Logan Carter

Last updated: 6/16/24 (not published)

Imagine This: Opaque DLMs are Reliable in the Context of Justification

1.0 – Introduction

Artificial intelligence (AI) and machine learning (ML) models have undoubtedly become useful tools in science. In general, scientists and ML developers are optimistic – perhaps rightfully so – about the potential that these models have in facilitating scientific progress. The philosophy of AI literature carries a different mood. The attention of philosophers remains on potential epistemological issues that stem from the so-called "black box" features of ML models. For instance, Eamon Duede (2023) argues that opacity in deep learning models (DLMs) is epistemically problematic in the context of justification, though not in the context of discovery.

In this paper, I will show that the epistemological concern regarding the black box features of DLMs is echoed in the epistemology of the imagination literature. A longstanding tradition in the philosophy of mind contends that it is epistemically problematic to rely on the imagination in the context of justification, but not in the context of discovery. This paper explores how far the analogy between the two literatures can be extended.

As such, the paper will be split into five parts. First, I will define "opacity" and "transparency" in AI and ML, with a specific focus on DLMs. Second, I will set out Duede's argument that the opacity of DLMs is epistemically problematic in the context of justification, albeit epistemically unproblematic in the context of discovery. Third, I will show that an analogous argument is made about the imagination. Fourth, I will suggest that the constraints-based approach to the imagination answers the epistemological concern in the imagination case. I will then argue by analogy that a constraints-based approach of sorts can be extended to answer the epistemological concern about opaque DLMs. Finally, I show that constraints on opaque DLMs can improve three different forms of transparency, thus easing the epistemological concern.

2.0 – Opacity and Transparency in DLMs

The complexity and network size of DLMs makes it difficult to determine exactly what functions and operations lead to their outputs. Even when experts know the fundamental details of the model, there can be elements of the model that are "opaque," meaning that there are elements or operations of the model that are not completely understood, explainable, or interpretable.

Paul Humphrey's (2009) analysis of computing systems in science serves as both an influential and helpful starting point in understanding opacity in DLMs. According to Humphrey, systems are "opaque relative to a cognitive agent *X* at time *t* just in case *X* does not know at *t* all of the epistemically relevant elements of the [system]" (618). Humphrey's analysis highlights two important aspects of opacity. First, opacity is not an intrinsic feature of a model, but an extrinsic feature that is agent-relative (Zednik, 269). Models in and of themselves are not opaque; they are opaque insofar as there are epistemically relevant elements of the system that *agents* cannot fully understand, explain, or interpret. Second, opacity refers to an agent's *lack of knowledge* of epistemically relevant elements of the system, i.e. elements of the system "which can be cited by [an agent] to explain the occurrence of some other element or of the system's overall output" (Zednik, 269). So, opacity is an epistemological problem because it prevents us from knowing all of the epistemically relevant elements of a model which renders us unable to understand how exactly the model produces its outputs.

Opaque models are typically contrasted with "transparent" models whose elements and operations are known, explainable, and interpretable. Kathleen Creel's (2020) analysis of transparency in complex computing systems involves three forms of transparency. *Functional transparency* concerns our knowledge of how the algorithm functions as a whole. *Structural transparency* concerns our knowledge of how the algorithm is realized in code. Finally, *run transparency* concerns our knowledge of a particular instantiation of the code. These forms of transparency can be exhibited independently (Creel, 581). The recent boom in "explainable AI" research comes from attempts to increase transparency in opaque DLMs so that we can be confident in the reliability of the model's outputs.

On the face of it, transparent DLMs are epistemically on better footing than opaque models because we can cite all of the epistemically relevant elements that lead to the outputs of transparent models; this gives us a reason to trust the reliability of the outputs of transparent models. Opaque DLMs, however, are cause for epistemological concern since we do not know all of the epistemically relevant elements that lead to an output, which calls into question the reliability of its outputs. Nevertheless, according to Duede, scientists trust the outputs of opaque DLMs in a diverse range of scientific contexts. Scientists often consider themselves justified in relying on DLMs and treat the outputs as "claims about the target systems upon which the models were trained" (Duede 2022, 491). How can scientists justifiably rely on the outputs of opaque DLMs?

3.0 – Duede on DLMs

How can scientists justifiably rely on the outputs of opaque DLMs, given that we do not know how the model produces the outputs that it does? To answer this question, Duede (2023) draws on the classic distinction in the philosophy of science between the context of justification and the context of discovery. According to Duede, access to the high-level logical rules that opaque DLMs follow is required for the outputs of the model to be evaluated in the context of justification (2023, 2). However, network opacity bars us from accessing these rules and is therefore epistemically problematic in the context of justification.

