The Knowledge Engineering Review, Vol. 33, e11, 1–26.
doi:10.1017/S0269888918000139

Knowledge machines

PAUL SMART

Electronics & Computer Science, University of Southampton, Southampton SO17 1BJ, UK;
e-mail: ps02v@ecs.soton.ac.uk

Abstract

The World Wide Web has had a notable impact on a variety of epistemically relevant activities, many of which lie at the heart of the discipline of knowledge engineering. Systems like Wikipedia, for example, have altered our views regarding the acquisition of knowledge, while citizen science systems such as Galaxy Zoo have arguably transformed our approach to knowledge discovery. Other Web-based systems have highlighted the ways in which the human social environment can be used to support the development of intelligent systems, either by contributing to the provision of epistemic resources or by helping to shape the profile of machine learning. In the present paper, such systems are referred to as *knowledge machines*. In addition to providing an overview of the knowledge machine concept, the present paper reviews a number of issues that are associated with the scientific and philosophical study of knowledge machines. These include the potential impact of knowledge machines for the theory and practice of knowledge engineering, the role of social participation in the realization of knowledge-based processes, and the role of standardized, semantically enriched data formats in supporting the *ad hoc* assembly of special-purpose knowledge systems and knowledge processing pipelines.

1 Introduction

Knowledge engineering is a discipline that concerns itself with the processes, methods, and tools by which knowledge is acquired, represented, and utilized, typically for the purposes of building and deploying knowledge-based systems (Studer *et al*., 1998; Schreiber *et al*., 2000). The World Wide Web (henceforth the Web) has had a profound impact on the shape of activities that lie at the heart of this endeavour (Gil, 2011; Schreiber, 2013). These include, most notably, the approach that is adopted with respect to the acquisition of knowledge, as well as the way in which many forms of intelligent, knowledge-based system are both developed and deployed. The impact of the Web is perhaps most keenly felt in the case of the Semantic Web—the set of Web-based resources that seek to make the semantic content of syntactic, computational elements both explicit and amenable to advanced forms of machine-based processing (Berners-Lee *et al*., 2001; Shadbolt *et al*., 2006). But the *Social* Web has also played its part in transforming the scope and focus of traditional knowledge engineering efforts. Crucially, with the advent of Web 2.0, the Web has emerged as an important platform for large-scale social participation, and this has arguably transformed our understanding of the role of socio-computational systems in the realization of knowledge processes. As Brian Gaines, one of the leading figures in knowledge engineering, rightly notes, '[i]n our era, computer technology and human–computer interaction have come to play a major role in knowledge processes, facilitating a level of knowledge generation, dissemination, access, and utilization beyond that we have ever known' (2013: 135).

One of the recent foci of research attention within the Web and Internet science community is a class of systems called *social machines* (Hendler & Berners-Lee, 2010; Smart & Shadbolt, 2014; Smart *et al*., 2014; Hendler & Mulvehill, 2016; Shadbolt *et al*., 2016). These are systems that feature the synergistic interaction of resources drawn from both the social and technological realms. An important category of social machines are concerned with the realization of knowledge processes and the generation of epistemic outcomes. In the present paper, such systems are referred to as *knowledge machines*. The epistemic power and potential of knowledge machines is evidenced by systems such as Wikipedia, which provides a compelling example of the role social machines can play in knowledge acquisition (see Shadbolt, 2013). It is also evidenced by a variety of lesser known systems that support the discovery of knowledge (see Section 3) or that implement knowledge processing routines (see Section 6).

The main aim of the present paper is to outline the concept of knowledge machines (see Section 2) and provide concrete examples of systems that qualify as knowledge machines (see Sections 3 and 6). The paper also seeks to highlight the potential transformational impact of knowledge machines with respect to the discipline of knowledge engineering (see Section 2), the nature of the mechanisms that realize knowledge-relevant processes (see Sections 5 and 7), the role of social participation in the realization of knowledge-oriented processing routines (see Section 6), and the role of standardized, semantically enriched data formats in supporting the *ad hoc* assembly of special-purpose knowledge systems and knowledge processing pipelines (see Section 8). The paper concludes by emphasizing the importance of knowledge machines as a focus for future scientific and philosophical work (see Section 9).

2 Knowledge machines

Knowledge machines are members of a class of systems that have been referred to as social machines. In order to help us make sense of the term 'knowledge machine' it is therefore useful to understand what is meant by the term 'social machine'.

There are, in fact, a variety of views as to the meaning of the term 'social machine'. According to one view—which I will dub the *content creation view*—social machines are Web-based socio-technical systems that feature a division of labour between the constituent social and technological elements. In particular, the human elements of such systems are deemed to play a role in the creation of online content, while the technological components are relegated to roles of a more administrative nature. This view was first proposed by Berners-Lee and Fischetti (1999) who were among the first to use the term 'social machine' in a Web-based context. In *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*, Berners-Lee and Fischetti write that:

> Real life is and must be full of all kinds of social constraint—the very processes from which society arises. Computers can help if we use them to create abstract social machines on the Web: *processes in which the people do the creative work and the machine does the administration* (1999: 172, 74 emphasis added).

There is undoubtedly something compelling about this idea of a social machine as a system in which the bulk of the online content is generated by the human community. Assuming that the notion of 'creative work' in the above quotation should be interpreted in terms of the creation of online content (e.g. uploading an image or writing some text), then it seems that Berners-Lee and Fischetti's characterization can be applied to many systems that have come to dominate the contemporary Web. These include, for example, systems such as Wikipedia, Twitter, Facebook, YouTube, Quora, and Flickr. Such systems are emblematic of an important shift in the way Web-based systems are developed. Instead of an online system being designed by a select group of individuals and pre-populated with content, we are now confronted with an alternative approach, one in which the bulk of the software engineering effort is geared to the provision of a platform that is specifically designed to support community-driven modes of content generation.

Despite its appeal, the content creation view has been criticized on the grounds that it cannot account for the functional diversity of the constituent elements of a social machine. In particular, it has been suggested that the functional roles of the social and technological elements should not be restricted to those of a purely 'creative' or 'administrative' nature (see Smart *et al*., 2014). An alternative to the content creation view comes in the form of the *mechanistic view* of social machines[1]. According to this view, a social machine is an online (i.e. Web- or Internet-based) system whose events, states, and processes are realized by the operation of socio-technical mechanisms, where a socio-technical mechanism is simply a mechanism whose constituent elements (i.e. components) are drawn from either the social or technological realms. In short, the mechanistic view embraces an idea that is common to many views of social machines, namely, the idea that social machines are socio-technical systems. Where it differs from other views is with respect to the way in which a socio-technical system is, itself, defined. The mechanistic approach to social machines is thus rooted in a specific area of philosophy whose intellectual remit is the study of mechanisms (Glennan, 2017; Glennan & Illari, 2018). According to the mechanistic view, a social machine is a system whose phenomena[2] (i.e. events, states, and processes) are realized by mechanisms that combine the material fabric of the online technological environment[3] with the resources of the human social environment. In more prosaic terms, a socio-technical mechanism is a mechanism that has people and technology as its component parts, and it is the inter-operation of these component parts that informs our understanding of system-level events, states, and processes.

Relative to this conception of social machines, a knowledge machine can be defined as follows:

> *A knowledge machine is a social machine that engages in a form of knowledge-related activity. Such knowledge-related activities include those associated with the elicitation, acquisition, and representation of knowledge, as well as those associated with the discovery of knowledge and the development of intelligent systems.*

In essence, knowledge machines are a particular kind of social machine. Their distinguishing feature is the nature of the phenomena that are realized by socio-technical mechanisms. In the case of knowledge machines, the processes of interest are those of the knowledge variety—that is, they are processes that support the discovery, representation, acquisition, and/or exploitation of knowledge.

The status of some system as a knowledge machine thus turns on two criteria: epistemic relevance and socio-technical hybridity. The first criterion (i.e. epistemic relevance) concerns the nature of the process that is the target of a mechanistically oriented explanatory effort: Do we have a process that counts as a knowledge-relevant (e.g. knowledge-producing) process? The second criterion (i.e. socio-technical hybridity) relates to the (material) nature of the mechanisms that realize the target process: Is the focal process realized by a socio-technical mechanism? If the answer to both these questions is positive, then we have a socio-technical system that qualifies as a *bona fide* member of the class of knowledge machines[4].

For the most part, we can regard a knowledge machine as a social machine that is involved in the production of epistemic outputs, many of which are likely to assume the form of propositional statements

[1] The mechanistic view and the content creation view are not the only views of social machines to have been discussed within the philosophical and scientific literature. Palermos (2017) for example, countenances a *cognitive systems view* of social machines, which casts social machines as distributed cognitive systems—that is, as socio-technical systems that perform cognitive tasks. A not altogether unrelated view is proposed by Hooper *et al*. (2016). They suggest that social machines should be regarded as problem-solving organizations (this is what we might call the *problem-solving view*). Finally, there are a number of views that highlight the role of the human social environment in shaping the capabilities and performances of AI systems. These include the *sociable machines view* (Hendler & Mulvehill., 2016) and the *socially situated machines view* (Smart & Madaan., 2017).

[2] More precisely, the phenomena of interest are what Kaiser and Krickel (2017) refer to as *object-involving occurrents*. These are said to consist of 'an object (or system) that is engaged in a certain occurrent' (Kaiser & Krickel., 2017: 24), where the term 'occurrent' is simply a shorthand way of referring to events, states, and processes.

[3] Given the status of social machines as *online* socio-technical systems, Internet communication protocols are likely to play a crucial role in supporting the flow of information between the components of a socio-technical mechanism.

[4] Note that the status of some system as a social machine is likely to require some form of empirical analysis. This is because the status of some system as a social machine depends on the mechanisms that are deemed to be responsible for system-level phenomena. For the most part, such mechanisms will be identified in the same way that mechanisms are identified in other scientific disciplines, for example, via the use of observational, experimental, and computer simulation techniques. Such forms of analysis may not be required in all cases, however. This is because the science of social machines is, at least in part, an engineering discipline and socio-technical mechanisms may be built from the ground up to realize a particular function. In this case, the material constituents of the mechanism will be relatively obvious, at least to those who are involved in the effort to build social machines.

regarding some body of domain-specific knowledge. Cast in this light, the notion of a knowledge machine is similar (but not identical) to a number of concepts that have appeared in the epistemological literature[5]. These include the concept of a *socio-epistemic engine* in social epistemology (Goldman, 2011) and the concept of an *epistemic group agent* in virtue epistemology (Palermos, 2015) (see also Section 7). The scope of the knowledge machine concept is, however, somewhat broader than either of these concepts. In particular, when it comes to knowledge machines, we are not merely concerned with systems that enable us to produce knowledge, we are also concerned with activities that involve, for example, the organization and representation of knowledge[6], as well as the development of intelligent systems that are able to behave in a manner that respects the epistemic infrastructure of some focal domain of interest.

The value of the knowledge machine concept comes from the way it helps us appreciate the power and potential of the Web from an epistemic perspective. A crucial point of interest here relates to the way in which knowledge machines are poised to transform our traditional approach to knowledge engineering. Note, for instance, that many of the activities in which knowledge machines are involved are also ones that we typically associate with the discipline of knowledge engineering. This includes a range of activities that seek to acquire, represent, and model domain-specific knowledge (Studer *et al*., 1998; Schreiber *et al*., 2000). Such forms of overlap encourage us to revise our views regarding the way that many knowledge engineering activities are (or at least could be) realized. With the knowledge machine concept to hand, we are thus able to adopt a somewhat 'distributed' approach to knowledge engineering[7], one in which the Web and (perhaps) society-at-large are poised to participate in the mechanistic realization of key knowledge engineering activities.

In order to help us better appreciate this point about the transformation of traditional approaches to knowledge engineering, imagine that you are tasked with the development of a palaeontological knowledge system, one that will serve as an online repository of information regarding the characteristics of dinosaur species. From the perspective of traditional knowledge engineering, you might attempt to approach this task by first engaging in an iterated sequence of conventional knowledge elicitation activities (Shadbolt & Smart, 2015). You might thus seek to acquire (and actively elicit) information from various sources (including expert palaeontologists). You would then, let us suppose, attempt to implement a database to store the acquired information and generate the code to display the information in, for example, a conventional Web browser.

But now let us consider an alternative approach to 'knowledge acquisition', one that is inspired by the sort of approach adopted by Wikipedia. In this case, you simply provide the technological infrastructure that is needed to enable the human user community to create and edit online content for themselves. It is then the human user community that is assigned the task of populating the relevant 'knowledge base' with appropriate content (see Section 5 for more on this).

