Journal of Social Computing

Volume 1 | Issue 1 Article 1

2020

Social Computing Unhinged

James Evans
University of Chicago, Chicago, IL 60637,and Santa Fe Institute, Santa Fe, NM 87501, USA

Recommended Citation
James Evans. Social Computing Unhinged. Journal of Social Computing 2020, 1(1): 1-13.

JOURNAL OF SOCIAL COMPUTING
ISSN 2688-5255 01/06 pp1–13
Volume 1, Number 1, September 2020
DOI: 10.23919/JSC.2020.0002

Social Computing Unhinged

James Evans*

Abstract: Social computing is ubiquitous and intensifying in the 21st Century. Originally used to reference computational augmentation of social interaction through collaborative filtering, social media, wikis, and crowdsourcing, here I propose to expand the concept to cover the complete dynamic interface between social interaction and computation, including computationally enhanced sociality and social science, socially enhanced computing and computer science, and their increasingly complex combination for mutual enhancement. This recommends that we reimagine Computational Social Science as Social Computing, not merely using computational tools to make sense of the contemporary explosion of social data, but also recognizing societies as emergent computers of more or less collective intelligence, innovation and flourishing. It further proposes we imagine a socially inspired computer science that takes these insights into account as we build machines not merely to substitute for human cognition, but radically complement it. This leads to a vision of social computing as an extreme form of human computer interaction, whereby machines and persons recursively combine to augment one another in generating collective intelligence, enhanced knowledge, and other social goods unattainable without each other. Using the example of science and technology, I illustrate how progress in each of these areas unleash advances in the others and the beneficial relationship between the technology and science of social computing, which reveals limits of sociality and computation, and stimulates our imagination about how they can reach past those limits together.

Key words: social computing; complex systems; computer supported cooperative work; computational social science; artificial intelligence; human computer interaction; human-centered computing

1 A Very Brief History of Social Computing

Regardless of definition, social computing is exploding in prevalence and intensity across the 21st Century world. The most common view of social computing concerns the intersection of human social behavior and computational systems that (re)construct social conventions and social contexts to enable interaction, informed decision-making, and collaboration. Early definitions restricted Social Computing to systems that distributed information indelibly connected to the identity of human contributors. In the 1994 special edition of the Communications of the Association for Computing Machinery, Douglas Schuler emphasized that Social Computing systems "support the gathering, representation, processing, use, and dissemination of information that is distributed across social collectivities such as teams, communities, organizations, and markets. Moreover, the information is not 'anonymous' but is significantly precise because it is linked to people, who are in turn linked to other people"[1]. This preservation of human provenance enabled systems to build on pre-existing relationships and reputations through technologies like email, Bulletin Board Systems (BBS) and multi-user gaming and socializing environments.

The first platform to showcase these principles was PLATO, a computer system designed at the University of Illinois at Urbana Champaign for teaching in 1960 ("Programmed Logic for Automatic Teaching Operations"), which broadened by the early 1970s to

• James Evans is with University of Chicago, Chicago, IL 60637, and Santa Fe Institute, Santa Fe, NM 87501, USA. E-mail: jevans@uchicago.edu.
∗ To whom correspondence should be addressed.
Manuscript received: 2020-09-28; accepted: 2020-10-08

include the first email and personal messaging ("Term-talk"), online news ("NewsReport"), group discussion ("Talkomatic"), newsgroups ("Notesfile"), message boards ("Pad"), and multi-user dungeons or MUDs ("dnd" and "Avatar"), space and battle games (e.g., "Spacewar!", "Empire", etc.)[2] These developments were followed by ARPANET, which came online in 1967 and by the 1970s had developed widespread interactions supported by elaborate standards of network etiquette ("netiquette") that evolved into the Internet. When the World Wide Web was added in the mid-1990s, tens of thousands of BBS systems located in the U.S. migrated to web platforms as weblogs or "blogs".

Social networking emerged with community websites like GeoCities (1994), which included personal profiles, friend lists and school affiliations. Explicit networking sites emerged later, replacing web of contacts sites like SixDegrees (1997) with those built around a model of social circles like Friendster and LinkedIn (2003). Open Diary (1998) and LiveJournal (1999) mixed profiles with user-generated content to become the first social media platforms. Improvements in compression technology enabled a broadening of shareable media from text to links to GIFs and memes, user-generated images (Facebook, 2004) and video (YouTube, 2005)*.

In 1994, the same year that Schuler restricted the definition of Social Computing to information tools tied to people, Ward Cunningham created a knowledge base and associated wiki code he titled "WikiWikiWeb" to facilitate the co-creation of content about design patterns among programmers at his company. Other wikis followed, including Wikipedia in 2001. Users have stable identities on Wikipedia, information-sharing and question-answering sites like Stack Overflow, and sharing economy sites like Uber and AirBnB, but they are effectively anonymous in their creation of content and provision of services. Nevertheless, reputation scores from prior interactions off-set anonymity to provide confidence in those with whom we interact[4]†.

Online services at the turn of the 20th Century began to aggregate completely anonymous information through collaborative filtering to increase the relevance of recommendations provided by web-based services. Collaborative filtering assumes that people who share known qualities likely share others unknown. Two people like *Toy Story*; one likes *Toy Story 2*; maybe the other would like *Toy Story 2* too? Widely used in recommendations about Amazon products, Netflix movies, Spotify songs, and Tinder romantic partners, information about others like the user is leveraged to help them effortlessly follow the most relevant crowd. Google used a similar social principle in its initial website scoring PageRank algorithm. By representing semantically relevant pages in proportion to the links they receive from other pages, Google enabled people new to a domain—outsiders with no local knowledge or connections—to visualize and follow the local crowd[5]. These developments have led to an update in conceptions of Social Computing, broadened to not only include "the use of computational devices to facilitate or augment the social interactions of their users", but also "to evaluate those interactions"—even indirect and anonymous ones—"in an effort to obtain new information"[6].

Beyond definitions, scholarship in social computing has focused on some issues and neglected others. For example, work and life among programmers receives disproportionate attention in social computing because this population has always been overrepresented online. From the origins of social messaging and media[7] to the contemporary widespread production of open source software on collaborative sites like GitHub[8], which enable unparalleled observation of online collaboration, programmers and technologists spend more of their lives online. Moreover, based on widespread connections between information systems, business, and business school scholars, there has been sustained interest in social computing about commercial content and brand management, knowledge sharing and discovery, and peer-to-peer influence, especially as relevant to enterprise[9].

Another subfield that has shone light on aspects of social computing at the intersection of collaboration, coordination and computing is Computer Supported Cooperative Work (CSCW). CSCW launched in the 1980s[10] to explore "how collaborative activities and their coordination can be supported by means of computer systems"[11]—networking, hardware, software, and services. Like social computing, CSCW

*Social networking and media, collaborative creation and tagging were part of a broad shift toward an interactive Web 2.0 that engaged users as creators and not merely consumers, bringing a greater proportion of social life—much of it anonymous—onto the internet[3].

