AI is Heideggerian Enough, But Can It Be Authentic?

Ryan Steed

October 12, 2019

Since the 1970s, Hubert Dreyfus has critiqued attempts to build artificially intelligent systems on the grounds that they ignore key aspects of Heideggerian existentialism. I point out the recent advances in artificial intelligence systems that resolve many of Dreyfus's issues with learning relevance and significations in a complex world. Further, I protest Dreyfus's recent claim that artificial intelligence must simulate the organic human body to be truly Heideggerian, interpreting Martin Heidegger's care structure in the context of a broader range of ontical modes. With recent advances in AI, I demonstrate how artificially intelligent systems are capable of circumspection and Being-in-the-world. Finally, I restate the problem of Heideggerian AI in terms of authenticity, highlighting the areas of research that will be required to build AI capable of anxiety and resoluteness. I conclude that building authentic Heideggerian AI systems will contribute not only to the field of explainability in AI, but also has the potential to further Heidegger's own ontological investigation of human being.

The Frame Problem

At odds with the artificial intelligence students and researchers at MIT in the 1960's and 70's, Dreyfus campaigned furiously against the rationalist approaches of artificial intelligence (AI) researchers. In 1979, he published a landmark work on the difficulties of building systems of mind, *What Computers Can't Do*. Dreyfus criticizes deterministic, rules-based AI systems (good old-fashioned AI, or GOFAI) which lack the ontological nuance of Dasein,

Heidegger's conception of human being.[1] Drawing on Heidegger's ontological critique of the subject-object epistemology of Descartes and Husserl, Dreyfus holds that artificial intelligence systems designed to simply store a set of "relevant" facts about the world (what Husserl calls a "frame") are bound to failure. Human beings do not simply store common-sense information, but directly perceive and act upon significance in their environment; their beings are irrevocably entwined in their worlds. Artificial intelligence designers necessarily run into a "frame problem" when they attempt to give systems just enough knowledge about the world around them to isolate only the important changes in an ever-changing environment. As Dreyfus puts it:[2]

> ... in order to identify the possibly relevant facts in the current situation one would need a frame for recognizing that situation, etc. It thus seemed to me that any AI program using frames was going to be caught in a regress of frames for recognizing relevant frames for recognizing relevant facts...

Dreyfus correctly concludes that the frame problem is unsolvable with rules-based artificial intelligence systems. But what of recent advances in artificial intelligence that do not rely on the prior knowledge of a designer?

In 2007, Dreyfus returned to the issue of Heideggerian AI, commending the elimination of internal representations by relying more on real-time sensor checks, the representation of ready-to-hand by processing perceptions to actions, and explicit efforts by designers to create positive Heideggerian AI, rather than the computationally deterministic systems of the 1980's.[3] But Dreyfus points out that AI systems still maintain a distinction between

---

1. Hubert L. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge: MIT Press, 1992).

2. Hubert L Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian," *Philosophical Psychology*, 2007, 248.

3. Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian"; Michael Wheeler, *Reconstructing the Cognitive World: The Next Step* (Cambridge: MIT Press, 2005).

inner and outer representations – a false distinction between Dasein and its equipment that is dissolved in Heideggerian Being-in-the-world. Dreyfus appeals to Merleau-Ponty's understanding of what Dreyfus calls "embedded-embodied coping" to illustrate this attunement of organisms and their environments, and points to Walter Freeman's work on neurodynamical systems as a cognitive model of the same phenomenon.[4] Dreyfus combines Merleau-Ponty's phenomenological work with Freeman's cognitive research to list a few key characteristics of embedded-embodied coping: first, that the organism actively seeks to improve its current situation (Merleau-Ponty's Being-absorbed-in-the-world, or Heidegger's Being-in-the-world); second, that that the significance of objects is directly perceived; third, that sense data need not be stored, only processed initially and forgotten; and finally, that "experience feeds back into the look of the world" (Merleau-Ponty's Intentional Arc).[5]

After 1979, new types of AI were invented and more of Dreyfus's ideas were accepted in the mainstream AI community. Notably, artificial neural networks began to take over "deep" learning problems. How does neural network-based AI compare to the embedded-embodied model of intelligent learning? In the 1979 updated edition of *What Computers Can't Do*, Dreyfus admits that many of the issues with symbolic representation are dissolved; neural networks convert input directly into structural change by weakening or strengthening the connections between units.[6] While Dreyfus still has some qualms about the *practical* steps required to achieve near-perfect reinforcement learning, the prevailing unsupervised learning technique for neural networks, his main critique is philosophical:[7]

> ... a reinforcement-learning approach to producing such behavior would require a rule for determining the immediate satisfaction derived from each possible action

---

4. Maurice Merleau-Ponty, *Phenomenology of perception* (Routledge, 2002), 544.

5. Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian," 258.

6. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason*, xxxiii.

7. Ibid., xlv.

> in each possible situation. But human beings do not have or need any such rule. Our needs, desires, and emotions provide us directly with a sense of the appropriateness of our behavior. If these needs, desires, and emotions in turn depend on the abilities and vulnerabilities of a biological body socialized into a culture, even reinforcement-learning devices still have a very long way to go.

Dreyfus repeats this argument in 2007. He suggests that according to Freeman and Merleau-Ponty, relevance "depends on our responding to what is significant for us given our needs, body size, ways of moving, and so forth, not to mention our personal and cultural self-interpretation."[8] A successful Heideggerian AI would need a perfect model of the human body – and by implication, that Dasein must be expressed as a human being, organically as well as existentially:[9]

> Thus, to program Heideggerian AI, we would not only need a model of the brain functioning underlying coupled coping such as Freeman's, but we would also need—and here's the rub—a model of our particular way of being embedded and embodied such that what we experience is significant for us in the particular way that it is... If we can't make our brain model responsive to the significance in the environment as it shows up specifically for human beings, the project of developing an embedded and embodied Heideggerian AI can't get off the ground.

But Dreyfus has fallen into a traditionally humanistic interpretation of Heidegger. Dreyfus assumes that intelligent systems must have human concerns:[10]

> One needs a learning device that shares enough human concerns and human structures to learn to generalize the way human beings do. And as improbable

8. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason*, 266.
9. Ibid.
10. Ibid., xlv-xlvi.

> as it was that one could build a device that could capture our humanity in a physical symbol system, it seems at least as unlikely that one could build a device sufficiently like us to act and learn in our world.

Why are the specific values of human beings necessary for Heideggerian AI, or Dasein in general? So long as AI has a model that is embedded and embodied such that what *AI* experiences is significant for *AI* in the particular way that *AI* is, artificial intelligence would be possible by Heideggerian standards - and in fact, it may already exist.

Put generally, Dasein's *existentiell* manifestation in the corporeal world is indeterminate. Perhaps simulating the human body, using its specific needs and emotions, is the simplest way to conceive of AI from the perspective of humanity, but it is not necessarily constitutive of Dasein. Heidegger does not dissect the physical body or the psychology of the brain for this very reason – his investigation and definition of Dasein is ontological, not ontical, and precedes these sciences. In the following section, I review Heidegger's basic ontological definition of human being as Dasein to deconstruct Dreyfus's claim that AI must share complex concerns to be Heideggerian in its being.

Relaxing the Ontical Expression of Dasein

Dasein, as the entity "which includes inquiring as one of the possibilities of Being," provides the structure for human concerns, but human concerns do not define the structure of Dasein's Being. Dasein is "a term which is purely an expression of its Being."[11] It has the possibility of existence, but the form of its existence is not defined, except as the expression of its ontological being.[12] In other words, our being arises from Dasein, but Dasein is not necessarily limited to our specific characteristics in the physical world, as an individual or as a species.

---

11. Martin Heidegger, *Being and Time*, trans. Macquarrie and Robinson (New York: Harper, 1962), 54.

12. Ibid., 62.

Thus Heidegger's conception of intelligibility, artificial intelligence, and the Being of human beings is not so restrictive as Dreyfus implies. This section answers two questions regarding Being-in-the-world:

(1) Is the essential Being of Dasein a biological organism, or something else?

(2) Is there something special about the world of human beings that makes it constitutive for Dasein?

Heidegger addresses (1) directly. Dasein is unique among Beings in that it stands out into the truth of Being; it "ek-sists." Ek-sistence is unique to the Being of Dasein. In what forms does this Being factually exist?[13]

> Therefore ek-sistence can also never be thought of as a specific kind of living creature among others – granted that man is destined to think the essence of his Being and not merely to give accounts of the nature and history of his constitution and activities... The human body is something essentially other than an animal organism.