Hopefully, this example helps to illustrate at least one of the ways in which the concept of knowledge machines is relevant to the theory and practice of knowledge engineering. Beyond this, however, the knowledge machine concept helps to reveal issues that connect a number of different disciplines (e.g. knowledge engineering, Web science, and contemporary epistemology). In subsequent sections, I attempt to provide an initial overview of these issues. I also aim to present examples of knowledge

[5] Historical precursors to the knowledge machine concept can also be found in literature of a more technologically oriented nature. Writing in 1992, for example, David Gelernter presents the notion of a *knowledge plant*, which is a computational system designed to extract benefit from network-mediated data streams. 'Instead of venting…data into the info-smog', Gelernter writes, 'we could treat our data sources as plunging waterfalls waiting to drive software powerplants that convert data into knowledge' (1992: 112). This quotation helps to reveal a crucial difference between Gelernter's notion of a knowledge plant and the concept of a knowledge machine. A knowledge plant is, as Gelernter himself notes, a system that is 'realized by a software ensemble' (1992: 112). This contrasts with a knowledge machine, which is a system whose epistemic processes are realized by a *socio-technical* mechanism.

[6] This is particularly evident when it comes to systems that aim to generate epistemic resources (e.g. computational ontologies) for the Semantic Web.

[7] The idea of a distributed approach to knowledge engineering is based on the notion of distributed cognition, as discussed in the cognitive science literature (Hutchins, 1995). The core idea is that socio-technical systems are able to implement some of the activities that we typically associate with knowledge engineering, for example, the attempt to elicit, acquire, model, and exploit human knowledge.

machines that are relevant to the processes of knowledge discovery (see Section 3), knowledge acquisition (see Section 5), and knowledge exploitation (see Section 6). Given the scope and complexity of this topic, it is clearly impossible to cover everything. Important omissions include a range of systems that participate in the generation of semantically rich content. Such systems include collaborative ontology authoring environments (Simperl & Luczak-Rösch, 2014), semantic wikis (Krötzsch *et al.*, 2007), and a variety of online multiplayer games (Siorpaes & Hepp, 2008). Given the reliance of these systems on socio-technical mechanisms and their role in the generation of epistemic outputs (e.g. Semantic Web resources), many of these systems are likely to qualify as *bona fide* knowledge machines[8].

3 Knowledge discovery

In recent years, an important class of systems has emerged to support the process of knowledge discovery. Such systems are typically referred to as *citizen science systems* (Lintott & Reed, 2013). One of the primary aims of these systems is to co-opt the efforts of human volunteers into the scientific process, typically by enabling large groups of individuals to perform scientifically relevant tasks, such as data acquisition and analysis.

Perhaps one of the best examples of a citizen science system is Galaxy Zoo (Lintott *et al.*, 2008). This is a system that was originally designed to support the morphological classification of nearly one million galaxies that were imaged as part of the Sloan Digital Sky Survey. This task, it should be clear, is one that requires a significant amount of time and effort. As a result, Galaxy Zoo was established as an online, Web-based system that enabled casual users to assist with the galaxy classification effort (see Figure 1).

As a citizen science system, Galaxy Zoo was a resounding success, yielding more than forty million individual galaxy classifications (see Lintott *et al.*, 2008). In addition, Galaxy Zoo has yielded a number of important scientific discoveries. These include an astronomical phenomenon known as 'Hanny's Voorwerp' (Lintott *et al.*, 2009) and a previously unknown class of greenish-coloured galaxies, aptly called Green Pea galaxies (Cardamone *et al.*, 2009).

Following the success of Galaxy Zoo, a number of other citizen science systems have been developed to support the analysis of astronomical data. These include the Planet Hunters system, which aims to support the detection of extra-solar planets. As with Galaxy Zoo, this system has yielded a number of important scientific discoveries, including the first circumbinary planet discovered in a four-star system (Schwamb *et al.*, 2013) and the discovery of more than 40 planet candidates in the habitable zone of their parent stars (Wang *et al.*, 2013).

Citizen science systems such as Galaxy Zoo and Planet Hunters attempt to recruit volunteers for the purpose of analyzing some body of data. Other systems, however, focus their attention on the acquisition of data, typically by using the human social environment as a form of biological sensing platform. An excellent example of such a system is eBird (Sullivan *et al.*, 2009). eBird harnesses the observational efforts of thousands of volunteers in order to gather information about the distribution and abundance of bird species. This provides a valuable source of real-time information about avian population dynamics. Given the scope and scale of the data acquisition effort, it is difficult to imagine how the body of data provided by eBird could be acquired in the absence of large-scale social participation. By 2013, for example, Sullivan *et al.* report that 'over 140 million observations had been submitted by 150,000 separate observers, who spent 10.5 million hours in the field collecting data' (2014: 32). Such a data set is an invaluable epistemic resource, helping to improve our understanding of the effects of climate change, pollution, and habitat loss on bird populations. It should also be clear that the dataset can be used as an aide to conservation efforts, providing information about species decline in particular regions and helping to

[8] As noted by an anonymous reviewer, another potential omission relates to systems that support collective forms of debate, reason, and argumentation (e.g. `http://www.debatepedia.org/`). Although such systems are commonly devoted to the expression and reconciliation of conflicting opinions, their outputs may, on occasion, qualify as epistemic outputs. In such cases, there seems little reason to discount the status of these systems as fully paid up members of the class of social machines.

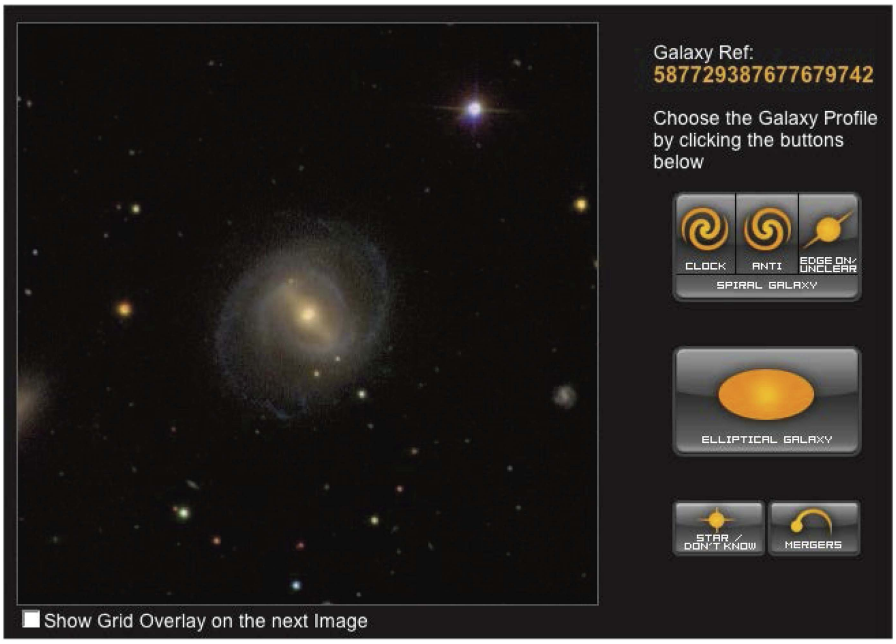

Figure 1 The interface of the Galaxy Zoo system. The system displays an image of a celestial object and asks participants to classify the object into one of six classes, namely elliptical galaxies, clockwise spiral galaxies, anticlockwise spiral galaxies, other spiral galaxies, stars, and mergers

coordinate clean-up operations in the face of environmental catastrophes (e.g. oil spills) (see Sullivan *et al*., 2009).

One of the things that makes eBird interesting as a citizen science system is the way it attempts to address issues of user motivation and community engagement. The problem is that human individuals are not required to participate in a citizen science system, and thus the long-term viability of the system is at risk if community interest should begin to wane. This is, of course, a problem that is not specific to citizen science systems. Inasmuch as we regard knowledge machines (and social machines, more generally) as socio-technical organizations, then it should be clear that social participation is the lifeblood of a knowledge machine: if an online system fails to attract the interest of a sufficient number of human participants, then it will never exist as a knowledge machine. And if the interest of existing participants cannot be maintained, then a knowledge machine will eventually cease to be. Social participation is not an optional add-on for knowledge machines; it is a feature of existential significance.

The challenge of attracting and maintaining human interest is one of the major targets of research into incentivization techniques or incentivization mechanisms (Tokarchuk *et al*., 2012; Scekic *et al*., 2013). Let us therefore dub this challenge the *incentivization challenge*. The incentivization challenge, it should be clear, is a challenge for all social machines, irrespective of whether or not they qualify as knowledge machines. There are, however, reasons to think that the incentivization challenge is of particular relevance to knowledge machines. One such reason is tied to the quality of epistemic outputs. As noted by Watson and Floridi (2018), issues of social scale are relevant to epistemic processing, with greater levels of social participation having an overall positive impact on the efficiency of evidentiary processing and the detection of anomalous data[9].

There have been a number of attempts to improve our understanding of the factors that motivate user participation in a variety of online systems, with altruism, a sense of community involvement, and social recognition emerging as particularly important factors (e.g. Tokarchuk *et al*., 2012). The design of eBird is

9 Watson and Floridi (2018) go on to note the virtues of the online environment with regard to issues of social scale. ‘Only online platforms’, they suggest, ‘offer the kind of scalability required to host hundreds of thousands of volunteers for any given project, and only at these volumes does the data processing power of untrained amateurs begin to compete with (or exceed) that of experts using traditional observation methods’ (Watson & Floridi, 2018: 753).

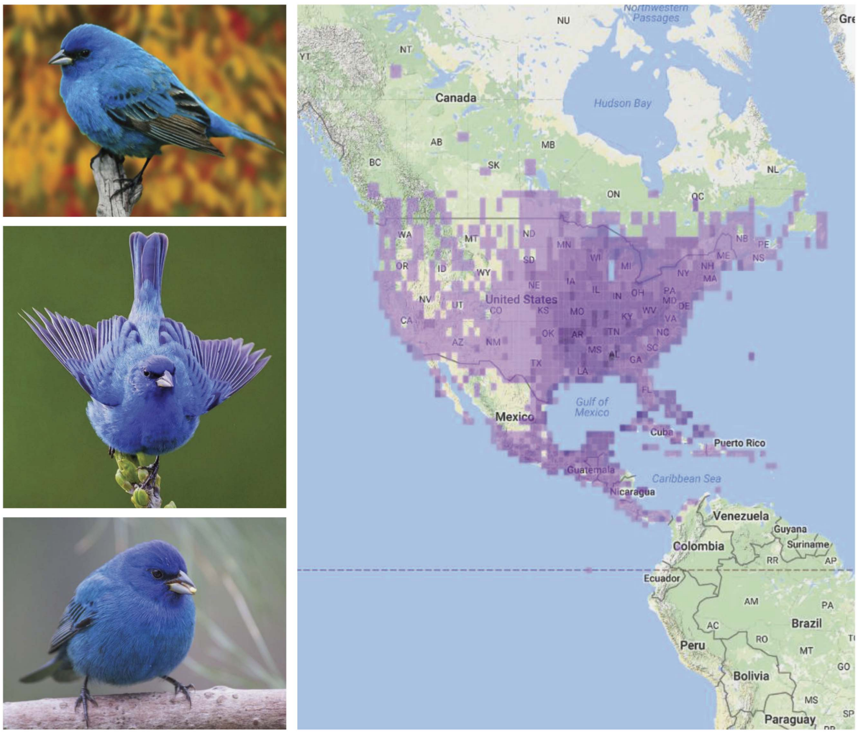

Figure 2 Distribution of sightings of the Indigo bunting (*Passerina cyanea*) across North America, aggregated across years and seasons (source: http://ebird.org)

certainly inspired by at least some of these factors[10]; however, eBird highlights another approach to ensuring the continued engagement of the user community. In particular, eBird attempts to use the data submitted by human observers as a means of providing services back to the user community. One such service comes in the form of a palette of easy-to-use data visualization and analysis tools that enables community members to identify the most likely places to spot particular birds of interest. Figure 2, for example, shows the frequency distribution of sightings of the Indigo bunting (*Passerina cyanea*).

Another approach to the problem of user motivation is illustrated by a class of social machines that go by the name of Games With a Purpose (GWAPs) (von Ahn, 2006; von Ahn & Dabbish, 2008; Savage, 2012). The general idea behind GWAPs is that peoples' game-playing actions can be used to perform a useful task. Given the apparent enthusiasm that people have for computer games[11], it seems that this approach has considerable promise in terms of harnessing human cognitive abilities for the purpose of tackling problems that lie beyond the current reach of AI algorithms. The problem, of course, is how to make a game sufficiently engaging to human game-players, while simultaneously allowing game-play actions to be exploited in the context of another task. There are, in general, two ways of approaching this problem. The first is to take the target task and attempt to make it as fun as possible, typically by scoring user performance and making lists of top-scoring players publicly available. For the sake of convenience, we can refer to this particular class of GWAPs as *goal-transparent GWAPs*. A second approach is to design the game in such a way that the relationship between game-player actions and the task to which such actions are applied is much less obvious. In this case, the fact that the game is being used to collect or analyze a body of, for example, scientific data is typically 'invisible' to the end-user—in fact, the human game-player may not even be aware that their game-play actions are being used to perform some other

[10] For example, eBird preserves the provenance of user contributions, enabling specific users to receive public recognition for important observations (e.g. the first sighting of a bird species in a particular geographic area).

