Conceptualization and Framework of Hybrid Intelligence Systems

Nikhil Prakash
nikhil07prakash@gmail.com

Kory W. Mathewson
korymath@gmail.com

Abstract

As artificial intelligence (AI) systems are getting ubiquitous within our society, issues related to its fairness, accountability, and transparency are increasing rapidly. As a result, researchers are integrating humans with AI systems to build robust and reliable hybrid intelligence systems. However, a proper conceptualization of these systems does not underpin this rapid growth. This article provides a precise definition of hybrid intelligence systems as well as explains its relation with other similar concepts through our proposed framework and examples from contemporary literature. The framework breakdowns the relationship between a human and a machine in terms of the degree of coupling and the directive authority of each party. Finally, we argue that all AI systems are hybrid intelligence systems, so human factors need to be examined at every stage of such systems' lifecycle.

1 Introduction

Artificial intelligence (AI) systems are moving out of the research laboratories into the real world. Systems are deployed in critical sectors such as medicine, judiciary, and transportation, as a result, its limitations are becoming profound and leading to disastrous consequences.[1] To overcome these shortcomings, researchers have begun to explore various ways in which humans can be integrated with AI systems. This recent surge has led to the development of multiple hybrid intelligence systems (Huang et al. 2018, Cranshaw et al. 2017, Lasecki et al. 2017). In a survey on the applications of crowdsourcing for machine learning research, Vaughan 2017 compiled a list of hybrid intelligence systems, from both the research community and industry, to explain the diversity in their use cases.

Although there has been a great deal of interest in the application of hybrid intelligence systems, little work has focused on its conceptual understanding, especially in relation to other similar concepts such as *interactive machine learning*, *AI-infused systems*, and *mixed-initiative interactions*. This lack of conceptualization and diversity in terminology not only impedes the progress of the field but also makes it difficult for AI practitioners to make efficient design decisions.

In this work, we propose a precise definition of hybrid intelligence systems and describe its proper interpretation using examples. To explain its relation with other similar concepts and study new upcoming hybrid intelligence systems in the context of older ones, we propose an intuitive framework and report its usage by explaining a few well-known intelligent systems. Finally, we argue, based on our proposed definition, that all AI systems are hybrid intelligence systems and discuss the importance of recognizing human factors at every stage of such systems' lifecycle.

2 Definition of Hybrid Intelligence Systems

Before providing our definition of hybrid intelligence systems, we first explain two other types of intelligent systems, *Human Computation* and *Self-Sufficient Artificial Intelligence* systems, and their

[1] Uber's self-driving operator charged over fatal crash (https://bbc.in/36knk3F)

Preprint. Under review.

relation with hybrid intelligence systems. The reason behind this is two-fold: it will establish proper boundaries and help to provide a broader yet concrete definition of hybrid intelligence systems. In this article, we use the definition of intelligence proposed by Albus 1991, "*as the ability of a system to act appropriately in an uncertain environment, where appropriate action is that which increases the probability of success, and success is the achievement of behavioral subgoals that support the system's ultimate goal*."

2.1 Human Computation

Although the term human computation has been used in multiple disciplines like philosophy and psychology for a very long time, its use in computer science is relatively recent (Quinn and Bederson 2011). According to Quinn and Bederson 2011, von Ahn's 2005 dissertation titled "Human Computation" was pivotal in instigating its usage in modern computer science literature, which defines it as "*a paradigm for utilizing human processing power to solve problems that computers cannot yet solve*." Consequently, the systems which solely utilize human computation, during its entire lifecycle, to complete its tasks are defined as human computation systems.

While in such systems, tasks are completed entirely through human intelligence, machines or software without intelligence could be utilized to assist humans. For instance, *Heteroglossia* is a crowd-powered system that assists story writers by eliciting novel story ideas from the crowd workers (Huang et al. 2020). The system utilizes software applications such as Google Docs's comment functionality for streamline interaction, but it does not involve any AI in supporting the user. Hence, it is a human computation system. Similarly, Mechanical Turk[2] can also be considered as a human computation system because the actual task of playing chess was performed by a human, and the machinery used to assist the human did not possess any intelligence. However, if there was any intelligence-possessing functionality in the machinery, then it would be regarded as a hybrid intelligence system.

