma, Y.; Gupta, S. Deep learning approaches for question answering system. *Procedia Comput. Sci.* **2018**, *132*, 785–794. [CrossRef]
8. Ford, M. *Architects of Intelligence: The Truth about AI from the People Building It*; Packt Publishing Ltd.: Birmingham, UK, 2018.
9. Müller, V.C.; Bostrom, N. Future progress in artificial intelligence: A survey of expert opinion. In *Fundamental Issues of Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 555–572.
10. Moor, J. The Dartmouth College artificial intelligence conference: The next fifty years. *Ai Mag.* **2006**, *27*, 87.
11. Mims, C. Why Artificial Intelligence Isn't Intelligent. 2021. Available online: <https://www.wsj.com/articles/why-artificial-intelligence-isnt-intelligent-11627704050> (accessed on 2 September 2021).
12. Parkhi, O.M.; Vedaldi, A.; Zisserman, A. Deep face recognition. In Proceedings of the BMVC 2015, Swansea, UK, 7–10 September 2015.
13. Dehghani, M.; Gouws, S.; Vinyals, O.; Uszkoreit, J.; Kaiser, L. Universal transformers. *arXiv* **2018**, arXiv:1807.03819.
14. Wang, Y.; Sun, A.; Han, J.; Liu, Y.; Zhu, X. Sentiment analysis by capsules. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 1165–1174.
15. Lockard, C.; Shiralkar, P.; Dong, X.L. Openceres: When open information extraction meets the semi-structured web. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, MN, USA, 3–5 June 2019; pp. 3047–3056.
16. Irvine, J.; Schieffelin, B.; Series, C.; Goodwin, M.H.; Kuipers, J.; Kulick, D.; Lucy, J.; Ochs, E. *Rethinking Context: Language as an Interactive Phenomenon*; Number 11; Cambridge University Press: Cambridge, UK, 1992.
17. Brézillon, P. Context in Artificial Intelligence: I. A survey of the literature. *Comput. Artif. Intell.* **1999**, *18*, 321–340.
18. Dey, A.K.; Abowd, G.D.; Salber, D. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Hum.-Comput. Interact.* **2001**, *16*, 97–166. [CrossRef]
19. Schaefer, K.E.; Oh, J.; Aksaray, D.; Barber, D. Integrating Context into Artificial Intelligence: Research from the Robotics Collaborative Technology Alliance. *AI Mag.* **2019**, *40*, 28–40. [CrossRef]
20. Singhal, A.; Luo, J.; Zhu, W. Probabilistic spatial context models for scene content understanding. In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA, 18–20 June 2003; Volume 1, p. I.
21. Rabinovich, A.; Vedaldi, A.; Galleguillos, C.; Wiewiora, E.; Belongie, S. Objects in context. In Proceedings of the 2007 IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
22. Zettlemoyer, L.S.; Collins, M. Learning Context-Dependent Mappings from Sentences to Logical Form. In Proceedings of the 47th Annual Meeting of the ACL, Suntec, Singapore, 2–7 August 2009.

23. Heitz, G.; Koller, D. Learning spatial context: Using stuff to find things. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 30–43.
24. Grice, H.P. Logic and conversation. In *Speech Acts*; Brill: Buckinghamshire, UK, 1975; pp. 41–58.
25. Mitchell, T.; Cohen, W.; Hruschka, E.; Talukdar, P.; Yang, B.; Betteridge, J.; Carlson, A.; Dalvi, B.; Gardner, M.; Kisiel, B.; et al. Never-ending learning. *Commun. ACM* **2018**, *61*, 103–115. [[CrossRef](#)]
26. Shawar, B.A.; Atwell, E. Chatbots: Are they really useful? *Ldv Forum* **2007**, *22*, 29–49.
27. Brandtzaeg, P.B.; Følstad, A. Chatbots: Changing user needs and motivations. *Interactions* **2018**, *25*, 38–43. [[CrossRef](#)]
28. Potter, J. Cognition as context (whose cognition?). *Res. Lang. Soc. Interact.* **1998**, *31*, 29–44. [[CrossRef](#)]
29. Alterman, R. Adaptive planning. *Cogn. Sci.* **1988**, *12*, 393–421. [[CrossRef](#)]
30. Öhman, A. Has evolution primed humans to “beware the beast”? *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16396–16397. [[CrossRef](#)]
31. Meaney, M.J. Nature, nurture, and the disunity of knowledge. *Ann. N. Y. Acad. Sci.* **2001**, *935*, 50–61. [[CrossRef](#)] [[PubMed](#)]
32. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
33. Sarker, M.K.; Zhou, L.; Eberhart, A.; Hitzler, P. Neuro-Symbolic Artificial Intelligence Current Trends. *arXiv* **2021**, arXiv:2105.05330.