[11] It is estimated that there are hundreds of millions of gamers worldwide who collectively spend more than 3 billion hours per week playing video games (see McGonigal, 2011).

task. Given that the real objectives of such games are invisible to the uninformed game-player, we can refer to this category of GWAPs as *goal-opaque GWAPs*[12].

An important example of a goal-transparent GWAP is the protein-folding game, Foldit (Cooper *et al*., 2010; Good & Su, 2011; Khatib *et al*., 2011). Foldit is an online multiplayer game that aims to derive accurate protein structure models via game-play responses. The game involves the presentation of improperly folded protein structures to human game-players. These are then manipulated using a combination of manual and automatic actions so as to maximize the score associated with a computed evaluation metric. The game is interesting because it provides a compelling example of the way in which social machines can be used to maximally exploit the distinctive capabilities of human and machine components (see Crouser *et al*., 2013). For example, in attempting to maximize their score, an individual user can interact with the protein structure—tugging and twisting the protein backbone as a means of exploring the target solution space. In doing so, the human game-players rely on a set of visual and spatial cognitive abilities that are, as yet, unmatched by the capabilities of existing AI systems. There is, however, an important role for machine-based processes in supporting the user's search for optimal protein conformations. In particular, the Foldit interface provides access to tools that implement so-called 'automatic moves'. These include, for example, a 'wiggle' routine that attempts to perform a localized search for high-scoring protein structures in the vicinity of the current structural candidate (see Cooper *et al*., 2010).

Perhaps one of the most impressive accomplishments of the Foldit system is its success in deciphering the crystal structure of the retroviral protease of the Mason-Pfizer monkey virus, a simian AIDS-causing virus (Khatib *et al*., 2011). The structure of this protein (an enzyme) had remained elusive despite attempts to solve the problem using conventional techniques. When assigned to the Foldit system, however, a group of Foldit players were able to produce an accurate three-dimensional (3D) model of the target protein within the space of just 3 weeks. This represents an important breakthrough for the biomedical research community, especially given the importance of retroviral proteases to HIV research (e.g. Kohl *et al*., 1988). The upshot is that we are provided with an important demonstration of the capacity of GWAPs to support the process of scientific discovery, and thus deliver advances in our communal body of scientific knowledge.

Compared to goal-transparent GWAPs, goal-opaque GWAPs are typically much harder to design. In particular, the game designer has to find a way of encouraging users to engage in actions that serve a dual purpose. First, the actions in question need to be consistent with the ludic objectives of the game. Second, the game-play actions need to be applicable to some form of scientifically relevant information processing. An interesting example of a game that manages to satisfy both of these constraints is Genes In Space (Coburn, 2014). Genes In Space is a game that was developed for the UK cancer research charity, Cancer Research UK. The purpose of the game, from the perspective of the human game-player, is to map a route through an asteroid-strewn landscape, collecting as much of a target substance (called Element Alpha) as they can (see Figure 3). In actual fact, by plotting (and then flying) a route through the virtual environment, the human game-players are assisting with the analysis of genomic data sets, helping scientists understand the genetic bases of (in this case) breast cancer. Importantly, from a knowledge machine perspective, Genes In Space is a game that succeeds in combining the respective capabilities of human agents and computational systems so as to yield a hybrid system with epistemically relevant properties (see Coburn, 2014).

The main purpose of Genes In Space is to support the analysis of a pre-existing body of scientific data. But not all GWAPs need to perform a data analytic function[13]. As is the case with citizen

[12] Despite the fact that citizen science systems and GWAPs are discussed separately in the present paper, there is nothing to prevent GWAPs (of either the goal-transparent or goal-opaque varieties) functioning as citizen science systems. In other words, there is no reason to assume that GWAPs and citizen science systems are disjoint categories of knowledge machine. There are clearly many cases where a game-oriented system can be used for serious scientific purposes, and, conversely, many forms of intellectual endeavour may benefit from the inclusion of ludic elements (e.g. Laszlo, 2004).

[13] Neither is the remit of GWAPs necessarily restricted to the scientific domain. One area of recent attention in the knowledge engineering community is the use of GWAPs to support the development of Semantic Web resources (Siorpaes & Hepp, 2008; Simperl *et al*., 2013).

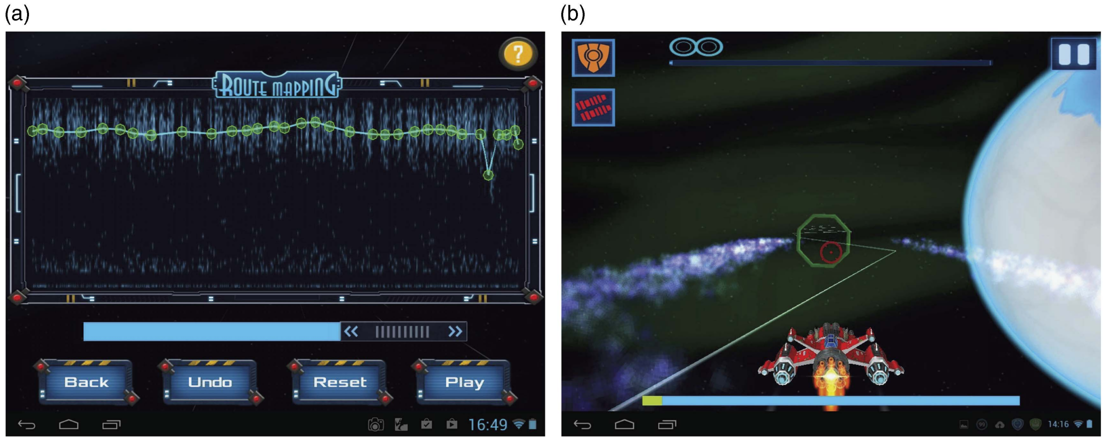

Figure 3 Two screenshots of the Genes In Space game. (a) The user plots a route, aiming to collect as much Element Alpha as possible. (b) The user attempts to pilot a spacecraft along the previously plotted route

science systems, GWAPs can also be used to collect as well as analyze bodies of scientific data. An interesting example of this comes in the form of a game called Sea Hero Quest[14] (Morgan, 2016, Spiers *et al.*, 2016). This is an example of what I have dubbed a goal-opaque GWAP, and thus the real purpose of the game (i.e. the acquisition of scientific data) is not immediately obvious to the uninformed game-player. In fact, the real purpose of the game is to provide information about a specific form of navigational competence, namely, the ability to orient oneself in a virtual 3D space and navigate to a target location. This is important, because impairments in spatial navigation ability are known to be one of the early signs of dementia. By thus understanding something about the normal parameters of navigational behaviour, the scientific community hopes to be able to detect the onset of dementia at an early stage.

The citizen science systems and GWAPs described in this section are undoubtedly an important expression of the growing interest in the computational power and potential of the human social environment. But should such systems be regarded as genuine members of the class of knowledge machines? I suggest they should, and the reason for this is that such systems meet the criteria presented in Section 2. First, there can be little doubt that these systems qualify as social machines. This is because such systems often involve complex forms of interaction between a set of social[15] resources (e.g. the game playing community) and a set of technological elements (e.g. the elements that are responsible for game execution and the tracking of user actions). Such forms of socio-technical entanglement with respect to the performance of a particular task (e.g. data analysis) are sufficient for us to see such systems as social machines. In addition to this, however, the systems are also ones that perform an epistemically relevant function: they support the acquisition and analysis of data, specifically for the purpose of expanding the epistemic horizons of the scientific community.

[14] See `http://www.seaheroquest.com/en/`

[15] The social status of these systems is not something that should be in doubt. For even GWAPs that fail to support direct player-to-player interactions still require a large number of participants in order to fulfil their epistemic purpose. In other words, important forms of knowledge discovery are often predicated on the contributions of *multiple* individuals. We see evidence of this in both the Genes in Space and Sea Hero Quest games. In the case of Genes In Space, multiple (independent) contributions are required in order to ensure the reliability of analytic outcomes. According to comments posted on the Cancer Research UK website, for instance, Genes In Space has been used to analyze the 'entire genomes of 1980 patients, each checked 50 times for accuracy' (see `http://www.cancerresearchuk.org/support-us/citizen-science/the-projects#citizenscience1`). A somewhat different role for multiple human contributions is apparent in the case of Sea Hero Quest. Here 'social' participation is a prerequisite for the success of the larger scientific effort, that is, the assembly of a normative dataset for the purposes of diagnostic testing.

4 Knowing us

The Web has provided a valuable opportunity for society-at-large to be recruited into an array of epistemically relevant activities. Citizen science systems provide us with a clear and unambiguous example of the sort of contribution that society can make to our attempts at knowledge discovery. There are, however, other ways of thinking about the epistemic implications of the Web. Just as the Web has provided the basis for large-scale forms of social participation in any number of online activities, so too it has opened the door to novel forms of social observation and analysis. Crucially, as our everyday social activities and endeavours become ever-more closely entwined with the online realm, so it becomes increasingly tempting to view the Web as part of the causally active physical fabric that *realizes* social processes (see Smart & Shadbolt, 2014). In other words, the Web presents us with a vision of society in which at least some kinds of social phenomena are subject (at least in part) to Web-based forms of computational realization. This raises a host of important issues regarding our ability to monitor, influence, and, indeed, create social processes. As noted by Strohmaier and Wagner:

> Today, the World Wide Web represents not only an increasingly useful reflection of human social behaviour, but everyday social interactions on the Web are increasingly mediated and shaped by algorithms and computational methods in general (2014: 84).

Using the Web as a platform for social observation and monitoring clearly raises concerns about privacy, surveillance, and social control. Nevertheless, the fact that important forms of social activity are now occurring in the online realm does provide us with a valuable opportunity to improve our understanding of society. This is important, because our contemporary society is a system of such complexity that its dynamical profile often resists our best attempts at prediction and explanation. In the wake of such complexity, it is perhaps tempting to think that the mechanistic underpinnings of social phenomena are doomed to forever lie beyond the reach of our (social) scientific grasp. But when we see the Web as part of the material fabric of society (i.e. as part of the physical machinery that realizes social phenomena), then we are afforded a much more positive perspective on the empirical and theoretical prospects of contemporary social science. This is because advances in mechanistic understanding (across all the sciences) are often linked to our ability to subject some target system to sophisticated forms of instrumentation and measurement. Perhaps, then, we can see the advent of the Web, and the current efflorescence of Web-enabled devices, as marking a potential sea change in our ability to establish an explanatorily and predictively potent grip on the social realm. Just as progress in other areas of science has followed hot on the heels of our ability to observe, measure, and monitor—consider the impact of the microscope and telescope on the fields of biology and astronomy—perhaps the Web is poised to progress the cause of the social sciences in a similar manner. In essence, what the Web gives us is an ability to observe (in more-or-less real time) the ebb and flow of social processes on a (potentially) global scale. As a result of such new-found observational abilities, we may, at last, be able to acquire the sort of data that informs the search for the mechanistic bases of (at least some kinds of) social phenomena.

It is at this point that claims about the epistemic significance of social machines begin to converge with the interests of the social science community. For the current interest and enthusiasm for mechanistic explanation in the social sciences (Hedström, 2005; Hedström & Ylikoski, 2010) dovetails perfectly with the current interest in the Web as a source of scientifically relevant information about the social environment (see Strohmaier & Wagner, 2014)[16]. Even if we retreat from the idea that the Web forms part of the material fabric that realizes social phenomena, there can be little doubt that the Web provides us with a significant, socially relevant observational ability, if only because so many of our everyday social activities are now tied up with the use of the Web. Such insights lie at the heart of a number of recent claims regarding the functional status of the Web (or parts thereof) as a form of 'digital socioscope'

[16] Such claims resonate with the idea that social machines serve as part of the realization base for social phenomena. In this respect, work into what are called Web Observatories is of particular interest, especially since such efforts often seek to observe and monitor the behaviour of social machines on the Web (Tiropanis *et al*., 2013; Tinati *et al*., 2015).

(Mejova *et al.*, 2015) or 'social observatory' (Caton *et al.*, 2015). They also lie at the heart of recent attempts to use the Web as a platform for what is called 'social mining' (Giannotti *et al.*, 2012).

When it comes to the role of knowledge machines in the process of knowledge discovery, therefore, we should not limit ourselves to the idea of the social environment as forming a literal part of the machinery that realizes parts of the scientific process; we can also think about the way in which knowledge machines are poised to provide us with a better understanding of the various forms of causal commerce that help to shape the structure of the social world.