†More anonymous than wiki-contributions or question answers, social bookmarking sites also emerged in the late 1990s, beginning with itList (1996) and WebTagger (1997), but expanding with de.li.cious (2003), which enabled collaborative tagging of websites and other content.

has a design orientation that has enlisted social and behavioral scientists, information and computer scientists, and applied researchers in education, health, and organizations committed to understanding and augmenting productive cooperation with computation.

Another area of social computing research considers the design of systems and formalisms that shape enable efficient online auctions[12], negotiations[13], allocations[14], transactions[15], and collective decision-making[16]. This tracks the explosion of research on mechanism design in economics, such as matching systems (e.g., the Kidney Exchange, medical residency placement) that achieve socially desirable distributional outcomes[17]. This includes attention to the emergence of novel protocols like blockchain, which facilitate automated consummation and persistent documentation of transactions through the interchange of smart contracts [18]. One of the most vibrant areas of social computing research—sometimes termed socially intelligent computing[19]—examines how systems unleash or inhibit collective intelligence among their members. This includes research on the design of crowdsourcing where crowds are assembled online to accomplish a simple or complex task[20]. This research also considers the emergence and intelligence of online communities and markets[21].

New frontiers of social computing are on the horizon, such as consequences of ubiquitous computing access and online social connectivity, richer and more immersive media, the explosion of computational speed promised by quantum computing, and most recently the catalysis of novel social interaction between humans and AI companions, assistants, chatbots, devices and appliances, and between these nonhuman intelligent agents‡.

This vast and growing swath of online activity has heretofore constituted social computing. But I argue that our understanding could be enriched through deeper interchange with related work in the social and computing sciences—not just to unhinge and expand social computing itself, but to broker a deeper compact between social and computing research. Here I propose to expand the concept of social computing to cover the complete dynamic interface between social interaction and computation.

‡ Some of these trends have been encapsulated in the term Web 3.0, coined by reporter John Markoff to include the semantic web, ubiquity, 3D immersiveness, connectivity, and AI[22].

2 What is Computing? What is Sociality?

In their most common usage, computers are machines built to externalize human cognition. Cognition is the mental process of acquiring knowledge and understanding through thought, fueled with the data of sense and experience. Computers were designed for human programmatic instruction to carry out sequences of mathematical and logical operations on their behalf. Charles Babbage proposed to design the first steampunk computer—the difference engine—because human computation of measurements to improve the British Nautical Almanac was so tedious and error prone[23]. Alan Turing, Claude Shannon and others updated computer designs for the digital age and vastly expanded the cognitive target of computation as not merely calculation but general, artificial intelligence[24,25].

If computation is machine cognition, then sociality and communication are human networking. Sociality is the inclination to companion with others; the property of being friendly that initiates and sustains positive social relations. Sociality is a core human inclination[26] and has increasingly been tied to humanity's global dominance over other species through large-scale cooperation[27]. It is notable that the same Babbage who designed the first computer with blueprints of unprecedented detail also designed the efficient breakdown of efficient social relations between human laborers in enterprise[28].

How do the principles underlying computing and sociality fit together? First they can be efficiently co-embedded. In the human realm, efficient economies involve specialization and trade, just as efficient enterprises and governments involve specialized roles linked through networks and hierarchies. In the realm of machines, efficient computation similarly involves networking or distributed processing, which trades space for time. At the extreme end is DNA computing, which exponentially splits and parallelizes arithmetic operations, doubling the pools required to store results with each new RNA replication[29]. In the analysis of data, random forests and deep neural networks are models designed to discover complex, nonlinear associations by fragmenting predictions into smaller operations then reassembling. Conversely, for humans, efficient communication requires the deployment of relevant language[30]. The efficient allocation of goods and services requires the design of relevant institutions[31]. Similarly, for computer systems, efficient networking requires computation to design, a central task of electrical engineering.

In the evolution of human cognition and sociality, which came first? The enlightenment answer was that reason led to collaboration: Social contracts emerged when individuals realized that they could do better together. In this account, cognition led to sociality as lawless and undersocialized agents agreed to be governed together[32]. This contrasts starkly with new thinking[33] and experimental evidence about this relationship. In the *Enigma of Reason*[34], cognitive scientists Hugo Mercier and Dan Sperber demonstrate persuasively that seeming defects in human reason resolve if reason evolved not to produce generalized knowledge, but to facilitate sociality by persuading that someone is a trustworthy member of one's group. In this conception, sociality preceded and drew forth reason, which increased sociality. Regardless of precise priority, sociality is not cognition's afterthought.

I argue for a vision of social computing that considers cognition and communication—computation and networking—between any combination of humans, machines and other agents (e.g., pets, states, corporations), because this approach will allow us a much broader field of comparisons to evaluate and better design those we want. Computation and networking could occur exclusively among machines as in the internet of things. Cognition and communication could occur exclusively among humans through politics, economics and society. Alternatively computation could take place inside humans, and networking through machines as in Web 1.0; or computation could take place inside machines, and networking through humans as in the use of mechanical calculators for the Nautical Almanac.

More realistic configurations involve not only computation and communication, but also sensing and action, which may be performed across complex networks of humans and machines. Consider the social computing involved in a simple supervised machine learning algorithm. Human sensors ingest data and cognitively classify it, then communicate classification to machines that compute a function to match and generalize those classes to new data. In crowdsourcing, a human agent designs a task, communicates it through a computational system to a crowd, as in Amazon Mechanical Turk, whose members use cognition (e.g., "rank the safety of these places"), human sensors (e.g., "draw/take pictures of (un)safe places"), or human actions (e.g., "physically assemble this kit"). In Alex 'Sandy' Pentland's winning entry for DARPA's red balloon challenge, he used a machine platform to contract with human spotters to place the location of 100 balloons DARPA had floated around the United States. Spotters could contract others, who could contract others, recursively, fragmenting the prize between them all[35]. In an open source software community, human coders produce computer code, comment on each other's code, like each other's comments, which likes are aggregated by algorithm to create visible popularity signals.

In summary, social computing could and should include much more than human sociality supported by machine cognition. It traces the possibility of the union of cognition and communication with humans and machines. This recommends we expand the notion of social computing to include computationally enhanced sociality and social science; socially enhanced computing and computer science; and their increasingly complex combination for mutual enhancement.

This expansion reimagines the emerging field Computational Social Science as Social Computing, inviting us to conceive of social systems as complex, collective computers. The results of such a project could, in turn, catalyze a more socially informed computer science, recommending deviations from core research paradigms such as the Artificial (humanoid) Intelligence project to another that seeks a human complementary Alien Intelligence. This rotation and inversion of social and computing reconfigures social computing as an extreme form of human computer interaction where machines and persons recursively combine to augment one another in generating collective intelligence, enhanced knowledge, and other social goods unattainable without each other. In the sections below I explore these possibilities, then conclude with a research example of how social computing could recompute science, but also politics, business and society.