2.2 Self-Sufficient Artificial Intelligence

Artificial Intelligence (AI) refers to the intelligence possessed by machines, thus sometimes also known as machine intelligence. Unlike humans, machines acquire intelligence through algorithmic techniques inspired from domains like statistics, mathematical optimization, cognitive science, and fueled by computer processing power and a large amount of data.

We define Self-Sufficient Artificial Intelligence systems as a type of intelligent system that does not utilize human computation at any point in its life cycle, i.e. human assistance is needed neither during its development nor when it is operational. The definition makes it evident that such intelligent systems could only be developed by either other self-sufficient artificial intelligence or hybrid intelligence systems. A key point to note here is that the self-sufficient artificial intelligence system created by a hybrid intelligence system is not considered a hybrid intelligence system because there is no direct involvement of human intelligence, and as a result, it will have inherently different characteristics. For instance, we don't have to consider human factors for the designing and development of such systems. Although we have not been able to find any such system, it is conceivable that large and complex hybrid intelligence systems would be able to create other intelligent systems, at least simpler than itself, without any direct human assistance in the near future, especially considering the recent developments in the field of Evolutionary Computation (Stanley et al. 2019, Miikkulainen 2020, Liang et al. 2019). While in automatic machine learning (AutoML) an intelligent system creates another one, it still requires data generated through human intelligence (He et al. 2020). Therefore, AutoML systems are not self-sufficient artificial intelligence systems.

2.3 Hybrid Intelligence

Hybrid intelligence bridges the enormous gap between human computation and self-sufficient artificial intelligence and is defined as the paradigm which utilizes both human and machine intelligence to solve problems. The systems utilizing hybrid intelligence at any point in their life cycle are called hybrid intelligence systems. Except for human computation and self-sufficient artificial intelligence systems, the proposed concept of hybrid intelligence systems forms the superset of all other kinds of

[2]https://en.wikipedia.org/wiki/The_Turk

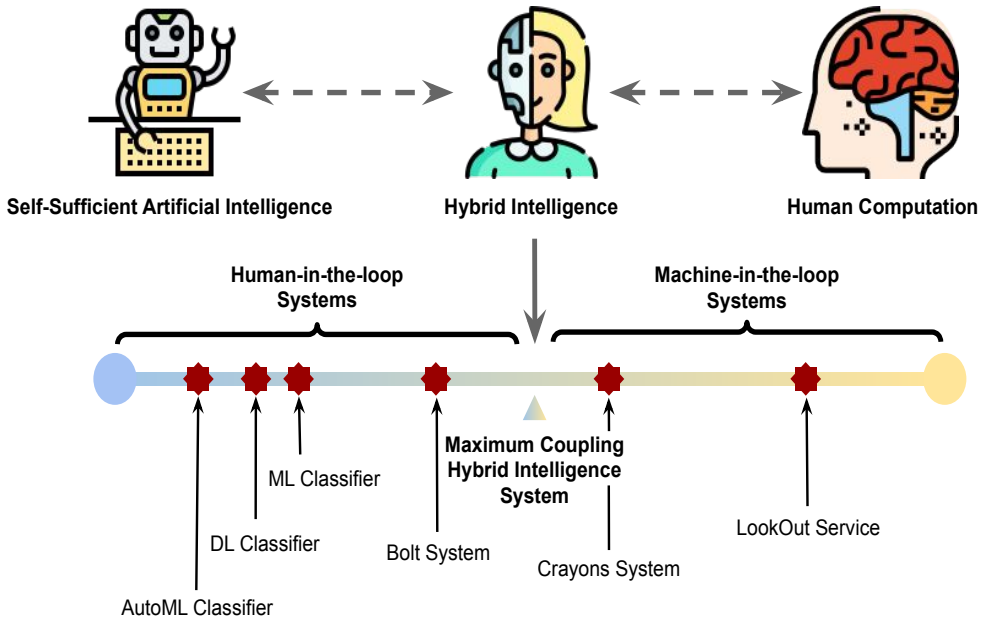

Figure 1: The continuum of hybrid intelligence systems with machine-in-the-loop and human-in-the-loop systems halves. Individual points on the continuum represent distinct hybrid intelligence systems with varying level of coupling between human and machine intelligence.

intelligent systems. The following set of intelligent systems is not intended to be exhaustive, but to provide an idea of how the proposed definition should be interpreted.