34. Floridi, L.; Chiriatti, M. GPT-3: Its nature, scope, limits, and consequences. *Minds Mach.* **2020**, *30*, 681–694. [[CrossRef](#)]
35. Coulmas, F.; Rabin, C.; Ibrahim, M.H.; Massamba, D.P.B.; Daswani, C.J.; Pasierbsky, F.; Takada, M.; Sugito, S.; Porksen, U.; Ehlich, K.; et al. *Language Adaptation*; Cambridge University Press: Cambridge, UK, 1989.
36. Attardo, S. *The Routledge Handbook of Language and Humor*; Taylor & Francis: New York, NY, USA, 2017.
37. Shamay-Tsoory, S.G.; Tomer, R.; Aharon-Peretz, J. The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology* **2005**, *19*, 288. [[CrossRef](#)]
38. Tagg, C. *Discourse of Text Messaging: Analysis of SMS Communication*; Bloomsbury Publishing: London, UK, 2012.
39. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)]
40. Zhong, G.; Wang, L.N.; Ling, X.; Dong, J. An overview on data representation learning: From traditional feature learning to recent deep learning. *J. Financ. Data Sci.* **2016**, *2*, 265–278. [[CrossRef](#)]
41. Zheng, A.; Casari, A. *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*; O’Reilly Media, Inc.: Sebastopol, CA, USA, 2018.
42. Feldman, R. Techniques and applications for sentiment analysis. *Commun. ACM* **2013**, *56*, 82–89. [[CrossRef](#)]
43. Sharma, A.; Dey, S. A comparative study of feature selection and machine learning techniques for sentiment analysis. In Proceedings of the 2012 ACM Research in Applied Computation Symposium, San Antonio, TX, USA, 23–26 October 2012; pp. 1–7.
44. Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; Kuksa, P. Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **2011**, *12*, 2493–2537.
45. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed representations of words and phrases and their compositionality. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; pp. 3111–3119.
46. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
47. Hsu, C.C.; Lin, C.W. Cnn-based joint clustering and representation learning with feature drift compensation for large-scale image data. *IEEE Trans. Multimed.* **2017**, *20*, 421–429. [[CrossRef](#)]
48. Joulin, A.; Grave, E.; Bojanowski, P.; Mikolov, T. Bag of Tricks for Efficient Text Classification. *arXiv* **2016**, arXiv:1607.01759.
49. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
50. Perozzi, B.; Al-Rfou, R.; Skiena, S. Deepwalk: Online learning of social representations. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 701–710.
51. Grover, A.; Leskovec, J. node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 855–864.
52. Kejriwal, M.; Szekely, P. Neural embeddings for populated geonames locations. In *International Semantic Web Conference*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 139–146.
53. Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. *arXiv* **2021**, arXiv:2101.01169.
54. DeRose, J.F.; Wang, J.; Berger, M. Attention flows: Analyzing and comparing attention mechanisms in language models. *IEEE Trans. Vis. Comput. Graph.* **2020**, *27*, 1160–1170. [[CrossRef](#)]
55. Minsky, M.L. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*; Simon & Schuster: New York, NY, USA, 2006; Google-Books-ID: AT0UIQEACAAJ.
56. Davis, E.; Marcus, G. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Commun. ACM* **2015**, *58*, 92–103. [[CrossRef](#)]
57. Kejriwal, M. An AI Expert Explains Why It’s Hard to Give Computers Something You Take for Granted: Common Sense. 2021. Available online: <https://theconversation.com/an-ai-expert-explains-why-its-hard-to-give-computers-something-you-take-for-granted-common-sense-165600> (accessed on 3 September 2021).

58. Lin, H.; Ng, V. Abstractive summarization: A survey of the state of the art. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 1 February 2019; Volume 33, pp. 9815–9822.