3 Computational Social Science as Social Computing

Before 2000, Computational Social Science referred to computing the consequences of theoretical assumptions in social simulations. This drew upon classic work like Thomas Schelling's *Micromotives and Macrobehavior*[36], which showed how outcomes like neighborhood ethnic segregation could arise from small and rare preferences for homophily that nevertheless passed a critical threshold. In the 1990s this program flourished into efforts that simulated artificial societies using agent-based models[37,38], generalizing the question of whether a given set of social rules was sufficient to generate observed social outcomes in the world.

After 2000, the rise of the web and profusion of social life on it turned computational social science to pursue computer-assisted analysis of the explosion of social data becoming available from blogs, social media, social networking and other traces of digital communication[39,40]. For social scientists the social data revolution includes "high throughput" archives, observatories, surveys, and experiments[41]. High throughput emerged as a desiderata of research at the turn of the 21st Century with experiments in drug discovery, genomics, biology, and chemistry that used robotics, advanced data processing and sensitive detectors to rapidly conduct millions of tests, potentially answering many questions simultaneously[42]. High throughput archives result from the massive digitization of historical materials, such as the Google Books project, a vast collaboration with academic libraries around the world, instigated not to produce free archives, but to teach Google how to read. It has been accompanied by innumerable focused projects, such as the digitization of historical newspapers by content providers like ProQuest, past parliamentary and congressional debates by the U.S. and European governments, and ancestral records by the Church of Jesus Christ of Latter-day Saints. Together, these provide novel views on historical culture, language, politics and kinship.

High throughput human observatories arise from the vast sensor array of mobile phones, social media, check-out scanners, credit card transactions, and fitness apps that provide unprecedented views of human behavior and interaction. High throughput surveys involve the use of crowdsourcing and deployment of simple information tasks at massive scales on digital platforms ranging from Wikipedia to XBox[43]. These often deploy active learning to make them adaptive and reduce sample size, focusing only on questions most relevant to respondents, information about which models are most uncertain, or both[44]. Finally, high throughput digital experiments are flourishing, facilitated by broadband that enables groups, teams and communities to interact with no perceivable latency. Because of the ubiquity of mobile phones, such experiments can be ecologically situated within natural contexts, like farmers whose experimental participation can be directed by real-time protocols[45].

These sources are creating big social data, which creates novel opportunities to study rare but consequential events like viral videos that spark collective attention[46], novel behavior that sets off cascades of copying, or network connections that bridge distant social or cultural communities[47]. For anything that sits in the tail of the frequency distribution, big data is small data and small data is no data. The supply of big social data has generated demand for tools that can turn unstructured digital information into forms ready for analysis. Models and tools have met this demand, many of them invented by research scientists in social media and networking companies that benefit from their analysis to better target ads and make recommendations. Most of these tools use deep neural networks applied to text, images, communication networks and arbitrary tabular information.

Analysis of social data has been associated with a social analytics revolution that now enables the modeling of high-dimensional data with dimension reduction techniques like LASSO ("least absolute shrinkage and selection operator") and approaches for discovering nonlinear interactions through random forests and deep learning. These efforts have led to models that in some cases sufficiently explain almost all of the variation in social outcomes of interest[48], while in other cases they cannot[49]. In successful cases, such models have enabled a new generation of simulations. High performance computing has fueled these advances, just as cloud-based storage has enabled computation over dispersed data, the creation of enclaves for protected data, pipelines that perform analysis as a service, and distributed adaptive surveys and experiments that draw on a common pool of models accessible from everywhere. Finally, artificial intelligence is helping to create hypotheses, and examine simulated agents in the wild[50].

In reimagining computational social science as social computing, we see that computational methods do not merely "plug-in" to pre-established social scientific pipelines, but come with epistemic entailments that alter what social science knows and who knows it. In an era of small data collected from expensive surveys, interviews and observation, social scientists relied on strong models, with many theoretical assumptions, in order to make inferences. With small data you would not mine that data for new insight, but reserve it for testing what you think you know. Big data invites us to weaken our models, ceding more intelligence and creativity to algorithms.

With big data, machines can help us discover hypotheses on some data, and test it on others. This does not mean theory is less important: keeping more theories in play will increase what we learn from data.

Computational social science as social computing would explicitly consider the computational cost or complexity of operations, which encourages a shift from random samples to optimized inference, a movement described by Suchow and Griffiths: "experimental design as algorithm design"[51]. If networking has no perceivable latency, then we can invest each new experimental task or survey question as a function of model uncertainty based on the unfolding stream of prior results, rather than rituals of reliability that spend samples unevenly across questions of more and less importance to our understanding.

Computational social science in the era of social computing also implies a new generative standard for social scientific epistemology, which refurbishes one forged in the era of data-less simulation: "don't trust it if you can't generate it"[38]. With large scale data and increasingly precise models, in some areas we can now develop social simulations sufficient and plausible enough to explain nearly all the phenomenon of interest[48]. This represents a shift in attention from discovering and testing mechanisms to predicting and generating "digital doubles" of complete social phenomena—from necessary but not sufficient explanations to sufficient but not necessary ones. This represents a corollary shift from social science to social technology. The design of a social algorithm that consistently produces desired outcomes may not hinge on accurate explanations of social systems in the wild, but it is highly relevant to the construction of human services on the web. This has likewise expanded concerns beyond causal inference to prediction—from strategies for minimizing model bias to also reducing model variance[52]. Nevertheless, to compute social policy, models will need to not only predict the most likely future, but identify factors that predictably change that future. This has led social scientists to recruit insights from machine learning into their quest for causal explanation[53] as many experts in machine learning have come to see causality as the next great frontier in their field[54] in a quest to yield simulations that effectively redesign society.

This focus on generation, simulation and prediction moves from singular explanations that account for a small percentage of variation in outcomes to combined and competing explanations within the same sufficient model[55]. This retunes social science for the solving of real world problems, but also leads to messier explanations, as one theory may not dominate others. Attention to multiple, simultaneous influences has led to new targets of statistical inference, like the minimization of false discovery rate for a collection of findings, rather than estimates for each individual one[56].

Finally, social computing adds a fundamentally new objective to computational social science. It poses the question, how do social systems compute solutions to their problems and how could they do it better? Work in data-driven animal collective behavior has identified how swarming fish efficiently compute defensive formations[57], just as human voting systems and deliberations are designed to socially compute maximum agreement. Understanding how and when social systems compute well and poorly in the noisy environments of the real world will hold secrets for a more socially informed and responsive computer science.

4 Computer Science as Social Computing

Networking is a central aspect of emerging computing systems, such as the internet of things that links appliances, houses, cars, stoplights and smart contracts to signal and transact with one another. Anticipating the thicket of interactions that result would benefit from generations of social science, especially as translated into the concerns of social computing.