Interactive Machine Learning systems Intelligent systems that are developed through rapid, focused, and incremental human interactions are considered as interactive machine learning systems (Amershi et al. 2014). There is a direct interaction between human and machine intelligence during the training phase in these systems, thus making them a special kind of hybrid intelligence system.

Machine / Deep Learning systems Although there is no direct interaction between human and machine intelligence in these systems, a considerable amount of human computation is still utilized at various stages such as data generation, feature extraction, result interpretation, and debugging. Thus, these systems also form a subset of the hybrid intelligence system.

AI-infused systems According to Amershi et al. 2019, systems that have features harnessing AI capabilities that are directly exposed to the end-user are referred to as AI-infused systems. In these intelligent systems, human computation is required not only during its development stages but also during its operations. Hence, these are a special kind of hybrid intelligence system as well.

3 Framework

The definition of hybrid intelligence systems provided in the previous section is broad and includes many different types of intelligent systems. As a result, it makes it difficult for AI researchers and practitioners to study and analyze various subsets of hybrid intelligence systems independently that share a number of common characteristics. Such targeted investigation could fast-track the progress of the field. Therefore, in this section, we explain our intuitive yet robust framework to overcome this challenge.

As shown in the figure 1, the domain of hybrid intelligence systems is conceptualized as a continuum and divided into two halves, representing *human-in-the-loop* and *machine-in-the-loop* systems.

Machine-in-the-loop systems These are the systems that utilize AI, to assist humans in solving their tasks. In this type of systems, the dominant component is human, and the minor component is AI. The endpoint on this half represents a hybrid intelligence system that has minimum AI and maximum human intelligence component.

Human-in-the-loop systems These are the systems in which human computation is utilized for assisting AI at any point in its life cycle. In this type of systems, the dominant component is AI, and the minor component is human. The endpoint on this half represents a hybrid intelligence system that has minimum human intelligence and maximum AI component.

The dimensions used to position hybrid intelligence systems on the continuum are:

1. The coupling between human and machine intelligence. By coupling, we mean the level of integration between both the intelligence. It represents the distance from the center of the continuum.
2. Who is playing the role of the dominant component? It represents the direction.

The center of the continuum represents the highest level of coupling between human and machine intelligence. Systems at the center have an equal contribution of human and machine intelligence, and both are highly dependent on each other. As the coupling decreases the position of the system moves away from the center and the direction is determined by the dominant component. If the dominant component is human, then the system will shift towards machine-in-the-loop systems half and human-in-the-loop systems half otherwise.

The following list of hybrid intelligence systems is not intended to be exhaustive, but to provide an idea of how the proposed framework should be utilized.

LookOut Service It is integrated with Microsoft Outlook and employs an intelligent agent to assist its users in reviewing their calendars and composing appointments when a new email is brought to focus (Horvitz 1999). This service is considered a hybrid intelligence system because it uses both machine intelligence (intelligent agent) and human intelligence (end-user) to complete the task of composing appointments. The dominant component is human as it makes the final decision using the suggestion made by the agent. The interdependence between both the intelligence is relatively low, thus making it a loosely coupled hybrid intelligence system. Appropriating both the dimensions, it is positioned on the machine-in-the-loop systems half, as shown in figure 1.

Crayons System Fails and Olsen 2003 proposed an interactive machine learning system to train a pixel classifier, called Crayons system. People iteratively provided input to the system through brushstrokes by delineating the foreground and background of the image that was utilized to train the model. Based on the user studies conducted to evaluate the system, the authors demonstrated that human input was influenced by the machine output. Thus, making both the involved intelligence in the system highly coupled. Although in tightly coupled hybrid intelligence systems, the boundary between the dominant and minor component becomes uncertain, we argue that in this system, human intelligence is the dominant component as it has control over the training process. Hence, this system is positioned near the center of the continuum on the machine-in-the-loop systems half, as shown in figure 1.