59. Lin, B.Y.; Zhou, W.; Shen, M.; Zhou, P.; Bhagavatula, C.; Choi, Y.; Ren, X. CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2020, Online, 16–20 November 2020; pp. 1823–1840. [CrossRef]
60. Gordon, A.S.; Hobbs, J.R. *A Formal Theory of Commonsense Psychology: How People Think People Think*; Cambridge University Press: Cambridge, UK, 2017. [CrossRef]
61. Santos, H.; Kejriwal, M.; Mulvehill, A.M.; Forbush, G.; McGuinness, D.L.; Rivera, A.R. An experimental study measuring human annotator categorization agreement on commonsense sentences. *Exp. Results* **2021**, *2*. [CrossRef]
62. Koehler, J.; Nebel, B.; Hoffmann, J.; Dimopoulos, Y. Extending planning graphs to an ADL subset. In *European Conference on Planning*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 273–285.
63. Koenig, S.; Simmons, R.G. Risk-sensitive planning with probabilistic decision graphs. In *Principles of Knowledge Representation and Reasoning*; Elsevier: Amsterdam, The Netherlands, 1994; pp. 363–373.
64. Singhal, A. Introducing the knowledge graph: Things, not strings. *Off. Google Blog* **2012**, *5*, 16.
65. Bellomarini, L.; Fakhoury, D.; Gottlob, G.; Sallinger, E. Knowledge graphs and enterprise AI: The promise of an enabling technology. In Proceedings of the 2019 IEEE 35th International Conference on Data Engineering (ICDE), Macau, China, 8–12 April 2019; pp. 26–37.
66. Xu, D.; Ruan, C.; Korpeoglu, E.; Kumar, S.; Achan, K. Product knowledge graph embedding for e-commerce. In Proceedings of the 13th International Conference on Web Search and Data Mining, Houston, TX, USA, 3–7 February 2020; pp. 672–680.
67. Kejriwal, M. Knowledge Graphs and COVID-19: Opportunities, Challenges, and Implementation. *Harv. Data Sci. Rev.* **2020**, *11*, 300.
68. Kejriwal, M. *Domain-Specific Knowledge Graph Construction*; Springer: Berlin/Heidelberg, Germany, 2019.
69. Kannan, A.V.; Fradkin, D.; Akrotirianakis, I.; Kulahcioglu, T.; Canedo, A.; Roy, A.; Yu, S.Y.; Arnav, M.; Al Faruque, M.A. Multimodal knowledge graph for deep learning papers and code. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Online, 19–23 October 2020; pp. 3417–3420.
70. McGuinness, D.L.; Van Harmelen, F. OWL web ontology language overview. *W3C Recomm.* **2004**, *10*, 2004.
71. Berners-Lee, T.; Hendler, J.; Lassila, O. The semantic web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
72. Iliovski, F.; Szekely, P.; Kejriwal, M. Commonsense Knowledge Graphs (CSKGs). 2020. Available online: <https://usc-isi-i2.github.io/ISWC20/> (accessed on 2 September 2021).
73. Garcez, A.D.; Gori, M.; Lamb, L.C.; Serafini, L.; Spranger, M.; Tran, S.N. Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. *arXiv* **2019**, arXiv:1905.06088.
74. Hagras, H. Toward human-understandable, explainable AI. *Computer* **2018**, *51*, 28–36. [CrossRef]
75. Hoffman, R.R.; Mueller, S.T.; Klein, G.; Litman, J. Metrics for explainable AI: Challenges and prospects. *arXiv* **2018**, arXiv:1812.04608.
76. Turek, M. Explainable Artificial Intelligence (XAI). 2016. Available online: <https://www.darpa.mil/program/explainable-artificial-intelligence> (accessed on 2 September 2021).
77. 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.
78. Valentino, M.; Thayaparan, M.; Freitas, A. Explainable natural language reasoning via conceptual unification. *arXiv* **2020**, arXiv:2009.14539.
79. Lim, S.; Lee, J.G. Motif-based embedding for graph clustering. *J. Stat. Mech. Theory Exp.* **2016**, *2016*, 123401. [CrossRef]
80. Zhang, S.; Hu, Z.; Subramonian, A.; Sun, Y. Motif-driven contrastive learning of graph representations. *arXiv* **2020**, arXiv:2012.12533.
81. Garnham, A. Representing information in mental models. In *Cognitive Models of Memory*; Studies in Cognition; The MIT Press: Cambridge, MA, USA, 1997; pp. 149–172.
82. Dhingra, B.; Zhou, Z.; Fitzpatrick, D.; Muehl, M.; Cohen, W.W. Tweet2vec: Character-Based Distributed Representations for Social Media. *arXiv* **2016**, arXiv:1605.03481.