5 Epistemic engineers

In Section 2 we encountered the idea of knowledge machines exploiting large-scale social participation for the purposes of knowledge acquisition. Wikipedia is, perhaps, the ultimate expression of this idea. By relying on a relatively simple set of human–computer interaction protocols, Wikipedia has emerged as a particularly important epistemic resource (Fallis, 2008, 2011), unrivalled with respect to its epistemic scope and roughly neck-and-neck with conventional encyclopaedias in terms of its epistemic reliability (Giles, 2005; Fallis, 2008). Wikipedia is, of course, a system that is intended to provide information for human consumption, and, in this respect, its epistemic outputs are unlike those encountered in traditional knowledge engineering projects. There is, in particular, no commitment to the sort of formal, machine-readable representations that are the typical outputs of conventional knowledge engineering efforts. This does not mean, however, that Wikipedia is irrelevant when it comes to the provision of such resources. DBpedia is one example where Wikipedia has helped to provide a structured epistemic resource that is relevant to the implementation of traditional knowledge-based systems (Auer *et al.*, 2007; Lehmann *et al.*, 2012)[17].

The case of Wikipedia serves as an important object lesson regarding the power and potential of knowledge machines to press maximal epistemic benefit from Web-based forms of social participation. The scale, scope, and complexity of Wikipedia exceeds anything that could have been developed by a single human individual, and it is for this reason that Wikipedia is often presented as an example of collective intelligence (Bonabeau, 2009; Malone *et al.*, 2010). In particular, the 'intelligence' of the human community with respect to the development of Wikipedia is sometimes seen to echo the intelligence exhibited by certain species of eusocial insect (e.g. Turner, 2011). Thus just as certain species of insect are able to coordinate their efforts so as to achieve feats of physical engineering that far outstrip the reach of their rather limited individual behavioural and cognitive repertoires, so too Wikipedia may be seen to represent a prodigious feat of socio-epistemic engineering, one that is, for the most part, beyond the ken of any single human individual.

What makes this comparison with insect societies all-the-more interesting is not just the scale of the collective achievements—the grand epistemic and physical edifices that emerge from the coordinated swirl of collective action—it is also the fact that such achievements may be grounded in the operation of similar *mechanisms*. Indeed, one kind of mechanism has proved to be of particular interest in the case of both insect societies and social machines. These are *stigmergic mechanisms*. The concept of stigmergy was first introduced by the French entomologist, Pierre-Paul Grassé, who used the term to account for the coordinated behaviour of termite colonies (see Theraulaz & Bonabeau, 1999). But it is not just the behaviour of eusocial insects that has been characterized in stigmergic terms; the notion of stigmergy has also been applied to systems that lie within the intellectual orbit of the social machine community. These include collaborative editing systems, such as Wikipedia (Parunak, 2005; Heylighen, 2016a), and open source software systems, such as Ushahidi (Marsden, 2013). This particular point of convergence helps to highlight the potential relevance of stigmergic mechanisms to our understanding of a variety of knowledge

[17] This does not mean that DBpedia should, itself, be viewed as a knowledge machine. What is crucial to the status of some system as a knowledge machine, at least from the standpoint of the mechanistic view, is the idea of some knowledge-relevant process being realized by a socio-technical mechanism. The output of such processes will typically be some form of epistemic product (e.g. a computational ontology), but this does not make the product, itself, a knowledge machine.

machines. Indeed, when it comes to systems like Wikipedia, the notion of stigmergy is important in helping us understand how the human social environment comes to play an explanatorily significant (and thus mechanistically relevant) role in knowledge acquisition processes. To help us see this, it will be useful to look at the notion of stigmergy in a bit more detail.

A useful definition of stigmergy is provided by Heylighen (2016a). Heylighen suggests that:

> …stigmergy is an indirect, mediated mechanism of coordination between actions, in which the trace of an action left on a medium stimulates the performance of a subsequent action (2016a: 6).

One thing that should be clear from this definition is that issues of environmental structuring and environmental mediation play a key role in stigmergic mechanisms. Typically, the concept of stigmergy implies that one or more agents will participate in the modification of some environmental resource, and this modification subsequently alters the response tendencies of other agents. The result is that complex structures and behavioural patterns emerge as a direct result of an agent's tendency to engage in actions that alter the conditions controlling the expression of other actions (including those expressed by agents that share the same 'local environment'[18]).

In applying the concept of stigmergy to Wikipedia, a couple of points are worth highlighting. The first is that the appeal to stigmergic mechanisms as a means of *explaining* the coordination of collective behaviours (and the subsequent emergence of complex structures) is a strategy that is in perfect accord with the mechanistic view of social machines outlined in Section 2. The mechanistic view of social machines, recall, focuses on the nature of the mechanisms that are responsible for some phenomenon of interest. This is precisely the sort of explanatory account that is provided by the appeal to stigmergic mechanisms. In the case of eusocial insects, the notion of stigmergy identifies a form of social (or perhaps eusocial!) mechanism that features the integration of forces and factors that are distributed across both the social (i.e. the insects) and the non-social (i.e. the stigmergic medium) realms. Similarly, in the case of the mechanistic view of social machines, what we are looking for is a materially hybrid mechanism that features a combination of both social and technological components. The only significant difference, here, is the nature of the stigmergic medium that works to coordinate collective behaviour. In this respect, it is important to note that the technological components of a socio-technical mechanism can sometimes play an active role in altering the configuration of a stigmergic medium. (Think of the way Wikipedia bots can alter the content of a Wikipedia page.) This, however, need not undermine the explanatory relevance of stigmergy as a means of accounting for the behavioural profile of knowledge machines. When it comes to Wikipedia, for example, Heylighen is quick to note that we can explain the behaviour of Wikipedia bots in the same sort of manner as we explain the behaviour of human editors:

> …a collectively edited website, like Wikipedia, may have some in-built procedures that automatically correct formatting errors, add links, or signal incoherencies. The fact that these actions are performed by computer programs (e.g. 'bots') does not fundamentally distinguish them from the actions of human contributors, since they all undergo the same stigmergic coordination (2016*b*: 52).

Stigmergy thus provides us with an important example of an explanatory account that is both mechanistic in spirit and oriented to the social domain. It is, as such, a concept that may be applicable to many kinds of social machines, especially those that involve the collaborative construction of online resources or the Web-mediated coordination of social activities.

A second point that is worth noting when it comes to stigmergic mechanisms and Wikipedia is the way in which the concept of stigmergy informs our understanding of Wikipedia *qua* knowledge acquisition system. To help us see this, let us turn our attention to one of the bugbears of traditional knowledge engineering: the problem of the *knowledge acquisition bottleneck* (Hayes-Roth *et al*., 1983). This is the problem of acquiring knowledge from a particular source (e.g. a human subject matter expert) in a manner

[18] The notion of a 'local environment' is typically understood in terms of spatial criteria. In the case of the Web, however, the local environment means the set of online resources (e.g. Wikipedia articles) that are accessed by multiple individuals.

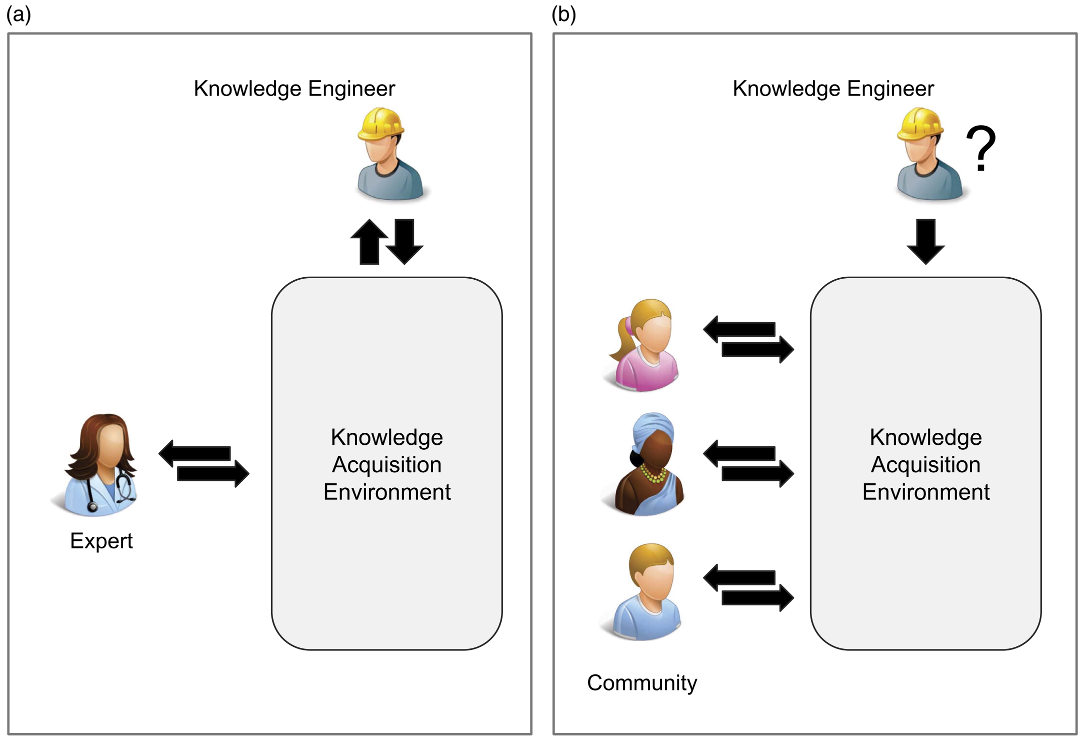

Figure 4 A situated view of knowledge acquisition from the perspective of (a) traditional knowledge engineering and (b) the perspective of a knowledge machine such as Wikipedia. In the case of traditional knowledge engineering (a), the knowledge engineer aims to create a situation that supports the elicitation of particular bodies of knowledge. In the case of Wikipedia (b), the role of the knowledge engineer, if there is one, is limited to the design of the technological system that supports the subsequent elicitation of collective knowledge via stigmergic mechanisms

that complies with, for example, the temporal and budgetary constraints of a knowledge engineering project. The approach to addressing the knowledge acquisition bottleneck is typically rooted in the careful selection and deployment of a range of knowledge elicitation techniques (Shadbolt & Smart, 2015). These are deemed to establish the sort of conditions that best support the elicitation of particular kinds of knowledge (see Hoffman & Lintern, 2006). A useful way of thinking about this state-of-affairs is to see the knowledge acquisition specialist—or knowledge engineer—as involved in the construction of situations that are of differential utility with regard to the elicitation of particular kinds of knowledge (see Figure 4a). In attempting to elicit procedural knowledge, for example, it may be of little practical value to establish a situation where the expert is required to respond, verbally, to a series of task-related questions. Instead, it may be much more appropriate to simply let the expert perform the relevant task and provide a running commentary as to the purpose of particular actions. What we end up with, in this case, is what might be called a *situated view of knowledge acquisition*: a way of thinking about the process of knowledge acquisition as the intelligent construction of situations that are of differential effectiveness with regard to the elicitation of particular kinds of knowledge.

Now consider how this situated view of knowledge acquisition is altered in the wake of systems like Wikipedia (see Figure 4b). Here, the role of the knowledge engineer—if indeed there is one—is reduced to the design of a technological system that is intended to solicit and record inputs from the user community. This is clearly important when it comes to the acquisition of particular bodies of knowledge. But note that once we turn our attention to the run-time operation of the system, we can see that an important form of 'situational engineering' is also being performed by the actual users of the system. The process of knowledge acquisition has, in a sense, become autocatalytic: as the actions of the user community progressively alter the structure of the online environment, so the cues, prompts, and affordances that work to elicit and structure further contributions are also altered. The result is that the community-at-large can be seen to play an active role in shaping the informational and social contexts that help (or, perhaps, hinder!) the acquisition of further knowledge.

When we think of knowledge acquisition from the perspective of knowledge machines like Wikipedia, we are thus provided with a vision in which the user community is engaged in a form of 'epistemic engineering'. Somewhat surprisingly, however, this role is not limited to the relatively straightforward idea of multiple individuals coming together to create an epistemically significant resource. There is also a sense in which we can see the user community as assuming the sort of role traditionally assigned to a knowledge engineer. The resulting vision of knowledge acquisition is thus one in which the human social environment is poised to play an important role in the progressive creation and configuration of situations that are relevant to the elicitation and acquisition of collective knowledge.

As a means of making this (admittedly awkward) idea a little clearer, consider the way in which the conversational exchanges between two people may provide the basis for a form of collaborative recall (see Sutton *et al*., 2010). Consider, for example, the following exchange (taken from Sutton *et al*., 2010) between a husband and wife discussing their honeymoon. In this particular exchange, the couple are trying to recall the name of a show they attended.

Wife: And we went to two shows. Can you remember what they were called?

Husband: We did. One was a musical, or were they both? I don't…no…one…

W: John Hanson was in it.

H: Desert Song.

W: Desert Song, that's it, I couldn't remember what it was called, but yes, I knew John Hanson was in it.

H: Yes.