There is rich historical precedent for the use of social computations to inspire machine ones. Oliver Selfrige was inspired by human election systems as inspiration for ensemble methods in machine learning, influencing the efforts to average individual decision trees into a "random forest". Each tree becomes a voter in the system, trained on a distinct but overlapping subset of experiences from data[58].

Seymour Pappert and Marvin Minsky began work on a compositional theory of intelligence they called "Society of Mind" in the early 1970s, which matured into Minsky's 1985 book of the same title[59]. In this view, intelligence emerges from the interaction of diverse and various cognitive agents, not from a singular mechanism. Directly influenced by child psychology, Minsky nevertheless constructed his motivation and guiding metaphor for the project from the division of labor in society.

Recent research on deep learning neural networks has begun to use network methods to perform social analysis of their structure. For example, deep convolutional graph models have recently been used to improve predictions about new ties in online social networks[60]. By incorporating hyperbolic geometry into those models, which captures the hierarchy intrinsic in social and biological networks[61], predictions markedly improved[62]. While it may be no surprise that adding social properties improves models of social data, work by some of the same computer scientists reveals that social properties may be important for arbitrary prediction. They generated a series of neural network models to perform object detection on images from the popular ImageNet and CIFAR databases, then showed that the most successful graph structures had high average path length and medium clustering coefficients, which they argue is similar to the structure of neurological networks[63].This also represents a core feature of social networks that balance weak and strong ties—novelty and cohesion. Other work demonstrates how randomly generated neural network designs perform well on a range of vision tasks, suggesting room for the design of improved networks[64] that could benefit from network properties based on social computing systems like deliberative human communities that achieve desirable outcomes. For example, a recent experiment with human teams showed that those with a higher intelligent quotient in solving problems did not have more intelligent members, but members sensitive to one another's contributions who contributed to team discourse in roughly equivalent turns[65]. These kinds of findings suggest possible regularization strategies for deep learning methods, such as a weight variance penalty, that could improve on heuristic approaches widely used today. I put forward the hypothesis that principles we discover to be consistently useful in social computing systems may also be useful for the design of networked computer systems more broadly.

Another principle that holds widely across social contexts involves the power of cognitive diversity for solving complex, nonroutine problems. In science, technology, design, or any domain in which quality is prized above speed, diverse ensembles rule. The rise of the web has seen a sharp increase in observational[66–70], theoretical and experimental social computing research[71–74] on the "wisdom of crowds" phenomenon, which implies that a collective's aggregate response exceeds the accuracy of its members. This research demonstrates that diverse crowds rule. Crowd wisdom hinges on the independence and diversity of information[66] and approach[75] held by its members. This is manifest in data science competitions like Kaggle where ensemble models (e.g., boosted forests or deep networks) have *always* won and never been bested by a best-fit single model. My own and others' work on collective achievement in science and technology suggests that diversity and independence plays a similarly critical role there, where dense communities slow the speed of advance by collapsing the space of ideas imagined and explored[70,76,77]. The wisdom of diverse crowds suggests the possibility that new diverse weight initialization approaches could improve the optimization of complex, networked models in deep learning. But it also suggests a deeper design challenge for computer science as social computing.

5 AI as Alien Intelligence

In 1955, two alternative and competing visions for AI were ratified in the same year. Young assistant professor John McCarthy from Dartmouth College applied to the Rockefeller Foundation for funding that would support 10 people for an 8 week study of "Artificial Intelligence". The workshop occurred the following summer involving Claude Shannon, Marvin Minsky, Oliver Selfridge, John Holland, Herbert Simon, Allen Newell and a few others in an event many have hailed as foundational for the field of Artificial Intelligence. The project covered topics ranging from neural networks and the theory of computation to national language processing, abstraction and human(oid) creativity. The view of AI that emerged from this meeting draws on the "imitation game" approach to assessing intelligence[25]: Intelligence is to mimic humans who represent the standard of intelligence. Artificial intelligence became even more deeply tied to human intelligence with Arthur Samuel's work to develop a computer checkers player in the late 1950s, and his coining of "machine learning"[78] to reflect algorithms that not only produced outcomes apparently indistinguishable from humans, but which directly learned from human activity (e.g., human checkers games).

And yet with 7 billion humans on earth and growing, is the production of more artificial humanoid intelligences the most interesting intellectual target? Not for computer science as social computing. The wisdom of diverse crowds suggests an alternative vision for AI as Alien Intelligence, not most but least like humans and human groups in order to achieve cognitive diversity for social computing—to help human collaborators think differently, bigger and better. Such a program would draw insight from computational social science as social computing to model how humans, groups and societies collectively think and act, then use novel objective functions like Bayesian Surprise[79], multiple objectives like surprise and support, or oppositionally structured objectives like those in Generative Adversarial Networks (GANs)[80] to grow computational Alien Intelligences that radically complement the humans with whom they collaborate.

The construction of Alien Intelligences (AIs) could enable heightened strategic action, where the concealment of purpose may be critical for competitive performance—where failing the Turing test is the best stealth. This approach could be used to cultivate strategies explicitly designed to interfere with contemporary theories of mind, and evolve as expectations update. Consider the game of "free chess", where competitors may be any combination of humans and machines. The first international tournament in 2005 was won by two amateur players running four off-the-shelf chess simulators on three cheap machines[81]. They likely beat Grand Masters Vladimir Kosyrev and Konstantin Landa running powerful single chess engines not only because their total system was better, but because they did not play an integrated game of chess. They played many such games simultaneously and short-circuited the expert ability to template a response. They were fundamentally unpredictable. Similarly, Alpha Go, Deep Mind's Go-playing robot was initially trained on hundreds of thousands of human games, but Alpha Go Zero was trained only against other naive adversaries like itself, growing what many master Go players described as an unfamiliar and decentering "alien" strategy[82]. Alpha Go Zero achieved this because GO is such a complex game that its training had been reduced to traditional, heuristic human strategies. Simultaneously incorporating data on human strategies, and the computable universe of alternatives could allow a player that intentionally and directly defies human expectation, with much less effort and even in cases where successful games cannot be simulated.

In order to radically, not incrementally, augment human intelligence to face complex challenges, AI as alien intelligence could provoke decision-makers, strategists, creators, researchers and teams of the same with ideas, approaches and analyses that humans and human groups could not have conceived of themselves. Financial, creative and intellectual arbitrage occurs when individuals with information from one market or domain of activity bring it to another where it is not yet present, but valuable. Acts of value-producing cognitive arbitrage can be accelerated with alien intelligences unconstrained by human incentives to over-coordinate or flock together. Consider a recent experiment in which Hirokazu Shirado and Nicholas Christakis staged a color coordination game where players sat in an online network and were each tasked with changing the color of their network node to contrast with those to whom they were connected. When they added robots to central locations in the graph, which did not follow human conventions but exhibited random noise in color choice, this disrupted local coordination but substantially improved global coordination, including game play between humans, leading to increased collective performance[83]. Because humans were most likely to coordinate with one another locally, the design of diverse robots that were worse at coordination, fluctuating with some randomness, best complemented human players and improved collective computation. One could also imagine the converse approach for training machine learning algorithms: strategically placing humans in the loop where they can productively perturb certainty within the network of components in the learning system.