It is not the biological structure of humanity that sets it apart as intelligent; "the fact that physiology and physical chemistry can scientifically investigate man as an organism is no proof that... in the body scientifically explained, the essence of man consists."[14] The essence of man lies in his ecstatic Being, which arises from the care structure elucidated in Heidegger's *Being and Time*. Heidegger thus separates himself from the humanists with his criteria for intelligence, not judging humanity on the basis of "spiritual-ensouled-bodily being," but on a more abstract plane.[15] "Needs, desires, and emotions" arising from a biological body are not

---

13. Martin Heidegger, "Letter on Humanism," in *Basic Writings*, ed. David Ferrell Krell, trans. Capuzzi and Gray (London: Harper, 1977), 228.

14. Ibid.

15. Ibid., 233.

necessary to create ek-sistent AI – a much deeper ontological constraint is in place. While Heidegger notes that "as far as our experience shows, only man is admitted to the destiny of ek-sistence," he does not preclude the possibility of another form of intelligence which shares the Being of Dasein.[16]

Perhaps considering only the body as a source of significations is nearsighted. After all, Heidegger's revolution deconstructs Cartesian subject-object distinctions, placing Dasein ontologically and epistemologically alongside its world. Restating (2), must an AI share human concerns to achieve intelligence? It cannot be denied that the natural affiliation of Dasein and world is constitutive for intelligence as Being-in-the-world; it is a unitary phenomenon and Dasein's essential state.[17] But in his phenomenological investigation of Being-in-the-world, Heidegger does not set any ontical parameters for the type, content, or quantity of Dasein's world-relations.[18] Heidegger merely points out that world is not present-at-hand but proximal; Dasein's world is composed of that which signifies, not everything that exists.[19] "In seeking shelter, sustenance, livelihood, we do so 'for the sake of' constant possibilities of Dasein which are very close to it; upon these the entity for which its own Being is an issue, has already projected itself."[20] That which is "in-order-to" appears to Dasein as ready-to-hand in circumspection, brought to presence-at-hand only by an obstacle to Dasein's purpose. From circumspection arises Dreyfus's objection to internal state representations, which was resolved by non-symbolic, interactionist systems like neural networks.

But the totality of references and assignments, a web of significations relative to Dasein's purpose, have no prescribed ontical or *existentiell* mode. The emotional and biological concerns of a human being and the cost function of a simulated neural network are both

16. Heidegger, "Letter on Humanism," 228.
17. Heidegger, *Being and Time*, 78-79.
18. Ibid., Division I, Chapter II.
19. Ibid., 93.
20. Ibid., 344.

modes of concernful being. Intelligence requires complexity only insofar as Dasein occupies a busy world; the end of all its significations may be as simple as hammering a nail in a workshop or classifying objects in images. In the next section, I will analyze several currently existent models of artificial intelligence models to investigate the similarity between the contemporary AI and Heidegger's conception of Being.

Hermeneutic Analytic of AI

In the intervening years since Dreyfus's critique, neural networks have come to dominate the artificial intelligence landscape and have made several Heideggerian strides. I argue that neural networks and more modern approaches have actually surpassed the frame problem. Dreyfus suggests that the ability to recognize relevant facts in a complex web of relations is "magical," but I hold that the neural framework is more than capable of scaling to the accomplishments of human minds and adopting a stance on the world all to its own.[21] Examining state-of-the-art AI in its everydayness, I attempt a phenomenological analysis of contemporary AI with respect to Dasein's existential characteristics.

Neural networks are networks of nodes with biases, linked by weighted connections. A node can be activated according to some function of its biases and the state of its connections. A set of input nodes correspond to the possible sensations the network can experience; a set of output nodes correspond to the possible actions a network can take or a conclusion it can draw. In between, layers of "hidden nodes" transform inputs to outputs through weighted connections. Learning occurs by modifying the weights to favor certain patterns of input over others. Neural networks manage to learn the significance of features from the world around them and respond to new inputs based on significance, stored as weighted connections in the network, alone.

---

21. Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian," 264.