This exchange is quite typical of our attempts to recall some piece of shared information in a social context, and it highlights something important about the nature of collaborative recall. Notice that the name of the show is successfully recalled as the result of the cross-cuing that occurs between the individuals. One person provides a cue, which by itself is inadequate to prompt the recall of the target information by either person. The cue does, however, provide the basis for the retrieval of another cue that is then fed back to the original person, and so on. This iterative cycle of reciprocal influence supports the progressive generation and elaboration of cues until, eventually, the conditions for recall are established.

What we see in the case of collaborative recall is thus somewhat reminiscent of the sorts of influence that occur in the context of (at least some) knowledge machines. Just as iterative cycles of information flow and influence support the progressive creation and elaboration of mnemonic cues that help to prompt the recall of target memories, so too the pattern of exchanges that occur between human agents and some form of online stigmergic medium can be seen to establish the situations that shape the structure of subsequent user contributions.

6 Expert systems

One of the major goals of traditional knowledge engineering was to support the development of systems that emulated the performance of human experts in some domain of interest. This was, by no means, the only objective of traditional knowledge engineering; nevertheless, the development of systems that embodied aspects of human expert knowledge (i.e. expert systems) was clearly one of the major drivers of knowledge engineering research throughout the 1980s and 1990s (Hayes-Roth *et al*., 1983; Hart, 1986; Kidd, 1987).

With the advent of the Web and, especially, the Semantic Web (Berners-Lee *et al*., 2001; Shadbolt *et al*., 2006), much of the early interest and enthusiasm in expert systems began to be eclipsed by a broader, and in some ways much grander, vision of the goal of knowledge engineering. In place of stand-alone systems that sought to embody the knowledge and expertise of individuals within narrow domains of interest, the Semantic Web yielded a vision of a globally distributed knowledge repository, one in which

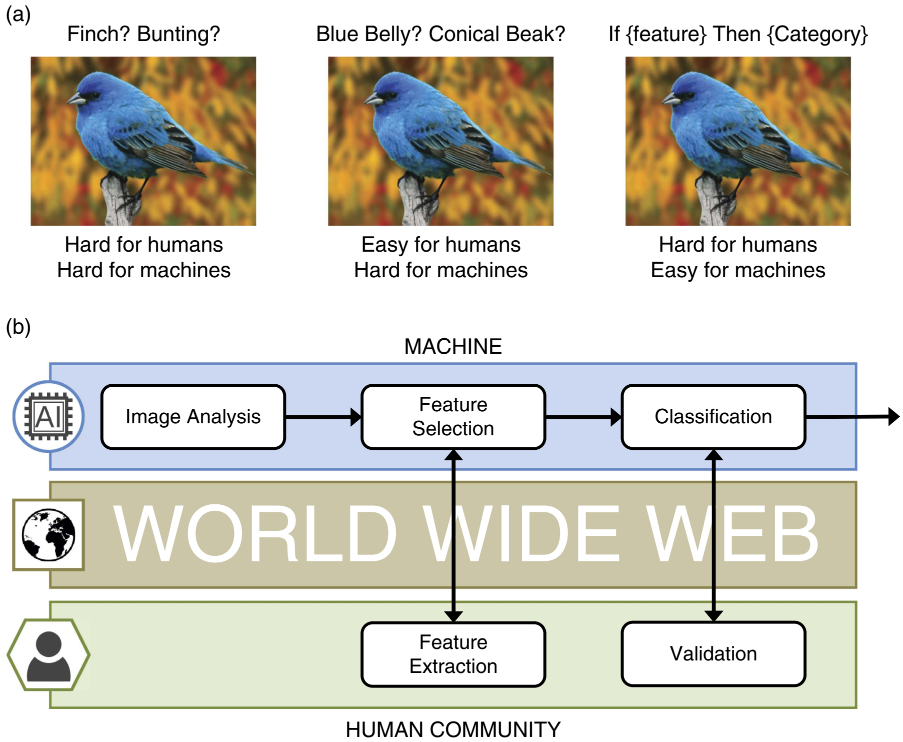

Figure 5 The socio-computational realization of knowledge-based processes. The classification of images based on the species of bird shown in the image is a task that is difficult for both humans and machines (a). The larger task can, however, be broken down into a series of smaller steps and assigned to the human and machine elements of a functionally integrated socio-technical system (b). The result is a hybrid knowledge-based system that relies on the distinctive (and, in this case, complementary) capabilities of the human and machine elements

the universe of human concepts could be represented in digital form. The result was that the focus of knowledge engineering research began to shift. The emphasis on detailed models of expert performance began to be replaced by an interest in the development of general-purpose computational ontologies (Gomez-Perez *et al*., 2004). For the most part, such ontologies were intended to be publicly accessible resources that were available for use in any number of knowledge-based systems and services.

There is clearly a sense in which the Semantic Web has had a profound impact on the scope and focus of knowledge engineering, both as an area of fundamental research and as an area of applied systems engineering (e.g. Gil, 2011). The nature of this impact should not be overstated, however. For even in the contemporary era, there is a sense in which the fundamental goals of knowledge engineering remain largely unchanged. Irrespective of whether our attention is focussed on traditional knowledge engineering or the attempt to build a Web-based semantic computing infrastructure, the goal of building intelligent systems remains a core focus of interest and concern. The Semantic Web has undoubtedly altered the way we seek to acquire, model, and exploit human knowledge, but it has not altered our sense of the fundamental importance of building systems that are poised to capitalize on the rich body of knowledge and experience that our species has managed to accumulate.

There are a number of ways in which the notion of knowledge machines is relevant to this vision. One example is provided by efforts that seek to create the sort of environment in which future intelligent systems are likely to emerge. Here we encounter a rich body of work that is concerned with the development of semantically rich resources, such as computational ontologies. At least some of the systems being developed in this space should arguably be counted among the ranks of knowledge machines; for such systems are often designed to co-opt the services of both humans and machines in delivering resources that support machine-based forms of reasoning and inference (Siorpaes & Hepp, 2008; Simperl *et al*., 2013).

Another way in which the knowledge machine concept is relevant to the implementation of intelligent systems is revealed by situations in which the knowledge machine itself functions as a form of knowledge-based system. A good example of such a system is described by Branson *et al.* (2014). Branson *et al.* were interested in combining the capabilities of human and machine elements in order to develop a hybrid system capable of identifying the species of bird depicted in a series of photographic images. This is a task that poses a significant challenge for both humans and machines (see Figure 5a). Branson *et al.*'s key insight, in this case, was to recognize the way in which the larger task of image classification could be decomposed into a series of smaller, more tractable steps, each of which could be assigned to the component elements of a materially hybrid, yet functionally integrated, socio-technical system (see Figure 5b). One of the steps in image classification, for example, concerns the extraction of specific features. These include features relating to, for example, the colour of the bird's plumage ('Does the bird have a blue belly?') and the shape of the bird's beak ('Does the bird have a beak that is conical in shape?'). Extracting such features from a natural scene is a task that is notoriously difficult for machine-based systems; it is, however, a task that is relatively easy for humans to perform (at least when the task does not rely on any *specialist* knowledge or ability). The result is that the feature extraction sub-task is one that can be delegated to the human elements of the larger system. The delegation task itself, however, is one that is far from straightforward. In particular, the efficiency of the larger classification process is one that depends on the intelligent coordination of the feature extraction sub-routine. There is, for example, no point in attempting to solicit information about beak shape if the machine-based components of the larger system can already infer that the beak can only be of one particular shape. Feature selection is thus something of a knowledge-intensive task in its own right, one that requires an ability to calculate the relative optimality (in an information theoretic sense) of different sequences of feature-oriented questions. This is a task that is highly amenable to machine-based processing, and it is for this reason that the task (or sub-task) of feature selection is one that ends up being assigned to the machine-based components of the larger information processing ensemble (see Figure 5b).

The result of all this is a knowledge-oriented socio-technical system that, I suggest, qualifies as a *bona fide* knowledge machine. In support of this claim, note that the system described by Branson *et al.* is one in which the activity of multiple human individuals is interleaved with the processing routines of a technological system. The system is, moreover, one that succeeds in delivering an epistemic outcome (a classification result) as a direct result of the operation of these hybrid information processing circuits. There is, in addition, no good reason to deny that the system is a genuine knowledge-based system, especially since the technological elements of the system trade in explicit encodings of the sort of knowledge that we might have otherwise expected to elicit from a human ornithological expert (e.g. knowledge relating to the optimal organization of the feature extraction task).

Branson *et al.*'s system thus provides us with a concrete example of a knowledge machine. It is, moreover, a knowledge machine that we can (and should) recognize as the modern equivalent of a classical expert system. The key difference between Branson *et al.*'s system and a conventional expert system lies not so much in the functional organization of their respective computational economies; neither is the difference to be found in the extent to which the systems rely on explicit encodings of domain-relevant knowledge. Instead, the primary difference relates to the way in which Branson *et al.*'s system relies on the human social environment as a source of readily available task-relevant knowledge. Thus rather than attempt to acquire, encode, and stockpile all the knowledge that is required to perform the classification task in advance of the task actually being performed, here we can see that the human social environment is being treated as a form of remote 'knowledge service', one that can be factored into the system's own computational routines as and when the need arises. The persistent presence of the human social environment, arising as a result of the Web, thus yields new approaches to the development of knowledge-based systems, enabling us to treat human agents as a potential source of epistemically relevant information[19].

[19] Such claims establish a useful point of contact with work in the cognitive sciences, especially work that highlights the role of just-in-time action as a means of exploiting the extra-organismic environment for cognitively relevant purposes (Clark, 2008; Myin & O'Regan, 2009).

7 Reliable mechanisms

When it comes to knowledge, issues of reliability are all-important. This is reflected in the fact that reliability lies at the heart of contemporary conceptions of knowledge within mainstream epistemology. A variety of conceptions of knowledge thus fall under the banner of what is known as reliabilism (Comesaña, 2011; Goldman, 2012). These include process reliabilism (Goldman, 1986), virtue reliabilism (Greco, 2010, 2012), and modal reliabilism (Pritchard, 2009: Ch. 2). In one way or another, all these forms of reliabilism appeal to the idea that knowledge-producing mechanisms should operate in a reliable manner; that is, in a manner that yields an overall preponderance of true (rather than false) beliefs.

It seems, therefore, that in order to be an acceptable producer of epistemic goods and services, the mechanisms associated with a knowledge machine will need to be reliable. Of course, in the case of knowledge machines, the mechanisms of interest are not ones that are located (solely) within the head of a single human individual. This marks an important difference with much of the work in mainstream epistemology. For the most part, the various forms of reliabilism presented above tend to focus on the individual human agent rather than a larger, functionally integrated nexus of social and technological elements. In spite of this, the role of reliability in resolving issues of positive epistemic standing does seem to be applicable to mechanisms that subtend the social and technological realms. Goldman (2011), for example, discusses the importance of reliability in relation to what he dubs *epistemic systems*[20]. Similarly, Palermos and Pritchard (2013) propose a social epistemological extension to virtue reliabilism, in which it is the reliability of the knowledge-producing *social mechanisms* (as opposed to a collection of intra-individual cognitive mechanisms) that represents the main focus of epistemological attention. Finally, Michaelian (2014) introduces the notion of distributed reliabilism as a means of extending the reach of reliabilist theory to the socio-technical realm. Distributed reliabilism, as defined by Michaelian, is thus an epistemological position that allows:

> …the process the reliability of which determines the epistemic status of a subject's belief to extend to include not only processing performed by other subjects but also processing performed by non-human technological resources (2014: 316).

This particular form of reliabilism, with its emphasis on socio-technical systems, is one that is clearly compatible with the idea of social machines functioning as knowledge-producing entities (i.e. as knowledge machines).

In addition to matters of a philosophical nature, reliability is clearly an issue of practical concern for those interested in knowledge machines. In general, we expect the mechanisms associated with a knowledge machine to operate in an epistemically desirable manner; that is, we expect mechanisms to produce outcomes that meet a range of epistemic desiderata, the most important of which is undoubtedly truth (see Goldman, 2002). In other words, the mechanisms associated with a knowledge machine should be organized in such a way as to yield informational outputs that are typically true. In addition to this, we might expect the larger system to modify its operation in the face of uncertainty. Ideally, the information processing economy of a knowledge machine should be sufficiently robust in the face of situations where the truth status of its outputs is at risk of being undermined, perhaps as the result of poor quality data or the sub-optimal performance of one or more of its component elements. In such situations, we might expect a knowledge machine to refrain from making any output (e.g. to suspend judgement). Alternatively, we might expect a knowledge machine to take remedial action and actively (re)configure its information processing economy so as to minimize the possibility of false outputs, perhaps by soliciting additional contributions from the human social environment or by switching to an alternative form of algorithmic processing.

[20] According to Goldman, epistemic systems are 'social system[s] that [house] social practices, procedures, institutions, and/or patterns of interpersonal influence that affect the epistemic outcomes of its members' (2011: 18). This is a concept that is broadly compatible with the idea of a knowledge machine. In particular, Goldman sees the epistemic standing of a social system as tied to the operation of one or more *social mechanisms* that are housed within the system. This much is clear from the emphasis that Goldman places on the role of organizational structure and patterns of inter-agent communication in the generation of epistemic outcomes.