Bolt System Lundgard et al. 2018 introduced the idea of *instantaneous crowdsourcing systems*: intelligent systems that present human computation at machine-level speed, typically in the order of milliseconds. They proposed a hybrid intelligence workflow called the *look-ahead approach* to develop such systems, which is based on the idea of prefetching crowd workers' response to the probable future states of the system. To demonstrate its efficacy, they built a hybrid intelligence system called the *Bolt* system, modeled on the Markov Decision Process to predict the upcoming states. For each predicted state, the system queried and fetched crowd worker's responses. When the system moves to any of these states, the prefetched responses are used to provide instantaneous actions. In this system, both intelligence contribute significantly to determine its output, thus coupling between intelligence is high. Further, we argue that AI is playing the role of the dominant component as it controls which states need to be annotated and in what order by the crowd workers. Hence, the system is positioned near the center of the continuum on the human-in-the-loop half.

Cat Image Classifier In a typical machine learning (ML) workflow for developing a cat image classifier, human intelligence is required at multiple stages like data collection, pre-processing, model design, and debugging, which makes it a hybrid intelligence system. In these systems, humans provide their assistance to develop the machine learning model, which performs the actual classification task. Hence, AI is considered as the dominant component, and the coupling between both intelligence is relatively low. Consequently, this system is positioned on the human-in-the-loop systems half on the continuum. Feature engineering is excluded in the deep learning (DL) workflow, which reduces the involvement of human intelligence, decreasing the integration between both intelligence. With the advent of AutoML, the requirement for human intelligence in the workflow has been significantly reduced, further decreasing the coupling between both the intelligence. This gradual reduction of coupling is represented as the system being shifted towards the endpoint of the human-in-the-loop systems in figure 1.

Evolving Hybrid Intelligence Systems Intelligent systems that evolve throughout their lifecycle, i.e. the contribution of human and machine intelligence change over time, can easily be explained through our proposed framework. One such system is *Calendar.help*, proposed by Cranshaw et al. 2017. It is a service that allows a user to schedule meetings conveniently and provide an experience similar to that of working with a human assistant. The user emails (carbon copy) the system's virtual assistant to use the service. After receiving the email, the virtual assistant follows up with the other person to decide upon the details, creates the meeting, and sends the invitation automatically. The proposed system's architecture was developed through multiple design iterations. Starting as a human computation system, it utilized domain experts to perform its tasks. In the next iteration, some of the tasks were transferred to crowdworkers. In the final version, machine intelligence was introduced to automate a few tasks. Authors have also noted that the dependence on human intelligence would be further reduced as more data is collected. This gradual progression during design iterations can be easily explained through our proposed framework as moving the system from human computation to tightly coupled hybrid intelligence and eventually to the human-in-the-loop systems endpoint. However, such a system could never be categorized as a self-sufficient artificial intelligence system, even if it could operate independently after some time because a considerable amount of human computation has gone into its development.

The exact unit of coupling between both intelligence, as well as examples that exist at the endpoints or the center of the continuum, have not been intentionally provided, as we believe that such statements can be made only when the field has matured and saturation is achieved concerning various types of hybrid intelligence systems. It is difficult to assert what kind of hybrid intelligence systems would have a maximum and minimum coupling of human and machine intelligence at this point. Therefore, these crucial types of hybrid intelligence systems can be thought of as theoretical limits.

4 Discussion

As mentioned in section 2, our proposed definition of hybrid intelligence systems is quite broad and forms a superset of all other kinds of intelligent systems. It also makes it evident that indeed all AI systems, except the self-sufficient artificial intelligence systems, are hybrid intelligence systems. And we believe that it is an important step towards acknowledging the much-needed fact that humans are an integral part of any AI workflow. It will compel designers, developers, and deployers to think about humans at every step in the process and obey the standard guidelines and principles in the domain (Amershi et al. 2019, Mathewson 2019). Such practices would encourage the development of human-centered intelligent systems that put humans before machines and help mitigate problems related to fairness, accountability, transparency, and ethics. More generally, this work aims to instigate more discussion on how AI practitioners could be encouraged and incentivized to give higher importance to human factors when designing, developing, and deploying intelligent systems.