6 Social Computing as Extreme Human Computer Interaction

The construction of alien intelligences that are not most, but least like human agents in order to assemble optimal diversity in human-computer groups defies not only core tenets of Artificial Intelligence, but its twin sister Augmented Intelligence. The same year that John McCarthy assembled his initial workshop on Artificial Intelligence, William Ashby detailed the possibility of "Amplifying Intelligence" in his *Introduction to Cybernetics*[84]. Soon after, Douglas Englebart put forward a program to "Augment Human Intellect"[85], which J.C.R. Licklider described as "Man-Computer Symbiosis"[86]. The vision behind amplified or augmented human intelligence was to accelerate human thinking by reducing friction and allowing unmediated access to information for reference and manipulation. If humanoid robots are the natural embodiment of artificial intelligence; screens, mice, file systems and hypertext are the embodiment of augmented intelligence with new human computer interfaces (eeg helmets, eye-tracking systems, etc.)

emerging to facilitate novel augmentations. These efforts gave rise to the fields of human-computer interaction (HCI) and Human-centered computing (HCC). HCI has historically focused on design principles that increase intuitive familiarity and decrease frictions when using an interface, while HCC has attended to the systems and technological practices that achieve human goals. If we reimagine HCI and HCC as social computing, then augmenting human intellect through seamless interfaces and integrated systems seems incremental and doomed to diminishing marginal returns. HCI and HCC as social computing will incorporate principles like diversity, which promotes collective creativity and intelligence through the creation of cognitive dissonance[87] and destabilizing conflict[88].

Interlocutors who disagree must nevertheless be able to communicate. Furthermore, Sinan Aral's research on communication through social networks reveals a trade-off between diversity and bandwidth[89]. Novel information comes through conversation with those who are different from you, but more information comes through conversation with those who are similar. HCI and HCC as social computing will seek to discover the optimal balance of information, perspective and friction required to compute the innovation or performance we desire.

The ultimate validation of this approach will be the fruitfulness of a research program inspired by the breaching experiments of sociologist and ethnomethodologist Harold Garfinkel. Garfinkel identified the existence of social and cultural norms by violating them and observing the reaction[90]. Research programs of this type will likely lead to more extreme human computer interaction, in which machines and persons recursively combine to provocatively augment one another and generate enhanced knowledge, collective intelligence and social goods unattainable and unimaginable without each other.

7 Social Computing for Science

My own recent work on collective achievement in science and technology suggests some of the synergies that could be achieved from integrating computational social science, socially inspired computer science, and human computer interaction as social computing. In science, dense communities constitute echo chambers that slow the speed of advance by collapsing the space of ideas imagined and explored[70,76]. Findings from connected communities are less likely to reproduce than those from diverse, disconnected ones[77]. The largest and most predictable factor in outsized, disruptive success in discovery or invention is when a small group[91] of scientific or inventive outsiders—people with backgrounds distant from a given field's problem, travel to that distant audience and solve that problem with patterns alien to the receiving field[48]. This could allow us to instrument scientific and technical abduction, where problems identified through the collision of deductive expectations and unexpected inductive findings are resolved with new, alien ideas and patterns that make the surprising unsurprising. With more data and better models than ever before on how scientists and inventors collectively think, we are now uniquely in a position to build alien intelligences with new objectives, which systematically avoid places that humans have overthought in order to identify promising leads that could not have been conceived before.

The use of diverse scientific viewpoints was drawn upon by pioneering information scientist Donald Swanson in his manual approach to "literature-based discovery" called the ABC model of hypothesis generation. If concepts A and B are studied in one literature, and B and C in another, Swanson assumed transitivity to hypothesize that A implies C, then demonstrated that novel A-to-C inferences were likely to be true, although unlikely to be arrived at via other means[92–95]. Through this approach, Swanson hypothesized that fish oil could lessen the symptoms of Raynaud's blood disorder and that magnesium deficits are linked to migraine headaches[96]. This heuristic relies on an implicit understanding of how diverse understandings could be combined into powerful new knowledge. Moreover, Swanson's approach acknowledged, but did not explicitly identify differences in the experiences and views of scientists it arbitraged to generate insights the scientific system would not have naturally discovered. Diversity in exposure to distinct cultural milieu has been shown relevant for scientific advance, with the observation that scientific collaborations from diverse countries receive more citations[97], but also the Foreman thesis in the history of science that national contexts constitute unique views on the world that shape what science can be imagined. Quantum mechanics emerged in the Weimar Republic as part of a philosophical movement against causality[98].

Now, consider a social computing approach that extends far-beyond the approach by Swanson, described above, by directly incorporating the distribution of diverse perspectives. A *Nature* article published in 2019 revealed how embedding chemicals and properties to a vector space from millions of prior research publications can be used to predict 40% of the novel associations more than two decades into the future[99]. These tools are now becoming widely used to generate new hypotheses in the biological, material and physical sciences. A neural embedding-based analysis ignores the patterns of collaboration and diversity underlying how past discoveries were made. Our preliminary experiments to incorporate information about the distribution of authors and their connections between fields dramatically improve the accuracy of novel discovery predictions by more than 100%. As suggested above, adding this information further allows us to make inferences regarding who would likely have been the discoverers, but also forecast promising discoveries that could not be made without a computational alien intelligence system of this kind.

8 Socially Computing Everything

Recomputing science by recursively exploring how scientific cognition and communication could be augmented through machine experimentation and extension recalls the legendary Ship of Theseus. Plutarch, the first century Greek historian, wrote a history of Theseus, the mythical founder of Athens, who traveled with other Athenian youths to Crete. They sought to free Athens, which had become a tributary to King Minos by defeating Minos' Minotaur, a half bull/half man who was to consume the Athenian tributes. With the help of Daedalus, who had created the Minotaur's labyrinth, and Ariadne, Minos' love struck daughter, Theseus killed the Minotaur and escaped with the youth. Theseus' ship returned to liberated Athens, Theseus died and slipped into legend, and the Athenians eventually "took away the old planks as they decayed, putting in new and stronger timber in their places, insomuch that this ship became a standing example among the philosophers, for the logical question of things that grow; one side holding that the ship remained the same, and the other contending that it was not the same"[100].

Like the decaying ship of Theseus, machines have been constructed to recompute human sensation, cognition and communication with increasing sensitivity, capacity and expressiveness. As more and more of Theseus' ship was replaced by newer, better planks, when did it cease to be Theseus' ship? As more and more scientific components are augmented by machines, when does it cease to be science? As more and more political processes are bolstered with machines, when does it cease to be a polity? As more and more business and commerce utilize machines, when does it cease to be an economy? It doesn't. But neither is it the same science, policy or economy as before. Social computing is the new ship of Theseus, inviting us to redesign legendary institutions for greater collective intelligence, flourishing, and innovation.