The upshot of all this is an interest in the mechanisms that support the reliable operation of knowledge machines. In Section 3, we encountered one of the challenges associated with the effective operation of knowledge machines, namely, the incentivization challenge. Here we encounter a second challenge, which centres on the reliable operation of the socio-technical mechanisms that underlie the realization of knowledge-relevant processes. Let us call this particular challenge the *reliability challenge*. In parallel with the interest in incentivization mechanisms, the scientific community has begun to explore a rich array of strategies (or perhaps better, reliability-enhancing mechanisms) to deal with this particular challenge[21]. Such strategies include, but are not necessarily limited to, the following:

- Human ability: Some systems attempt to measure the ability of human agents with respect to a particular task and then weight user contributions accordingly. In the case of Galaxy Zoo, for example, Lintott *et al*. (2008) discuss the use of a weighting method that assigns greater weight to users who consistently agree with the majority of users. Another approach to human ability assessment is to measure the performance of human agents with respect to a set of problems for which the correct outputs are already known (see Weld *et al*., 2015). Such problems yield something of a 'gold standard' that provides insight into the performance capabilities of particular individuals.
- Human consistency: Given a typical classification task, such as that encountered in the case of Galaxy Zoo, it is entirely possible that the responses of a particular individual may be made at random. Perhaps, therefore, the success of an individual on a particular classification task is merely down to luck: perhaps the agent in question does not really have a genuine ability to discriminate galaxy types; rather, their (occasional) epistemic successes are merely due to some fortunate happenstance. One way of tackling this issue is to determine the consistency of responses made by a given individual with regard to the same or similar inputs. In the case of Galaxy Zoo, for example, the same galaxy image may be presented to the same individual multiple times across the course of their participation in the system. This tells us something about the consistency of the human classifier: it tells us whether the human agent is responding to particular features of the input stimuli in a standard way. This looks to be important, even in situations where the human agent consistently chooses the incorrect response: an individual that is consistently wrong can, in some cases, be just as informative as an individual who is always right.
- Process monitoring: Some systems attempt to monitor the behaviour of human users and adjust their operation to accommodate departures from behaviours that are deemed truth conducive. One example is The Milky Way Project, which forms part of the Zooniverse collection of citizen science projects. In this case, the system monitors the digital tools used by individual users and discounts contributions from users who fail to use all the tools provided (Simpson *et al*., 2012). Another form of process monitoring occurs in relation to systems that record the specific steps that are taken by an individual information processing element (e.g. a human agent) to perform a particular task. Here we see one of the virtues of socio-technical hybridization: by situating human action in a technological space that provides opportunities for the detailed monitoring of specific action sequences, we are provided with an opportunity to assess the reliability of the individual human 'components' of the larger knowledge-producing system. One example of this particular form of micro-monitoring comes from a study by Rzeszotarski and Kittur (2012). They describe a system called CrowdScape, which is designed to monitor the detailed behaviour of human agents as they engage in online tasks. Such capabilities are deemed to yield a 'digital fingerprint' of a task, which can then be used for the purposes of reliability assessment and quality evaluation.
- Adaptive coupling: Adaptive coupling mechanisms are mechanisms that support the active reconfiguration of a knowledge machine's information processing architecture so as to improve its chances of producing a veridical outcome. These mechanisms come in a variety of flavours. They include the adaptive routing of information to specific individuals at particular points in time (see Smart *et al*., 2010), as well as the intelligent (knowledge-driven) assignment of individuals to particular tasks (see Kamar *et al*., 2012). Shadbolt *et al*., (2016) also hint at a form of adaptive coupling when they note

[21] See Weld *et al*. (2015) and Steyvers and Miller (2015) for a useful overview of some of the mechanisms that can be used to enhance the reliability of socio-computational systems. Watson and Floridi (2018) also provide a useful overview of some of the quality control procedures adopted by the Galaxy Zoo system.

that '[c]urrent experimental work at the Zooniverse projects…pairs people who are gifted in particular tasks with others with complementary skills to achieve higher accuracy and task completion efficiency' (2016: 110).

- Ecological assembly: An important means of improving the reliability of knowledge-relevant mechanisms is to limit the material constitution of the mechanisms to those elements that are likely to yield the best overall level of performance[22]. This kind of reliability-enhancing mechanism differs from adaptive coupling in the sense that it features the proactive selection of elements that will be involved in the performance of a particular task. Ecological assembly is thus concerned with the initial formation of a knowledge-relevant mechanism, as opposed to the adaptive configuration of an existing mechanism. Typically, a system will attempt to recruit those elements (e.g. human agents) that are most suited to the task in question. Crowd building (see Demartini, 2015: 9) is one example of ecological assembly. In this case, the recruitment process is directed solely to the social realm and involves the attempt to evaluate the skills and expertise of particular human individuals (see also Bozzon *et al*., 2013).
- Social verification: One of the benefits of large-scale social participation is that the social environment can sometimes be relied on to support the verification of uncertain information. Examples of this occur in the case of Wikipedia, where the user community participates in the corrective editing of online factual content. Another example comes from Ushahidi, a crisis management and disaster relief platform (Okolloh, 2009; Gao *et al*., 2011). In this case, users of the system are able to click on a verification button in order to confirm the accuracy of existing reports (Gao *et al*., 2011). This feature is essential in disaster relief situations, where a variety of factors (including the changing nature of the situation itself) conspire to undermine the validity of previously submitted information.
- External verification: Resources external to a knowledge machine can sometimes be used to check and verify information. A particularly interesting example of this is provided by Lehmann *et al.* (2012). They discuss the use of DBpedia (a resource derived from Wikipedia) to check the validity of Wikipedia content. Given that DBpedia is amenable to various forms of machine-based processing, including logical consistency checking, it is able to detect semantic anomalies that appear in the original Wikipedia articles. Consider, for example, the unfortunate state-of-affairs in which an individual's date of birth is entered erroneously, so that it is represented as occurring *after* the individual's death. Here we have a rather delightful example of a situation in which a derivative knowledge resource (i.e. DBpedia) is used to check the epistemic integrity of the resource from which it derives (i.e. Wikipedia).
- Agent agreement: Agent agreement mechanisms rely on the consensus that is established by participants as a result of performing a task. One example of agent agreement comes in the form of *output agreement*. This occurs in situations where common responses are taken to be an indication of output validity. The ESP image labelling game is one example of output agreement (von Ahn & Dabbish, 2004). An alternative to output agreement is (you've guessed it!) *input agreement* (Law & von Ahn, 2009). This is used when the chances of multiple individuals converging on a common response is undermined as a result of high levels of descriptive entropy (i.e. the target resource can be described in many different ways). Input agreement relies on the ability of agents to determine whether they are processing the same resource based solely on the descriptive information that is supplied by other agents. An example of output agreement comes in the form of a system called TagATune whose aim is to solicit descriptive tags in respect of online audio resources (Law & von Ahn, 2009).

Note that many of these reliability-enhancing mechanisms are ones that span both the social and technological domains. In many cases, it is thus the interplay between the human agents and the technological elements that determines the truth status of the system's epistemic outputs. In this sense, a

[22] It should be noted that there is a potentially important parallel here with the notion of ecological assembly in the cognitive sciences. The focus in a cognitive scientific context is typically on the mechanisms that enable a particular cognitive agent to select and assemble a set of extra-organismic resources into some larger problem-solving whole. Clark (2008) provides a useful characterization of the idea in the form of the *Principle of Ecological Assembly*. According to this principle, 'the canny cognizer tends to recruit, on the spot, whatever mix of problem-solving resources will yield an acceptable result with a minimum of effort' (Clark, 2008: 13).

knowledge machine may be said to possess an epistemically relevant ability that is of a genuinely hybrid nature[23]. In other words, the ability is one that depends on the joint operation of both its constituent social and technological elements. This emphasis on systemic abilities and epistemic outcomes is one that establishes an interesting point of contact with recent work in epistemology, especially that which goes under the heading of virtue epistemology (Greco, 2007, 2010; Palermos & Pritchard, 2013).

8 Mechanical links, epistemic connections

For the most part, the focus of the paper up to this point has been on systems that function in isolation from one another. There is no attempt, for example, to integrate or embed the functionality of Branson *et al.*'s (2014) bird classification system within a larger economy of online systems, services, and applications. Despite this, we can clearly imagine situations where such forms of integration might be of epistemic value. Consider, for example, the hypothetical state-of-affairs depicted in Figure 6. Here we have a form of distributed processing that combines multiple knowledge machines into a functionally integrated information processing pipeline. In the initial stages of the depicted process, geotagged photos, as collected by the general public, are passed to Branson *et al.*'s classification system in order to determine the species of bird depicted in the photo (see Section 6). The resulting body of annotated photos is then made available to the eBird system (see Section 3) for assimilation into the eBird database. Finally, the content of the eBird database is itself made available in a form that permits flexible forms of integration, combination, and juxtaposition with a range of other data-driven applications and services. Such a capability would clearly be of tremendous value in respect of a broad range of epistemic endeavours. Imagine, for example, that you want to examine the impact of meteorological factors on avian population dynamics. In this case, an ability to juxtapose data regarding, for example, seasonal precipitation records with bird sighting density could be of crucial epistemic importance. Indeed, the statistical analysis of such data could lead to new hypotheses regarding the nature of the underlying mechanisms that are responsible for the observed correlations[24]. Perhaps, for example, low precipitation is associated with an increase in wild-fires and this destroys the breeding habitat of a particular bird species. Given access to appropriate bodies of data, you can go on to test this hypothesis, integrating data (from eBird) regarding seasonal fluxes in avian population dynamics with data obtained from fire mapping agencies (e.g. the Active Fire Mapping Program[25]).

The idea of linking otherwise independent online systems together to form ever-larger and more useful problem-solving organizations is one that is sometimes encountered in the literature on technology-mediated social participation (see Michelucci & Dickinson, 2016). It is also an idea that lies at the heart of Hendler and Berners-Lee's (2010) vision of the problem-solving potential of future social machines. Hendler and Berners-Lee note that today's social machines are somewhat limited with respect to their ability to exchange data across individual system boundaries. In response to this they suggest that we should move towards an era in which social machines are poised to participate in ever-larger information processing economies, serving as the loosely coupled constituents of systems that are dynamically assembled to meet the needs of specific problems.

Central to this vision of flexible integration and the *ad hoc* construction of special-purpose information processing pipelines is the idea of Web-optimized data formats that provide built-in support for knowledge-oriented processing. This is, of course, one of the key objectives of the Semantic Web initiative

[23] Palermos (2017) raises a similar point in respect of the Wikipedia system. In this case, Palermos suggests that the relevant form of reliability is 'an emergent, distributed property that belongs to Wikipedia as an overall, integrated system and cannot be accounted for in terms of the reliability of its underlying mechanistic and organismic components. Its reliability arises, instead, out of the actual and potential synergetic interactions of all components—organismic and mechanistic ones alike—at the same time' (2017: 972–973).

[24] This highlights the way in which one kind of mechanism (i.e. a socio-technical mechanism) can assist with the epistemic analysis of other mechanisms (e.g. the mechanisms responsible for species distribution). There is, of course, no reason why a socio-technical mechanism cannot be used to further our understanding of other socio-technical mechanisms. Indeed, one of the objectives of a knowledge machine may very well be to improve our understanding of the mechanisms that realize the epistemic performances of other knowledge machines.