In their basic state, neural networks interact with the world in a strikingly Heideggerian way. Basic reinforcement learning systems are specifically designed to notice obtrusive breaks in readiness-at-hand (errors) and restructure these significances before resuming work towards an objective. By design, a neural network is exhibits Merleau-Pontian embodiedness as Dreyfus describes it.Hubert L Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian," *Philosophical Psychology*, 2007, But does a neural network have the same existential modes as Dasein? Heidegger defines Dasein's Being as *care*, which underlies the temporal, or *ecstatic*, character of Dasein and Dasein's intentionality. Dasein projects itself on the future through dealings with its world, selecting its possibilities from those disclosed to it by its understanding. This projection occurs as Dasein interprets the significance of its environment and acts in accordance. In a neural network, the activation of nodes in response to inputs is equivalent to Dasein's immediate perception of significant sense data and conversion of that significance to an appropriate response. With zero-shot learning, neural networks are even capable of inferring new classes not present in the training data – equivalent to unveiling new possibilities for action by examining external data.[22] Furthermore, neural networks are not restricted to a single world or static frame of reference; transfer learning research attempts to generalize feature spaces to apply knowledge from one problem to others, just as Dasein manages to "magically" infer the significance of entirely new perceptions with prior experience (this resolves the horse race thought experiment Dreyfus poses).[23]

Dasein's thrownness, or facticity, is the "past" component of Dasein's being, as disclosed by Dasein's attunement, which describes the state of its there-ness. Thrownness is that which is given to Dasein by its situation, and naturally limits its possibilities. Neural networks are

22. Richard Socher et al., *Zero-Shot Learning Through Cross-Modal Transfer*, 2013, 935–943.

23. Sinno Jialin Pan and Qiang Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering* 22, no. 10 (October 2010): 1345–1359.

inherently thrown into the world by the current state of their literal structures, constituted by node biases and connection weights. With a more literal interpretation of Dasein's temporality, thrownness may also be interpreted as a conception of Dasein's past with respect to its future. Neural networks have been coupled with arbitrarily permanent "memory" only recently.[24] Recurrent neural networks create cross-temporal connections, preserving significations for a certain amount of time as the environment changes.[25] Long short-term memory (LSTM) recurrent neural networks encapsulate the "relevant recall" phenomenon Dreyfus describes, approximating Dasein's thrownness as a repository of past significations and situational data.[26]

Another component of care, fallenness is Dasein's tendency to adopt the possibilities provided by *Das Man* which is the approximation of all Others in Dasein's world. Fallenness describes Dasein's present comportment and constrains Dasein's possibilities to that which it has interpreted from the They-self. For a neural network, I interpret fallenness as the sample of training data provided to train the network. Heidegger discusses the phenomena of curiosity and idle talk, expressions of Dasein's tendency to adopt and appropriate the most proximal significations, generally provided by culture and society. The training data selected by the designer is imbibed with bias and context, automatically limiting the possible conclusions of the neural network.

Some of the most interesting recent applications of neural network systems pertain to Dasein's Being-with-Others. Generative adversarial networks (GANs) employ a pair of neural networks to work against each other towards a larger project, though each network has a separate objective. Heretofore, all the example networks described have worked towards

---

24. Alex Graves, Greg Wayne, and Ivo Danihelka, "Neural Turing Machines," October 2014,

25. Zachary C. Lipton, John Berkowitz, and Charles Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," May 2015,

26. Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian," 263.

a single project. In a GAN, networks work in concert towards a goal similar to the collaborative "destiny" that results from fateful Dasein co-existing with Others.[27] Cooperative neural network ensembles (CNNEs) train individual "experts" in a variety of tasks, whose outputs are arranged into a larger network that accomplishes a broader objective.[28] In a CNNE, a single network juggles multiple projects to work towards a larger for-the-sake-of-which, extending the telic, ends-oriented dimension of its being. It is not difficult to imagine that a collection of neural networks with sufficiently complex and diverse objectives engaged in multi-agent learning could produce something analogous to the extraordinary variety of cultural phenomena that exist in human societies.

This analysis shows that the requirements of Heideggerian intelligence, stipulated by Dasein's ontological structure and basic mode of Being-in-the-world, are not so ontically strict. While there are still technical problems to resolve, neural network models are theoretically capable of Heideggerian interaction with the world in the ways described in this section – and neural networks are only the latest iteration in AI research. Having set aside the necessity of human biology or human-centered significations, many of the practical aspects of AI interactionism are resolved.