As a field, social computing will be unhinged and expanded if we consider the complete dynamic interface between social interaction and computation. It will be enriched by engaging more deeply with computational social science, which will help to provoke and explore the possibility of a socially informed computer science. But the greatest contribution may come from convening other fields, from cognitive science to communication to computing, together in conversation to design systems that generate social goods impossible to conceive alone. With the power to create goods comes the capacity to unleash bads, Matrix-like scenarios with humans trapped in webs of machines that update nightmares of child labor from the industrial revolution for the age of AI. And this is why social computing must also convene the critical theorists, philosophers, and artists, to imagine and warn of danger.

Social computing will be most productive as an impure science—in tension with technology. Science identifies stable patterns in the social and natural worlds, but the view that technology simply flows from scientific insight[101] ignores the vast history of successful, but inexplicable inventions and curiosities that provoke science[102]. These include the Bessemer Process to remove carbon in making steel or Edison's etheric force, later rediscovered as the wireless transmission of electromagnetic radiation[103]. Successful techniques and novelties suggest scientific regularities to be discovered and harnessed. By exploring the limits of society and computation—investigating with science; insulting with technology—social computing can unhinge our imaginations and focus our effort on how they can reach past those limits together. It is in the hope of this chaotic conversation, partly beyond human comprehension, certainly at risk of peril, that we launch the journal of social computing.

References

[1] D. Schuler, Social computing, *Commun. ACM*, vol. 37, no. 1, pp. 28–29, 1994.

[2] B. Dear, *The Friendly Orange Glow: The Untold Story of the Rise of Cyberculture*. Knopf Doubleday Publishing Group, 2017.

[3] G. Cormode and B. Krishnamurthy, Key differences between Web 1.0 and Web 2.0, *First Monday*, vol. 13, no. 6, 2008.

[4] B. Abrahao, P. Parigi, A. Gupta, and K. S. Cook, Reputation offsets trust judgments based on social biases among Airbnb users, *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 37, pp. 9848–9853, 2017.

[5] I. Rogers, The Google pagerank algorithm and how it works, https://cs.wmich.edu/gupta/teaching/cs3310/lectureNotes_cs3310/Pagerank%20Explained%20Correctly%20with%20Examples_www.cs.princeton.edu_~chazelle_courses_BIB_pagerank.pdf, 2002.

[6] N. Hemmatazad, Social computing, in *Encyclopedia of Information Science and Technology*, 3rd ed. M. Khosrow-Pour, ed. Hershey, PA, USA: Information Science Reference, 2015, pp. 6754–6761.

[7] F. Turner, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago, IL, USA: University of Chicago Press, 2010.

[8] M. Biazzini and B. Baudry, May the fork be with you: Novel metrics to analyze collaboration on GitHub, in *Proceedings of the 5th International Workshop on Emerging Trends in Software Metrics*, Hyderabad, India, 2014, pp. 37–43.

[9] Y. B. Li and K. D. Joshi, The state of social computing research: A literature review and synthesis using the latent semantic analysis approach, https://aisel.aisnet.org/amcis2012/proceedings/DecisionSupport/13/, 2012.

[10] J. Grudin, Computer-supported cooperative work: History and focus, *Computer*, vol. 27, no. 5, pp. 19–26, 1994.

[11] P. H. Carstensen and K. Schmidt, Computer supported cooperative work: New challenges to systems design (1999), http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.43.5157, 2020.

[12] R. B. Myerson, Optimal auction design, *Mathematics of OR*, vol. 6, no. 1, pp. 58–73, 1981.

[13] I. Roussaki, M. Louta, and L. Pechlivanos, An efficient negotiation model for the next generation electronic marketplace, in *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No.04CH37521)*, Dubrovnik, Croatia, 2004, pp. 615–618.

[14] D. Gale and L. S. Shapley, College admissions and the stability of marriage, *null*, vol. 69, no. 1, pp. 9–15, 1962.

[15] S. T. Kirsch, Secure, convenient and efficient system and method of performing trans-internet purchase transactions, US Patent 5963915, Oct. 5, 1999.

[16] S. P. Lalley and E. G. Weyl, Quadratic voting: How mechanism design can radicalize democracy, *AEA Papers and Proceedings*, vol. 108, pp. 33–37, 2018.

[17] A. E. Roth and E. Peranson, The redesign of the matching market for American physicians: Some engineering aspects of economic design, *Am. Econ. Rev.*, vol. 89, no. 4, pp. 748–780, 1999.

[18] M. Iansiti and K. R. Lakhani, The truth about blockchain, https://hbr.org/2017/01/the-truth-about-blockchain, 2017.

[19] A. Shaw, Teaching socially intelligent computing principles in introductory computer science courses, in *Proceedings of the 50th Annual Southeast Regional Conference*, Tuscaloosa, Alabama, 2012.

[20] D. Retelny, S. Robaszkiewicz, A. To, W. S. Lasecki, J. Patel, N. Rahmati, T. Doshi, M. A. Valentine, and M. Bernstein, Expert crowdsourcing with flash teams, in *Proceedings of the 27th annual ACM symposium on User interface software and technology*, Honolulu, HI, USA, 2014, pp. 75–85.

[21] D. Adjodah, Y. Leng, S. K. Chong, P. Krafft, and A. Pentland, Social Bayesian learning in the wisdom of the crowd, arXiv: 1712.09960v1, 2017.

[22] K. Patel, Incremental journey for World Wide Web: introduced with Web 1.0 to recent Web 5.0—A survey paper, *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 10, 2013.

[23] W. J. Ashworth, The calculating eye: Baily, Herschel, Babbage and the business of astronomy, *Br. J. Hist. Sci.*, vol. 27, no. 4, pp. 409–441, 1994.

[24] C. E. Shannon, A symbolic analysis of relay and switching circuits, *Univ. Polytehnica Bucharest Sci. Bull. Ser. C: Electr. Eng.*, vol. 57, no. 12, pp. 713–723, 1938.

[25] A. M. Turing, Computing machinery and intelligence, *Mind*, doi: 10.1093/mind/lix.236.433.

[26] D. Brooks, *The Social Animal*: *The Hidden Sources of Love, Character, and Achievement*. Random House Trade Paperbacks, 2012.

[27] Y. N. Harari, *Sapiens*: *A brief history of humankind*. Random House, 2014.

[28] C. Babbage, On the economy of machinery and manufactures, *null*, vol. 1, no. 3, pp. 208–213, 1832.

[29] A. Currin, K. Korovin, M. Ababi, and K. Poper , Computing exponentially faster: Implementing a non-deterministic universal Turing machine using DNA, *J. R. Soc. Interface*, vol. 14, no. 128, p. 20160990, 2017.

[30] G. K. Zipf, Human behavior and the principle of least effort, https://doi.org/10.1525/aa.1950.52.2.02a00290, 1950.