[25] See `https://fsapps.nwcg.gov/afm/`

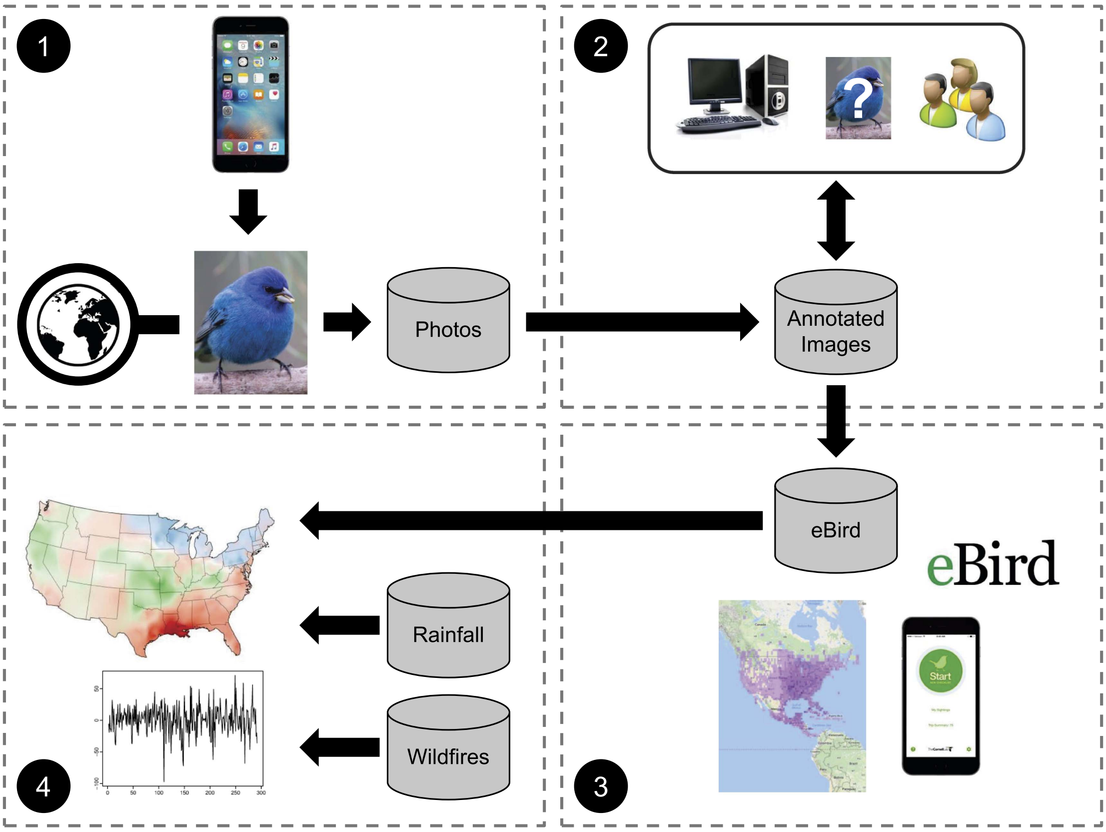

Figure 6 The flow of data through an information processing pipeline assembled from multiple knowledge machines and other online resources. (1) Geotagged photos are captured using a smartphone device and posted to an online repository. (2) The photos are classified using a variant of Branson *et al.*'s (2014) bird classification system. (3) Classified images are interpreted as 'sighting' data and imported into eBird. (4) Data from eBird data is combined with other data assets to test specific research hypotheses

(Berners-Lee *et al.*, 2001), and it is precisely for this reason that Hendler and Berners-Lee (2010) advocate the use of Semantic Web technologies as part of the effort to build future social machines (see also Gruber, 2008). The value of being able to effortlessly and automatically transfer data between physically and functionally disparate systems should not be underestimated here. Such forms of fluid informational exchange are often seen as relevant to efforts that seek to exploit the latent potential of so-called big data assets. Interestingly, data itself has come to be viewed as a form of commodity, on a par perhaps with the more conventional commodities (e.g. coffee, oil, and copper) that fuel the global economy. It is for this reason that data is sometimes presented as the 'new oil' (see Hirsch, 2014). Such a metaphor no doubt appeals to our intuitions regarding the potential economic value of big data. But there is, I suggest, an alternative way to view this oil-related metaphor. In this case, we can view data as a form of 'lubricant' that helps to ease the inevitable 'friction' that occurs at the points of contact between the mechanical elements of a complex, dynamic, and articulated piece of knowledge processing machinery. Data, in this sense, is the thing that enables individual knowledge machines to be assimilated into much larger computational organizations, some of which may themselves function as knowledge machines in their own right. None of this should force us to renege on the basic vision of knowledge machines as distinct, bounded systems that are able to function independently of other (online) systems. It does, however, serve as a useful reminder of the fact that knowledge machines can have a parallel existence as the material elements—the components, if you will—of much larger information processing mechanisms. It is in this sense, perhaps, that we can begin to appreciate the value of a commitment to standardized, semantically expressive data formats, such as those championed by proponents of the Semantic Web. For it is at this level—the level where individual knowledge machines are merged into larger mechanisms—where we see a role for suitably enriched data formats in lubricating the mechanical linkages between the pumps, pistons, and pulleys of ever-larger and ever-more sophisticated epistemic engines.

9 Conclusion

Knowledge machines are online systems that are concerned with the socio-technical realization of knowledge-related processes, many of which lie at the heart of traditional knowledge engineering efforts. These include, for example, the processes associated with the acquisition, representation, and exploitation of knowledge. The knowledge machine concept is thus intended to help us understand the potential transformative impact of the Web in relation to traditional forms of knowledge engineering. In essence, the knowledge machine concept is intended to prompt something of a ‘socio-technical turn’ in both the theory and practice of knowledge engineering. This does not mean, of course, that traditional forms of knowledge engineering are to be abandoned. Rather, the knowledge machine concept is intended to sensitize us to the opportunities emerging as a result of the ever-increasing social penetration of the Internet and Web. By highlighting the way in which knowledge engineering objectives can be subject to alternative (i.e. socio-technical) forms of mechanistic realization, the knowledge machine concept helps us appreciate some of the ways in which the online world might be pressed into useful epistemic service.

The present paper provides an initial overview of the knowledge machine concept. In Section 2, we saw how a mechanistic approach to social machines supports a philosophically robust characterization of knowledge machines—one that is rooted in the nascent sub-discipline of mechanical philosophy (Glennan, 2017; Glennan & Illari, 2018). From the standpoint of the mechanistic view, therefore, knowledge machines are Web- or Internet-based systems whose knowledge-relevant phenomena (e.g. knowledge-producing processes) are realized by socio-technical mechanisms. One of the virtues of this approach is that it reveals a number of opportunities for cross-disciplinary work. The interest in socio-technical mechanisms, for instance, serves a useful point of contact with work in both the social (e.g. Hedström & Ylikoski, 2010) and the cognitive (e.g. Palermos, 2017) sciences. It is, moreover, an interest that tallies with the socio-technical orientation of recent work in both contemporary epistemology (Carter *et al*., 2018) and the philosophy of science (e.g. Watson & Floridi, 2018).

In addition to an exploration of the knowledge machine concept, the present paper also identified specific examples of knowledge machines (see Section 3), discussed some of their associated mechanisms (see Section 5), and highlighted the role of Web technologies in supporting the emergence of ever-larger knowledge processing organizations (see Section 8). An additional point of interest surfaced in respect of some of the challenges that confront the attempt to ensure the effective operation of knowledge machines. One such challenge concerns the need to ensure that sufficient numbers of people engage with a knowledge machine. This was dubbed the incentivization challenge (see Section 3). A second challenge is to ensure that socio-technical mechanisms operate in a reliable (i.e. truth-conducive) manner. This was dubbed the reliability challenge (see Section 7). These two challenges do not exhaust the kind of challenges that are likely to arise in future work. In addition to incentivization and reliability, for example, it will no doubt be important to develop techniques that protect knowledge machines from various forms of malign intervention, including those that aim to derail or sabotage socio-epistemic processes.

Knowledge machines are likely to be an important focus area for future theoretical and practical work. Given that our success as a species is, at least to some extent, predicated on our ability to manufacture, represent, communicate, and exploit knowledge (see Gaines, 2013), there can be little doubt about the significance of knowledge machines for our own species. But in addition to their ability to harness the cognitive and epistemic capabilities of the human social environment, knowledge machines also provide us with a potentially important opportunity to scaffold the development of new forms of machine intelligence (see Smart & Madaan, 2017). Thus just as much of our own human intelligence may be rooted in the fact that we are born into a superbly structured and deliberately engineered environment (see Sterelny, 2003), so too the next generation of synthetic intelligent systems may benefit from a rich and structured informational environment that houses the sum total of human knowledge. In this sense, knowledge machines are not just important with respect to the continued epistemic prosperity of our own species, they may also work to support the emergence of a new class of epistemic prosumers—agents whose cognitive and epistemic successes are founded on the rich body of knowledge that our species has long sought to create and codify.

Acknowledgements

This work is supported under SOCIAM: The Theory and Practice of Social Machines. The SOCIAM Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1 and comprises the Universities of Southampton, Oxford and Edinburgh.

References

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R. & Ives, Z. 2007. DBpedia: a nucleus for a Web of open data. *Lecture Notes in Computer Science* **4825**, 722–735.

Berners-Lee, T. & Fischetti, M. 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. Harper Collins.

Berners-Lee, T., Hendler, J. & Lassila, O. 2001. The Semantic Web. *Scientific American* **284**(4), 34–43.

Bonabeau, E. 2009. Decisions 2.0: the power of collective intelligence. *MIT Sloan Management Review* **50**(2), 45–52.

Bozzon, A., Brambilla, M., Ceri, S., Silvestri, M. & Vesci, G. 2013. Choosing the right crowd: expert finding in social networks. In *16th International Conference on Extending Database Technology*, N. W. Paton, G. Guerrini, B. Catania, M. Castellanos, P. Atzeni, P. Fraternali & A. Gounaris(eds). ACM, 637–648.

Branson, S., Horn, G., Wah, C., Perona, P. & Belongie, S. 2014. The ignorant led by the blind: a hybrid human-machine vision system for fine-grained categorization. *International Journal of Computer Vision* **108**(1), 3–29.

Cardamone, C., Schawinski, K., Sarzi, M., Bamford, S. P., Bennert, N., Urry, C. M., Lintott, C., Keel, W. C., Parejko, J. & Nichol, R. C. 2009. Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society* **399**(3), 1191–1205.

Carter, A. J., Clark, A., Kallestrup, J., Palermos, O. S. & Pritchard, D. (eds) 2018. *Socially Extended Epistemology*. Oxford University Press.

Caton, S., Hall, M. & Weinhardt, C. 2015. How do politicians use Facebook? An applied Social Observatory. *Big Data & Society* **2**(2), 1–18.

Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.

Coburn, C. 2014. Play to cure: genes in space. *Lancet Oncology* **15**(7), 688.

Comesaña, J. 2011. Reliabilism. In *The Routledge Companion to Epistemology*, Bernecker, S. & Pritchard, D. (eds). Routledge, 176–186.

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z. & Foldit, Players 2010. Predicting protein structures with a multiplayer online game. *Nature* **466**(7307), 756–760.

Crouser, R. J., Ottley, A. & Chang, R. 2013. Balancing human and machine contributions in human computation systems. in *Handbook of Human Computation*, Michelucci, P. (ed.). Springer, 615–623.

Demartini, G. 2015. Hybrid human-machine information systems: challenges and opportunities. *Computer Networks* **90**, 5–13.

Fallis, D. 2008. Toward an epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology* **59**(10), 1662–1674.

Fallis, D. 2011. Wikipistemology. in *Social Epistemology: Essential Readings*, Goldman, A. I. & Whitcomb, D. (eds). Oxford University Press, 297–313.

Gaines, B. R. 2013. Knowledge acquisition: past, present and future. *International Journal of Human-Computer Studies* **71**(2), 135–156.

Gao, H., Barbier, G. & Goolsby, R. 2011. Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems* **26**(3), 10–14.

Gelernter, D. 1992. *Mirror Worlds*. Oxford University Press.

Giannotti, F., Pedreschi, D., Pentland, A., Lukowicz, P., Kossmann, D., Crowley, J. & Helbing, D. 2012. A planetary nervous system for social mining and collective awareness. *The European Physical Journal Special Topics* **214**(1), 49–75.

Gil, Y. 2011. Interactive knowledge capture in the new millennium: how the Semantic Web changed everything. *The Knowledge Engineering Review* **26**(1), 45–51.

Giles, J. 2005. Internet encyclopaedias go head to head. *Nature* **438**(7070), 900–901.

Glennan, S. 2017. *The New Mechanical Philosophy*. Oxford University Press.

Glennan, S. & Illari, P. M. (eds) 2018. *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. Routledge.

Goldman, A. I. 1986. *Epistemology and Cognition*. Harvard University Press.

Goldman, A. I. 2002. Précis of knowledge in a social world. *Philosophy and Phenomenological Research* **64**(1), 185–190.

Goldman, A. I. 2011. A guide to social epistemology. In *Social Epistemology: Essential Readings*, Goldman, A. I. & Whitcomb, D. (eds). Oxford University Press, 11–37.