83. Wang, L.; Gao, C.; Wei, J.; Ma, W.; Liu, R.; Vosoughi, S. An Empirical Survey of Unsupervised Text Representation Methods on Twitter Data. *arXiv* **2020**, arXiv:2012.03468.
84. Schwag, V.; Bhagoji, A.N.; Song, L.; Sitawarin, C.; Cullina, D.; Chiang, M.; Mittal, P. Analyzing the robustness of open-world machine learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, London, UK, 15 November 2019; pp. 105–116.
85. Cambel, A.B. *Applied Chaos Theory: A Paradigm for Complexity*; Elsevier: Amsterdam, The Netherlands, 1993.
86. Zimmerman, M. Living with uncertainty: The moral significance of ignorance. *Analysis* **2009**, *69*, 785–787.
87. Edgington, D. Vagueness by degrees. In *Vagueness: A Reader*; Keefe, R., Smith, P., Eds.; MIT Press: Cambridge, MA, USA, 1997.
88. Achinstein, P. Circularity and induction. *Analysis* **1963**, *23*, 123–127. [CrossRef]
89. Senator, T. Science of Artificial Intelligence and Learning for Open-World Novelty (SAIL-ON). 2019. Available online: <https://www.darpa.mil/program/science-of-artificial-intelligence-and-learning-for-open-world-novelty> (accessed on 3 September 2021).
90. Nakaoka, F.; Oda, N. Some applications of minimal open sets. *Int. J. Math. Math. Sci.* **2001**, *27*, 471–476. [CrossRef]
91. Sennott, L.I. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Oper. Res.* **1989**, *37*, 626–633. [CrossRef]

92. Langley, P. Open-world learning for radically autonomous agents. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 13539–13543.
93. Kotz, A.M.; Dittrich, K.R.; Mülle, J.A. Supporting semantic rules by a generalized event/trigger mechanism. In *International Conference on Extending Database Technology*; Springer: Berlin/Heidelberg, Germany, 1988; pp. 76–91.
94. Batista-Navarro, R.T.; Kontonatsios, G.; Mihăilă, C.; Thompson, P.; Rak, R.; Nawaz, R.; Korkontzelos, I.; Ananiadou, S. Facilitating the analysis of discourse phenomena in an interoperable NLP platform. In *International Conference on Intelligent Text Processing and Computational Linguistics*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 559–571.
95. Keren, G. *Perspectives on Framing*; Psychology Press: London, UK, 2011.
96. Kühberger, A. Framing. In *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory*, 2nd ed.; Psychology Press: New York, NY, USA, 2017; pp. 79–98.
97. Gottschalk, F.; Van Der Aalst, W.M.; Jansen-Vullers, M.H.; La Rosa, M. Configurable workflow models. *Int. J. Coop. Inf. Syst.* **2008**, *17*, 177–221. [[CrossRef](#)]
98. Van Der Aalst, W.; Van Hee, K.M.; van Hee, K. *Workflow Management: Models, Methods, and Systems*; MIT Press: Cambridge, MA, USA, 2004.
99. NSF. National Artificial Intelligence (AI) Research Institutes. 2020. Available online: <https://beta.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes> (accessed on 2 September 2021).
100. Chemodanov, D.; Esposito, F.; Sukhov, A.; Calyam, P.; Trinh, H.; Oraibi, Z. AGRA: AI-augmented geographic routing approach for IoT-based incident-supporting applications. *Future Gener. Comput. Syst.* **2019**, *92*, 1051–1065. [[CrossRef](#)]
101. Launchbury, J. DARPA Perspective on AI. 2017. Available online: <https://www.youtube.com/watch?v=-O01G3tSypU&t=198s> (accessed on 2 September 2021).
102. Gunning, D. Machine common sense concept paper. *arXiv* **2018**, arXiv:1810.07528.
103. Kejriwal, M.; Knoblock, C.A.; Szekely, P. *Knowledge Graphs: Fundamentals, Techniques, and Applications*; MIT Press: Cambridge, MA, USA, 2021.
104. De Nicola, A.; Karray, H.; Kejriwal, M.; Matta, N. Knowledge, semantics and AI for risk and crisis management. *J. Contingencies Crisis Manag.* **2020**, *28*, 174–177. [[CrossRef](#)]