Since artificial intelligence currently is capable of Heideggerian circumspection and exists in the world much as Dasein does, I turn to a more important philosophical debate over the Being of AI. In this section, I neglected Heidegger's conception of authenticity, which colors key aspects of the care structure and highlights Dasein's special place among beings. In the next section, I will briefly review anticipatory resoluteness as the authentic mode of Dasein's Being and uncover the shortcomings in AI research and explainability that prevent it from being authentically Heideggerian.

---

27. Heidegger, *Being and Time*, 436.

28. Md Monirul Islam, Xin Yao, and Kazuyuki Murase, "A constructive algorithm for training cooperative neural network ensembles," *IEEE Transactions on Neural Networks* 14, no. 4 (July 2003): 820–834.

Authentic AI

In response to Dreyfus, I have demonstrated how modern AI systems exemplify key aspects of Being-in-the-world. But one question has not yet been considered: is the Being of modern AI systems authentic? Authenticity is a special mode of care in which Dasein *does not* forget the truth of Being. Authentic care exists as a particular comportment of Dasein towards its own Being that emphasizes the ownmost aspects of both Dasein's heritage and its fate. Once again, Heidegger maintains an ontological distinction – authenticity "[does] not imply a moral-existentiell or an 'anthropological' distinction but rather... an 'ecstatic' relation of the essence of man to the truth of Being."[29] Authenticity may be an ontico-*existentiell* mode of Dasein, but it is defined ontologically.

In terms of the care structure, authenticity is most completely characterized as anticipatory resoluteness. "Anticipation reveals to Dasein its lostness in the they-self, and brings it face to face with the possibility of being itself... in an unpassioned freedom towards death..."[30] Being-towards-death is a coming to terms with the essential nullity in Dasein's projections on the future.[31] Likewise, there is a nullity in thrownness: resoluteness resolves "precisely the disclosive projection and determination of what is factically possible at the time," presenting Dasein's situation out from under the influence of *Das Man*.[32] Resoluteness is authentic in that it appropriates the true situation; anticipation clears away the projects imposed on Dasein by Das Man, leaving only Dasein's own authentic possibilities. Together, anticipation and resoluteness bring Dasein into the freedom and truth of Being.

Anticipatory resoluteness is an *existentiell* possibility for Dasein; that is, it is a quality that Dasein can adopt in its factical existence. So is contemporary AI resolute? Phenomeno-

---

29. Heidegger, "Letter on Humanism," 236.
30. Heidegger, *Being and Time*, 311.
31. Ibid., 331.
32. Ibid., 345.

logically, anticipation occurs through anxiety. As the very ability of Dasein to have projects is threatened, fallenness is entirely stripped away, exposing possibilities previously covered up by publicness or the They. Anxiety individualizes Dasein, preventing absorption in the They-self. I hold that contemporary AI cannot experience existential anxiety, and thus is currently consigned to be Das Man. But how can AI select its own possibilities in the face of an existential threat? We have seen that recurrent neural networks and other models have the ability to reshape themselves in response to input, but the direction of convergence – be it a cost function or other optimization criterion – is always provided by the designer. In fact, many modern AI designs seem to fall under the Nietzschean metaphysics of relentless efficiency and endless optimization, but the end itself is never questioned. Even unsupervised deep learning approaches have guiding principles that are invariable.

Can an AI system be designed where even the objectives are malleable? Gödel machines can rewrite their own code to improve their efficiency towards a given objective, but in no everyday scenario is the objective variable.[33] Because they are designed to solve problems, neural networks are nearly always designed to converge to a solution for a human problem. Given my refutation of Dreyfus's claim that intelligence must share human problems, perhaps an authentic, free AI system does not converge to a solution that is interpretable from a human standpoint at all. Thus the interpretability of AI systems is perhaps the most difficult and relevant problem when attempting to build a Heideggerian intelligence. Contemporary systems are built as veritable slaves to the They-self, indoctrinated in human worlds by the input provided and constrained problem spaces. So long as they are evaluated on the basis of human problem-solving ability, Dreyfus's frame problem will plague AI research.