Goldman, A. I. 2012. *Reliabilism and Contemporary Epistemology: Essays*. Oxford University Press.
Gomez-Perez, A., Fernandez-Lopez, M. & Corcho, O. 2004. *Ontological Engineering*. Springer.
Good, B. M. & Su, A. I. 2011. Games with a scientific purpose. *Genome Biology* **12**(135), 1–3.
Greco, J. 2007. The nature of ability and the purpose of knowledge. *Philosophical Issues* **17**(1), 57–69.
Greco, J. 2010. *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.
Greco, J. 2012. A (different) virtue epistemology. *Philosophy and Phenomenological Research* **85**(1), 1–26.
Gruber, T. 2008. Collective knowledge systems: where the Social Web meets the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web* **6**(1), 4–13.
Hart, A. 1986. *Knowledge Acquisition for Expert Systems*. Kogan Page.
Hayes-Roth, F., Waterman, D. A. & Lenat, D. B. 1983. *Building Expert Systems*. Addison-Wesley.
Hedström, P. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press.
Hedström, P. & Ylikoski, P. 2010. Causal mechanisms in the social sciences. *Annual Review of Sociology* **36**, 49–67.
Hendler, J. & Berners-Lee, T. 2010. From the Semantic Web to social machines: a research challenge for AI on the World Wide Web. *Artificial Intelligence* **174**, 156–161.
Hendler, J. & Mulvehill, A. M. 2016. *Social Machines: The Coming Collision of Artificial Intelligence, Social Networking, and Humanity*. Apress.
Heylighen, F. 2016a. Stigmergy as a universal coordination mechanism I: definition and components. *Cognitive Systems Research* **38**, 4–13.
Heylighen, F. 2016b. Stigmergy as a universal coordination mechanism II: varieties and evolution. *Cognitive Systems Research* **38**, 50–59.
Hirsch, D. D. 2014. The glass house effect: big data, the new oil, and the power of analogy. *Maine Law Review* **66**(2), 373–395.
Hoffman, R. R. & Lintern, G. 2006. Eliciting and representing the knowledge of experts. In *The Cambridge Handbook of Expertise and Expert Performance*, Ericsson, K. A., Charness, N., Feltovich P. & Hoffman, R. R. (eds). Cambridge University Press, 165–191.
Hooper, C., Bailey, B., Glaser, H. & Hendler, J. 2016. Social machines in practice: solutions, stakeholders and scopes. In *8th International ACM Web Science Conference*, Nejdl, W., Hall, W., Parigi P. & S. Staab, (eds). ACM, 156–160.
Hutchins, E. 1995. *Cognition in the Wild*. MIT Press.
Kaiser, M. I. & Krickel, B. 2017. The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science* **68**(3), 745–779.
Kamar, E., Hacker, S. & Horvitz, E. 2012. Combining human and machine intelligence in large-scale crowdsourcing. In *11th International Conference on Autonomous Agents and Multiagent Systems*, van der Hoek, W., Padgham, L., Conitzer, V. & Winikoff, M. (eds). **1**, IFAAMAS 467–474.
Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovic, Z., Baker, D. & Foldit, Players 2011. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences* **108**(47), 18949–18953.
Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popovic, Z., Jaskolski, M. & Baker, D. 2011. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology* **18**(10), 1175–1177.
Kidd, A. L. (ed.) 1987. *Knowledge Acquisition for Expert Systems: A Practical Handbook*. Plenum Press.
Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A. F., Scolnick, E. M. & Sigal, I. S. 1988. Active human immunodeficiency virus protease is required for viral infectivity. *Proceedings of the National Academy of Sciences* **85**(13), 4686–4690.
Krötzsch, M., Vrandečić, D., Völkel, M., Haller, H. & Studer, R. 2007. Semantic Wikipedia. *Journal of Web Semantics* **5**(4), 251–261.
Laszlo, P. 2004. Science as play. *American Scientist* **92**(5), 398–400.
Law, E. & von Ahn, L. 2009. Input-agreement: a new mechanism for collecting data using human computation games. In SIGCHI Conference on Human Factors in Computing Systems, Olsen, D. R., Arthur, R. B., Hinckley, K., Morris, M. R., Hudson, S. & Greenberg, S. (eds). ACM, 1197–1206.
Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P. & Auer, S. 2012. DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* **6** (2), 167–195.
Lintott, C. J. & Reed, J. 2013. Human computation in citizen science. In *Handbook of Human Computation*, Michelucci, P. (ed.). Springer, 153–162.
Lintott, C. J., Schawinski, K., Keel, W., Van Arkel, H., Bennert, N., Edmondson, E., Thomas, D., Smith, D. J. B., Herbert, P. D., Jarvis, M. J., Virani, S., Andreescu, D., Bamford, S. P., Land, K., Murray, P., Nichol, R. C., Raddick, M. J., Slosar, A., Szalay, A. & Vandenberg, J. 2009. Galaxy Zoo: 'Hanny's Voorwerp', a quasar light echo? *Monthly Notices of the Royal Astronomical Society* **399**(1), 129–140.

Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., Andreescu, D., Murray, P. & van den Berg, J. 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society* **389**(3), 1179–1189.

Malone, T. W., Laubacher, R. & Dellarocas, C. 2010. The collective intelligence genome. *MIT Sloan Management Review* **51**(3), 21–31.

Marsden, J. 2013. Stigmergic self-organization and the improvisation of Ushahidi. *Cognitive Systems Research* **21**, 52–64.

McGonigal, J. 2011. *Reality is Broken: Why Games Make us Better and How They Can Change the World*. Penguin Books Ltd.

Mejova, Y., Weber, I. & Macy, M. W. (eds) 2015. *Twitter: A Digital Socioscope*. Cambridge University Press.

Michaelian, K. 2014. JFGI: from distributed cognition to distributed reliabilism. *Philosophical Issues* **24**(1), 314–346.

Michelucci, P. & Dickinson, J. L. 2016. The power of crowds. *Science* **351**(6268), 32–33.

Morgan, J. 2016. Gaming for dementia research: a quest to save the brain. *The Lancet Neurology* **16**(13), 1313.

Myin, E. & O'Regan, J. K. 2009. Situated perception and sensation in vision and other modalities: a sensorimotor approach. In *The Cambridge Handbook of Situated Cognition*, Robbins, P. & Aydede, M. (eds). Cambridge University Press, 185–200.

Okolloh, O. 2009. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory Learning and Action* **59**(1), 65–70.

Palermos, O. & Pritchard, D. 2013. Extended knowledge and social epistemology. *Social Epistemology Review and Reply Collective* **2**(8), 105–120.

Palermos, S. O. 2015. Active externalism, virtue reliabilism and scientific knowledge. *Synthese* **192**(9), 2955–2986.

Palermos, S. O. 2017. Social machines: a philosophical engineering. *Phenomenology and the Cognitive Sciences* **16** (5), 953–978.

Parunak, H. V. D. 2005. A survey of environments and mechanisms for human-human stigmergy. In *International Workshop on Environments for Multi-Agent Systems', Vol. 3830 of Lecture Notes in Artificial Intelligence*, D. Weyns, P. H. V. Dyke & F. Michel (eds). Springer-Verlag, 163–186.

Pritchard, D. 2009. *Knowledge*. Palgrave Macmillan.

Rzeszotarski, J. & Kittur, A. 2012. Crowdscape: interactively visualizing user behavior and output. In *25th Annual ACM Symposium on User Interface Software and Technology*, Miller, R., Benko, H. & Latulipe, C. (eds). ACM, 55–62.

Savage, N. 2012. Gaining wisdom from crowds. *Communications of the ACM* **55**(3), 13–15.

Scekic, O., Truong, H.-L. & Dustdar, S. 2013. Incentives and rewarding in social computing. *Communications of the ACM* **56**(6), 72–82.

Schreiber, G. 2013. Knowledge acquisition and the Web. *International Journal of Human-Computer Studies* **71**(2), 206–210.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N. R., Van de Velde, W. & Weilinga, B. 2000. *Knowledge Engineering and Management: The CommonKADS Methodology*. MIT Press.

Schwamb, M. E., Orosz, J. A., Carter, J. A., Welsh, W. F., Fischer, D. A., Torres, G., Howard, W., Crepp, J. R., Keel, W. C. & Lintott, C. J. 2013. Planet Hunters: a transiting circumbinary planet in a quadruple star system. *The Astrophysical Journal* **768**(127), 1–21.

Shadbolt, N., Hall, W. & Berners-Lee, T. 2006. The Semantic Web revisited. *IEEE Intelligent Systems* **21**(3), 96–101.

Shadbolt, N., Van Kleek, M. & Binns, R. 2016. The rise of social machines: the development of a human/digital ecosystem. *IEEE Consumer Electronics Magazine* **5**(2), 106–111.

Shadbolt, N. R. 2013. Knowledge acquisition and the rise of social machines. *International Journal of Human-Computer Studies* **71**(2), 200–205.

Shadbolt, N. R. & Smart, P. R. 2015. Knowledge elicitation: methods, tools and techniques. In *Evaluation of Human Work*, 4th edition Wilson, J. R. & Sharples, S. (eds). CRC Press, 163–200.

Simperl, E., Acosta, M. & Flock, F. 2013. Knowledge engineering via human computation. In *Handbook of Human Computation*, Michelucci, P. (ed.). Springer, 131–151.

Simperl, E. & Luczak-Rösch, M. 2014. Collaborative ontology engineering: a survey. *The Knowledge Engineering Review* **29**(1), 101–131.

Simpson, R. J., Povich, M. S., Kendrew, S., Lintott, C. J., Bressert, E., Arvidsson, K., Cyganowski, C., Maddison, S., Schawinski, K. & Sherman, R. 2012. The Milky Way Project first data release: a bubblier galactic disc. *Monthly Notices of the Royal Astronomical Society* **424**(4), 2442–2460.

Siorpaes, K. & Hepp, M. 2008. Games with a purpose for the Semantic Web. *IEEE Intelligent Systems* **23**(3), 50–60.

Smart, P. R., Huynh, T. D., Braines, D. & Shadbolt, N. R. 2010. Dynamic networks and distributed problem-solving. In *Knowledge Systems for Coalition Operations (KSCO'10).*

Smart, P. R. & Madaan, A. 2017. The social scaffolding of machine intelligence. *International Journal On Advances in Intelligent Systems* **10**(3/4), 261–279.

Smart, P. R. & Shadbolt, N. R. 2014. Social machines. In *Encyclopedia of Information Science and Technology*, Khosrow-Pour, M. (ed.). IGI Global, 6855–6862.

Smart, P. R., Simperl, E. & Shadbolt, N. R. 2014. A taxonomic framework for social machines. In *Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society*, Miorandi, D., Maltese, V., Rovatsos, M., Nijholt A. & Stewart J. (eds). Springer, 51–85.

Spiers, H. J., Manley, E., Silva, R., Conroy Dalton, T., Wiener, J. M., Hoolscher, C., Bohbot, V. & Hornberger, M. 2016. Spatial navigation ability assessed in over 1 million people globally. In *Neuroscience*.

Sterelny, K. 2003. *Thought in a Hostile World: The Evolution of Human Cognition*. Blackwell Publishing.

Steyvers, M. & Miller, B. 2015. Cognition and collective intelligence. In Malone, T. W. & Bernstein, M. S. (eds). *Handbook of Collective Intelligence*. MIT Press, 119–137.

Strohmaier, M. & Wagner, C. 2014. Computational social science for the World Wide Web. *IEEE Intelligent Systems* **29**(5), 84–88.

Studer, R., Benjamins, V. R. & Fensel, D. 1998. Knowledge engineering: principles and methods. *Data & Knowledge Engineering* **25**(1), 161–197.

Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A. A., Dietterich, T. & Farnsworth, A. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation* **169**, 31–40.

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D. & Kelling, S. 2009. eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* **142**(10), 2282–2292.

Sutton, J., Harris, C. B., Keil, P. G. & Barnier, A. J. 2010. The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences* **9**(4), 521–560.

Theraulaz, G. & Bonabeau, E. 1999. A brief history of stigmergy. *Artificial Life* **5**(2), 97–116.

Tinati, R., Wang, X., Tiropanis, T. & Hall, W. 2015. Building a real-time Web observatory. *IEEE Internet Computing* **19**(6), 36–45.

Tiropanis, T., Hall, W., Shadbolt, N., De Roure, D., Contractor, N. & Hendler, J. 2013. The Web Science Observatory. *IEEE Intelligent Systems* **28**(2), 100–104.

Tokarchuk, O., Cuel, R. & Zamarian, M. 2012. Analyzing crowd labor and designing incentives for humans in the loop. *IEEE Internet Computing* **16**(5), 45–51.

Turner, J. S. 2011. Termites as models of swarm cognition. *Swarm Intelligence* **5**(1), 19–43.

von Ahn, L. 2006. Games with a purpose. *Computer* **39**(6), 96–98.

von Ahn, L. & Dabbish, L. 2004. Labeling images with a computer game. In *SIGCHI Conference on Human Factors in Computing Systems*, Dykstra-Erickson, E. & Tscheligi, M. (eds). ACM, 319–326.

von Ahn, L. & Dabbish, L. 2008. Designing games with a purpose. *Communications of the ACM* **51**(8), 58–67.

Wang, J., Fischer, D. A., Barclay, T., Boyajian, T. S., Crepp, J. R., Schwamb, M. E., Lintott, C., Jek, K. J., Smith, A. M. & Parrish, M. 2013. Planet Hunters. V. A confirmed Jupiter-size planet in the habitable zone and 42 planet candidates from the Kepler archive data. *The Astrophysical Journal* **776**(1), 1–18.

Watson, D. & Floridi, L. 2018. Crowdsourced science: sociotechnical epistemology in the e-research paradigm. *Synthese* **195**(2), 741–764.

Weld, D. S., Lin, C. H. & Bragg, J. 2015. Artificial intelligence and collective intelligence. In *Handbook of Collective Intelligence*, Malone, T. W. & Bernstein, M. S. (eds). MIT Press, 89–114.