Our proposed framework can help researchers to organize, group, and classify different types of existing and upcoming intelligent systems. It would lead to a better understanding of existing terminologies and bring consistency in the field. Developers and practitioners will also benefit from the organization as it will help them to make a better system design decision based on their requirements. It can also assist policymakers by targeting various types of hybrid intelligence systems to build better governance mechanisms.

5 Related Work

Dellermann et al. 2019 provided an initial conceptualization of the term hybrid intelligence systems by defining it as "*systems that have the ability to accomplish complex goals by combining human and artificial intelligence to collectively achieve superior results than each of them could have done in separation and continuously improve by learning from each other*". In this definition, the condition of continuous learning by both humans and machines restricts its applicability to most of the intelligent systems, which utilize both human and machine intelligence. For example, Teachable Machine,[3] an interactive machine learning system in which the end-user trains a classifier through direct interaction (iteratively provides input and evaluates output), will not be considered as a hybrid intelligence system as the human does not learn anything new by teaching the model.

Pilarski et al. 2017 proposed that humans as well as prosthetics and other technologies for human enhancement should be treated as autonomous agents of an information processing system. They provided a schema to analyze the relationship between the ability of each agent to have, seek, and achieve its goals and the overall human-machine performance. Based on the schema and existing works in the literature, they concluded that overall human-machine performance can be significantly improved by increasing the capabilities of prosthetic agents.

Kamar 2016b studied the reasoning capabilities of AI systems to determine when and how should it query humans to maximize its benefit from human computation, with a focus to enhance its applicability. Lasecki 2019 has postulated that hybrid intelligence systems could bridge the socio-technical gap between the capabilities of systems and varying demands of people. Kamar 2016a also addressed the challenges and opportunities for developing better crowdsourcing platforms with the rise of hybrid intelligence systems.

6 Conclusion and Future Work

In this work, we provided a precise definition of hybrid intelligence systems and argued that it forms a superset of all other similar concepts such as interactive machine learning, AI-infused systems, and mixed-initiative interactions. We also provided a simple intuitive framework to compare and contrast existing and upcoming hybrid intelligence systems based on the degree of coupling between human and machine intelligence and the directive authority of each party, along with examples from the literature. It would be interesting to investigate various properties, strengths, and limitations of systems positioned at different locations on the continuum. It would also be worth investing time to investigate the exact unit of coupling and precise definition of the maximum and minimum coupling between human and machine intelligence with examples.

Broader Impact

This article presents theoretical work on the precise definition and intuitive framework of hybrid intelligence systems to underpin the rapid growth in the development of such systems. It investigates the interaction of human and machine intelligence to build accountable, reliable, and trustworthy intelligent systems. It also advocates the idea of examining human factors at every stage of the intelligent systems' lifecycle. This research will not only benefit researchers but also AI practitioners in the industry. This work does not include any system building and is not related to any particular dataset. We believe that this work has no adverse impact on society.

Acknowledgments

We thank anonymous reviewers for their helpful suggestions and constructive feedback.

References

J. S. Albus. Outline for a theory of intelligence. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3):473–509, 1991.

[3] https://teachablemachine.withgoogle.com/

S. Amershi, M. Cakmak, W. B. Knox, and T. Kulesza. Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4):105–120, Dec. 2014. doi: 10.1609/aimag.v35i4.2513. URL https://www.aaai.org/ojs/index.php/aimagazine/article/view/2513.

S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen, J. Teevan, R. Kikin-Gil, and E. Horvitz. Guidelines for human-ai interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. doi: 10.1145/3290605.3300233. URL https://doi.org/10.1145/3290605.3300233.

J. Cranshaw, E. Elwany, T. Newman, R. Kocielnik, B. Yu, S. Soni, J. Teevan, and A. Monroy-Hernández. Calendar.help: Designing a workflow-based scheduling agent with humans in the loop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 2382–2393, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346559. doi: 10.1145/3025453.3025780. URL https://doi.org/10.1145/3025453.3025780.