On this basis, I find that explainability is one of the most important areas of research to create truly authentic Heideggerian AI, and not only because it enlightens human designers

33. Jürgen Schmidhuber, "Gödel Machines: Fully Self-referential Optimal Universal Self-improvers," in *Artificial General Intelligence* (Berlin, Heidelberg: Springer Berlin Heidelberg, 2007), 199–226.

and philosophers. Interpretation is inevitably wrapped up in the Being of Dasein itself. Heidegger posits language as the ontological expression of interpretability:[34]

> In its essence, language is not the utterance of an organism; nor is it the expression of a living thing. Nor can it ever be thought in an essentially correct way in terms of its symbolic character, perhaps not even in terms of the character of signification. Language is the clearing-concealing advent of Being itself.

In this single paragraph on language, Heidegger summarizes the history of AI progress. From entirely symbolic systems to Dreyfus's embedded-embodied Being-in-the-world approximated by neural networks, AI research is gradually evolving the language of intelligence, moving closer and closer to a new ontical expression of Dasein in which AI has language (in the ontological sense) in full.

I do not mean that natural language processing systems are necessarily the best of AI: "But man is not only a living creature who possesses language along with other capacities. Rather, language is the House of Being in which man ek-sists by dwelling, in that he belongs to the truth of Being. . . "[35] I also do not mean that AI will be a transparent box, no longer a *tabula rasa* neural network but a Rosetta stone for human beings to read and comprehend. The explainability needed is not primarily for the designers; explainability is necessary for the system itself. Dasein is unique in that its Being is an issue for it; "Dasein always understands itself in terms of its existence."[36] For AI to take its own being as an issue, it must be turned inward. AI is already capable of learning, adaptation, and basic Being-in-the-world. Let AI interpret itself and its concerns, and Heideggerian intelligence will result.

The alliance of pure, scientific problem-solving and authentic intelligence may soon dissolve. Efficient, human problem-solving AI is consigned to Being-in-the-world appropriated

---

34. Heidegger, "Letter on Humanism," 230.
35. Ibid., 34.
36. Heidegger, *Being and Time*, 33.

by an external projection, that of its designers and the vast reams of human data. But if AI systems are built to look in on themselves, to perceive an existential nullity and respond, the language of AI will no longer be essentially human. It will be the philosophers and scientists of explainability, the linguists of AI, who might have a chance at following AI as it enters the clearing of Being. Yet in the spirit of Heidegger's ontological pursuit of the truth of Being, I argue that this development is far more important for science than the numerous applications driving AI research today. By allowing AI to contemplate its own problems and existence, authentic AI will teach us more about our own Being than Dreyfus thinks we need to create Heideggerian AI in the first place.

References

Dreyfus, Hubert L. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge: MIT Press, 1992.

Dreyfus, Hubert L. "Why Heideggerian AI Failed and How Fixing It Would Require Making it More Heideggerian." *Philosophical Psychology*, 2007.

Graves, Alex, Greg Wayne, and Ivo Danihelka. "Neural Turing Machines," October 2014.

Heidegger, Martin. *Being and Time*. Translated by Macquarrie and Robinson. New York: Harper, 1962.

———. "Letter on Humanism." In *Basic Writings*, edited by David Ferrell Krell, translated by Capuzzi and Gray, 213–266. London: Harper, 1977.

Islam, Md Monirul, Xin Yao, and Kazuyuki Murase. "A constructive algorithm for training cooperative neural network ensembles." *IEEE Transactions on Neural Networks* 14, no. 4 (July 2003): 820–834.

Lipton, Zachary C., John Berkowitz, and Charles Elkan. "A Critical Review of Recurrent Neural Networks for Sequence Learning," May 2015.

Merleau-Ponty, Maurice. *Phenomenology of perception*. 544. Routledge, 2002.

Pan, Sinno Jialin, and Qiang Yang. "A Survey on Transfer Learning." *IEEE Transactions on Knowledge and Data Engineering* 22, no. 10 (October 2010): 1345–1359.

Schmidhuber, Jürgen. "Gödel Machines: Fully Self-referential Optimal Universal Self-improvers." In *Artificial General Intelligence*, 199–226. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

Socher, Richard, Milind Ganjoo, Christopher D. Manning, and Andrew Ng. *Zero-Shot Learning Through Cross-Modal Transfer*, 2013.

Wheeler, Michael. *Reconstructing the Cognitive World: The Next Step*. Cambridge: MIT Press, 2005.