[31] D. C. North, Institutions and economic growth: An historical introduction, *World Dev.*, vol. 17, no. 9, pp. 1319–1332, 1989.

[32] T. Hobbes, Leviathan (1651), *Glasgow 1974*, 1980.

[33] M. Granovetter, Economic action and social structure: The problem of embeddedness, *Am. J. Sociol.*, vol. 91, no. 3, pp. 481–510, 1985.

[34] H. Mercier and D. Sperber, *The Enigma of Reason*. Harvard University Press, 2017.

[35] J. C. Tang, M. Cebrian, N. A. Giacobe, H.-W. Kim, T. Kim, and D. Wickert, Reflecting on the DARPA Red Balloon Challenge, *Commun. ACM*, vol. 54, no. 4, pp. 78–85, 2011.

[36] T. Connolly, Micromotives and macrobehavior, *Administrative Science Quarterly*, vol. 24, no. 3. pp. 500–504, 1979, doi: 10.2307/2989930.

[37] J. M. Epstein and R. Axtell, *Growing Artificial Societies*: *Social Science from the Bottom Up*. Washington, DC, USA: Brookings Institution Press, 1996.

[38] J. M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ, USA: Princeton University Press, 2006.

[39] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al., Life in the network: The coming age of computational social science, *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[40] D. M. J. Lazer, A. Pentland, D. J. Watts, S. Aral, S. Athey, N. Contractor, D. Freelon, S. Gonzalez-Bailon, G. King, H. Margetts, et al., Computational social science: Obstacles and opportunities, *Science*, vol. 369, no. 6507, pp. 1060–1062, 2020.

[41] J. Evans and J. G. Foster, Computation and the sociological imagination, *Contexts*, vol. 18, no. 4, pp. 10–15, 2019.

[42] J. Inglese and D. S. Auld, High Throughput Screening (HTS) techniques: Applications in chemical biology, in *Wiley Encyclopedia of Chemical Biology*, T. P. Begley, ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2007, p. 1.

[43] W. Wang, D. Rothschild, S. Goel, and A. Gelman, Forecasting elections with non-representative polls, *Int. J. Forecast.*, vol. 31, no. 3, pp. 980–991, 2015.

[44] N. Sengupta, N. Srebro, and J. Evans, Simple surveys: Response retrieval inspired by recommendation systems, *Social Science Computer Review*, doi: 10.1177/0894439319848374.

[45] G. Grossman and D. Baldassarri, The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in Uganda, *Am. J. Pol. Sci.*, vol. 56, no. 4, pp. 964–985, 2012.

[46] S. Goel, A. Anderson, J. Hofman, and D. J. Watts, The structural virality of online diffusion, *Manag. Sci.*, vol. 62, no. 1, pp. 180–196, 2016.

[47] P. S. Park, J. E. Blumenstock, and M. W. Macy, The strength of long-range ties in population-scale social networks, *Science*, vol. 362, no. 6421, pp. 1410–1413, 2018.

[48] F. Shi and J. Evans, Science and technology advance through surprise, arXiv: 1910.09370, 2019.

[49] M. J. Salganik, I. Lundberg, A. T. Kindel, C. E. Ahearn, K. Al-Ghoneim, A. Almaatouq, D. M. Altschul, J. E. Brand, N. B. Carnegie, R. J. Compton, et al., Measuring the predictability of life outcomes with a scientific mass collaboration, *Proc. Natl. Acad. Sci. USA.*, vol. 117, no. 15, pp. 8398–8403, 2020.

[50] P. M. Krafft, M. Macy, and A. Pentland, Bots as virtual confederates: Design and ethics, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, Portland, OR, USA, 2017, pp. 183–190.

[51] J. W. Suchow and T. L. Griffiths, Rethinking experiment design as algorithm design, in *Proc. 30th Conf. Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 2016, pp. 1–8.

[52] J. Kleinberg, J. Ludwig, S. Mullainathan, and Z. Obermeyer, Prediction policy problems, *American Economic Review*, vol. 105, no. 5, pp. 491–495, 2015.

[53] S. Athey and G. Imbens, Recursive partitioning for heterogeneous causal effects, *Proc. Natl. Acad. Sci. USA*, vol. 113, no. 27, pp. 7353–7360, 2016.

[54] J. Pearl, The seven tools of causal inference, with reflections on machine learning, *Commun. ACM*, vol. 62, no. 3, pp. 54–60, 2019.

[55] D. J. Watts, Should social science be more solution-oriented? *Nature Human Behaviour*, vol. 1, no. 1, p. 0015, 2017.

[56] Y. Hochberg and Y. Benjamini, More powerful procedures for multiple significance testing, *Statistics in Medicine*, vol. 9, no. 7. pp. 811–818, 1990.

[57] A. M. Hein, S. B. Rosenthal, G. I. Hagstrom, A. Berdahl, C. J. Torney, and I. D. Couzin, The evolution of distributed sensing and collective computation in animal populations, *Elife*, vol. 4, p. e10955, 2015.

[58] M. L. Jones, How we became instrumentalists (again) data positivism since World War II, *Hist. Stud. Nat. Sci.*, vol. 48, no. 5, pp. 673–684, 2018.

[59] M. Marvin, *The Society of Mind*. New York, NY, USA: Simon and Shusier, 1985.

[60] M. Niepert, M. Ahmed, and K. Kutzkov, Learning convolutional neural networks for graphs, in *International Conference on Machine Learning*, 2016, pp. 2014–2023.

[61] F. Papadopoulos, M. Kitsak, M. Á Serrano, M. Boguñá, and D. Krioukov, Popularity versus similarity in growing networks, *Nature*, vol. 489, no. 7417, pp. 537–540, 2012.

[62] I. Chami, R. Ying, C. Ré and J. Leskovec, Hyperbolic graph convolutional neural networks, *Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 4869–4880, 2019.

[63] J. X. You, J. Leskovec, K. M. He, and S. N. Xie, Graph structure of neural networks, arXiv: 2007.06559, 2020.

[64] S. N. Xie, A. Kirillov, R. Girshick, and K. M. He, Exploring randomly wired neural networks for image recognition, in *Proc. 2019 IEEE/CVF Int. Conf. Computer Vision*, Seoul, South Korea, 2019, doi: 10.1109/iccv.2019.00137.

[65] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone, Evidence for a collective intelligence factor in the performance of human groups, *Science*, vol. 330, no. 6004. pp. 686–688, 2010, doi: 10.1126/science.1193147.

[66] J. Surowiecki, *The Wisdom of Crowds*: *Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business*, *Economies*, *Societies and Nations*. New York, NY, USA: Doubleday, 2004.

[67] A. Kittur and R. E. Kraut, Harnessing the wisdom of crowds in wikipedia: Quality through coordination, in *Proc. 2008 ACM Conf. Computer Supported Cooperative Work*, San Diego, CA, USA, 2008, pp. 37–46.