D. Dellermann, A. Calma, N. Lipusch, T. Weber, S. Weigel, and P. Ebel. The future of human-ai collaboration: A taxonomy of design knowledge for hybrid intelligence systems. In T. Bui, editor, *HICSS*, pages 1–10. ScholarSpace, 2019. URL http://dblp.uni-trier.de/db/conf/hicss/hicss2019.html#DellermannCLWWE19.

J. A. Fails and D. R. Olsen. Interactive machine learning. In *Proceedings of the 8th International Conference on Intelligent User Interfaces*, IUI '03, page 39–45, New York, NY, USA, 2003. Association for Computing Machinery. ISBN 1581135866. doi: 10.1145/604045.604056. URL https://doi.org/10.1145/604045.604056.

X. He, K. Zhao, and X. Chu. Automl: A survey of the state-of-the-art, 2020.

E. Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, page 159–166, New York, NY, USA, 1999. Association for Computing Machinery. ISBN 0201485591. doi: 10.1145/302979.303030. URL https://doi.org/10.1145/302979.303030.

C.-Y. Huang, S.-H. Huang, and T.-H. K. Huang. Heteroglossia: In-situ story ideation with the crowd. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–12, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376715. URL https://doi.org/10.1145/3313831.3376715.

T.-H. K. Huang, J. C. Chang, and J. P. Bigham. Evorus: A crowd-powered conversational assistant built to automate itself over time. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–13, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3173869. URL https://doi.org/10.1145/3173574.3173869.

E. Kamar. Hybrid intelligence and the future of work. In *In Proceedings of the Productivity Decomposed: Getting Big Things Done with Little Microtasks Workshop at CHI 2016*, May 2016a. URL https://www.microsoft.com/en-us/research/publication/hybrid-intelligence-future-work/.

E. Kamar. Directions in hybrid intelligence: Complementing ai systems with human intelligence. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, IJCAI'16, page 4070–4073. AAAI Press, 2016b. ISBN 9781577357704.

W. S. Lasecki. On facilitating human-computer interaction via hybrid intelligence systems. 2019.

W. S. Lasecki, C. D. Miller, I. Naim, R. Kushalnagar, A. Sadilek, D. Gildea, and J. P. Bigham. Scribe: Deep integration of human and machine intelligence to caption speech in real time. *Commun. ACM*, 60(9):93–100, Aug. 2017. ISSN 0001-0782. doi: 10.1145/3068663. URL https://doi.org/10.1145/3068663.

J. Liang, E. Meyerson, B. Hodjat, D. Fink, K. Mutch, and R. Miikkulainen. Evolutionary neural automl for deep learning. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2019)*, 2019. URL http://nn.cs.utexas.edu/?liang:gecco19.

A. Lundgard, Y. Yang, M. L. Foster, and W. S. Lasecki. Bolt: Instantaneous crowdsourcing via just-in-time training. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–7, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3174041. URL https://doi.org/10.1145/3173574.3174041.

K. W. Mathewson. A human-centered approach to interactive machine learning, 2019.

R. Miikkulainen. Neuroevolution. In D. Phung, C. Sammut, and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Science, 3rd Edition*. Springer, New York, 2020. URL http://nn.cs.utexas.edu/?miikkulainen:encyclopedia20-ne.

P. M. Pilarski, R. S. Sutton, K. W. Mathewson, C. Sherstan, A. S. R. Parker, and A. L. Edwards. Communicative capital for prosthetic agents, 2017.

A. J. Quinn and B. B. Bederson. Human computation: A survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 1403–1412, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450302289. doi: 10.1145/1978942.1979148. URL https://doi.org/10.1145/1978942.1979148.

K. O. Stanley, J. Clune, J. Lehman, and R. Miikkulainen. Designing neural networks through evolutionary algorithms. *Nature Machine Intelligence*, 1:24–35, 2019. URL http://nn.cs.utexas.edu/?stanley:naturemi19.

J. W. Vaughan. Making better use of the crowd: How crowdsourcing can advance machine learning research. *J. Mach. Learn. Res.*, 18(1):7026–7071, Jan. 2017. ISSN 1532-4435.