[68] J. Lorenz, H. Rauhut, F. Schweitzer, and D. Helbing, How social influence can undermine the wisdom of crowd effect, *Proc. Natl. Acad. Sci. USA*, vol. 108, no. 22, pp. 9020–9025, 2011.

[69] H. L. Chen, P. De, Y. Hu, and B. H. Hwang, Wisdom of crowds: The value of stock opinions transmitted through social media, *The Review of Financial Studies*, vol. 27, no. 5, pp. 1367–1403, 2014.

[70] F. Shi, M. Teplitskiy, E. Duede, and J. A. Evans, The wisdom of polarized crowds, *Nat. Hum. Behav.*, vol. 3, no. 4, pp. 329–336, 2019.

[71] A. E. Mannes, R. P. Larrick, and J. B. Soll, The social psychology of the wisdom of crowds, in *Frontiers of Social Psychology. Social Judgment and Decision Making*, J. I. Krueger, ed. New York, NY, USA: Psychology Press, 2012.

[72] J. Becker, D. Brackbill, and D. Centola, Network dynamics of social influence in the wisdom of crowds, *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 26, pp. E5070–E5076, 2017.

[73] J. Becker, E. Porter, and D. Centola, The wisdom of partisan crowds, *Proc. Natl. Acad. Sci. USA*, vol. 116, no. 22. pp. 10 717–10 722, 2019.

[74] B. Golub and M. O. Jackson, Naïve learning in social networks and the wisdom of crowds, *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.

[75] S. E. Page, *The Diversity Bonus*: *How Great Teams Pay Off in the Knowledge Economy*. 2nd ed. Princeton, NJ, USA: Princeton University Press, 2019.

[76] A. Rzhetsky, J. G. Foster, I. T. Foster, and J. A. Evans, Choosing experiments to accelerate collective discovery, *Proc. Natl. Acad. Sci. USA*, vol. 112, no. 47, pp. 14569–14574, 2015.

[77] V. Danchev, A. Rzhetsky, and J. A. Evans, Meta-research: centralized scientific communities are less likely to generate replicable results, *eLife*, vol. 8, pp. e43094, 2019.

[78] A. L. Samuel, Some studies in machine learning using the game of checkers, *IBM J. Res. Dev.*, vol. 3, no. 3. pp. 210–229, 1959.

[79] L. Itti and P. Baldi, Bayesian surprise attracts human attention, *Vision Res.*, vol. 49, no. 10, pp. 1295–1306, 2009.

[80] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in *Proc. 27th Int. Conf. Neural Information Processing Systems*, Montreal, Canada, 2014, pp. 2672–2680.

[81] G. Kasparov, The chess master and the computer, *New York Rev. Books*, vol. 57, no. 2, pp. 16–19, 2010.

[82] D. Chan, The AI that has nothing to learn from humans, *Atlantic*, vol. 20, 2017.

[83] H. Shirado and N. A. Christakis, Locally noisy autonomous agents improve global human coordination in network experiments, *Nature*, vol. 545, no. 7654, pp. 370–374, 2017.

[84] W. R. Ashby, *An Introduction to Cybernetics*. New York, NY, USA: J. Wiley, 1956.

[85] D. E. Englebart, Augmenting human intellect, *originally published in* "*The Augmentation Papers*", *Bootstrap Institute*, pp. 64–90, 1962.

[86] J. C. R. Licklider, Man-computer symbiosis, *IRE Transactions on Human Factors in Electronics*, vol. HFE-1, no. 1, pp. 4–11, 1960.

[87] D. Stark, *The Sense of Dissonance*: *Accounts of Worth in Economic Life*. Princeton, NJ, USA: Princeton University Press, 2011.

[88] R. I. Sutton, *Weird Ideas That Work*: *11 1/2 Practices for Promoting*, *Managing*, *and Sustaining Innovation*. New York, NY, USA: Free Press, 2002.

[89] S. Aral and M. Van Alstyne, The diversity-bandwidth trade-off, *Am.J. Sociol.*, vol. 117, no. 1, pp. 90–171, 2011.

[90] H. Garfinkel, *Studies in Ethnomethodology*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1967.

[91] L. F. Wu, D. S. Wang, and J. A. Evans, Large teams develop and small teams disrupt science and technology, *Nature*, vol, 566, no. 7744, pp. 378–382, 2019.

[92] D. R. Swanson, Fish oil, Raynaud's syndrome, and undiscovered public knowledge, *Perspect. Biol. Med.*, vol. 30, no. 1, pp. 7–18, 1986.

[93] D. R. Swanson, Medical literature as a potential source of new knowledge, *Bull. Med. Libr. Assoc.*, vol. 78, no. 1, pp. 29–37, 1990.

[94] D. R. Swanson and N. R. Smalheiser, An interactive system for finding complementary literatures: A stimulus to scientific discovery, *Artif. Intell.*, vol. 91, no. 2, pp. 183–203, 1997.

[95] M. Weeber, H. Klein, L. T. W. De Jong-Van Den Berg, and R. Vos, Using concepts in literature-based discovery: Simulating Swanson's Raynaud-fish oil and migraine-magnesium discoveries, *J. Am. Soc. Inf. Sci. Technol.*, vol. 52, no. 7, pp. 548–557, 2001.

[96] J. Evans and A. Rzhetsky, Machine science, *Science*, vol. 329, no. 5990, pp. 399–400, 2010.

[97] R. B. Freeman and W. Huang, Collaboration: Strength in diversity, *Nature*, vol. 513, no. 7518, p. 305, 2014.

[98] P. Forman, Weimar culture, causality, and quantum theory, 1918–1927: Adaptation by German physicists and mathematicians to a hostile intellectual environment, *Hist. Stud. Phys. Sci.*, vol. 3, pp. 1–115, 1971.

[99] V. Tshitoyan, J. Dagdelen, L. Weston, A. Dunn, Z. Q. Rong, O. Kononova, K. A. Persson, G. Ceder, and A. Jain, Unsupervised word embeddings capture latent knowledge from materials science literature, *Nature*, vol. 571, no. 7763, pp. 95–98, 2019.

[100] Plutarch, *Life of Theseus*. Independently Published, 2019.

[101] V. Bush, Science the endless frontier: A report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development, United States Government Printing Office, Washington DC, USA, 1945.

[102] N. Rosenberg and R. Nathan, *Exploring the Black Box*: *Technology, Economics, and History*. Cambridge, UK: Cambridge University Press, 1994.

[103] I. Wills, Edison and science: A curious result, *Stud. Hist. Philos. Sci. B Stud. Hist. Philos. Modern Phys.*, vol. 40, no. 2, pp. 157–166, 2009.

James Evans is Professor of Sociology, Faculty Director of Computational Social Science and Director of Knowledge Lab at the University of Chicago and the Santa Fe Institute. His research uses large-scale data, machine learning and generative models to understand how collectives think and what they know, with a special focus on innovation in science, technology, ideology, and culture.