While Duede raises epistemological concerns about opacity in DLMs in the context of justification, he does not think that opacity in DLMs is epistemically problematic in the context of discovery. This is because when situated in the context of discovery, the outputs of opaque DLMs are treated merely as guides for scientific research. When treated in this way, the outputs of opaque DLMs are similar to abduction and problem solving-heuristics in that they "can serve to facilitate discovery without their outputs or internal logic standing in need of justification…[and thus] can be treated as situated in the 'context of discovery'" (Duede 2023, 5-6). Thus, opacity is epistemically unproblematic in the context of discovery because the model's outputs "can be used to guide attention and scientific intuition toward more promising hypotheses but do not, themselves, stand in need of justification…[the] outputs of opaque models serve to provide reasons to or evidence for pursuit of particular paths of inquiry over others" (Duede 2023, 6).

Though Duede proposes that opacity in DLMs is epistemically problematic in the context of justification, he acknowledges that scientists nevertheless trust the reliability of DLMs beyond the context of discovery. That is, scientists typically treat the outputs of opaque models as "claims about the target systems upon which the models were trained" rather than mere guides for scientific research (Duede 2022, 491). We return to a familiar question: what justifies scientists' reliance on opaque DLMs? Duede (2022) appeals to three philosophically familiar justifications in the philosophy of science: (i) brute inductive considerations, (ii) scientific instruments, and (iii) trust in scientific experts, and concludes that the justification for relying on opaque DLMs is reducible to none of the three philosophically familiar justifications.

So, Duede is arguing that opacity in DLMs is epistemically problematic in the context of justification insofar as there is not yet a philosophically familiar form of justification for our relying on the model's outputs. Duede is *not* arguing that scientists *cannot* be justified in relying on opaque DLMs or that scientists' trust in opaque models is unwarranted. Rather, the use of opaque DLMs in science has presented an opportunity for novel approaches to reliability in the philosophy of science. The project for the remaining sections is to offer such a novel approach, one that will begin by appealing to the epistemology of the imagination.

4.0 – Intro to Imagination

Now that we have seen Duede's argument that opacity in DLMs is epistemically problematic in the context of justification, albeit epistemically unproblematic in the context of discovery, I will now show that there is an analogous argument made in the epistemology of imagination literature.

It is commonplace in the epistemology of imagination literature to hold that whatever justificatory power the imagination has, if any at all, must be limited to the modal sphere. According to this tradition, "the imagination is, at most, a capacity for generating ideas, a means for coming up with candidates for knowledge…[that] are then to be evaluated on the relevant evidence, irrespective of their provenance" (Kinberg & Levy, 2). In other words, the imagination plays a role in *generating* beliefs but does not play a role in *justifying* those beliefs. Amy Kind (2018) characterizes this sentiment as *the charge of epistemic irrelevance*: while the imagination might justify our beliefs about what is possible, the imagination cannot justify our beliefs about what is actual (227).

The charge of epistemic irrelevance can be reconstructed in a way that closely resembles Duede's epistemological concern: according to the charge of epistemic irrelevance, relying on the imagination is epistemically unproblematic in the context of discovery since its outputs can be treated merely as guides for further inquiry, whether it be in everyday experiences, scientific settings, or other. For example, one might be led by their imagination to believe that a couch can fit through a doorway, but to be sure, they do a pen-and-paper calculation that reinforces their belief. After all, "no one denies that an act of imagination can lead to interesting innovations, discoveries, or new directions of research" (Kind & Kung 2016, 147).

However, according to the charge of epistemic irrelevance, relying on the outputs of the imagination *is* epistemically problematic in the context of justification since the imagination in and of itself does not serve to justify the beliefs that it generates. Thus, an analogous epistemological concern about opaque DLMs and the imagination has surfaced: our reliance on the imagination is epistemically unproblematic in the context of discovery, but epistemically problematic in the context of justification. Analogously, opacity in DLMs is epistemically unproblematic in the context of discovery, but epistemically problematic in the context of justification.

However, not all accounts of the imagination fall into the traditional camp which endorses the (reconstructed) charge of epistemic irrelevance. Recently, popular accounts have argued that the imagination *does* play a role in justifying our beliefs and, as such, is a method for knowledge acquisition in both practical and theoretical matters. The idea that the imagination can be a method for obtaining knowledge "runs contrary to a fairly entrenched tradition in epistemology and philosophy of mind, associated with the distinction between the context of discovery and the context of justification" (Kinberg & Levy, 2). Two views argue that it is epistemically unproblematic to rely on the imagination in the context of justification: *black-box reliability arguments* for the imagination and *constraints-based approaches* to the imagination.

Black-box reliability arguments turn out to be non-starters for the current discussion since these arguments seek to establish the reliability of the imagination without explaining how the imagination obtains its reliability. Thus, black-box reliability arguments are not going to help answer the question of how we can justifiably rely on the imagination in the context of justification since they do not rescue us from the 'black box territory,' so to speak, that we are aiming to escape. What remains, then, is the constraints-based approach which has the resources to dissolve the concern that the reliance on the imagination is epistemically problematic in the context of justification.

4.1 – Constraints-Based Approach

Unlike the black-box reliability arguments, constraints-based approaches attempt to establish *and* explain the imagination's reliability by appealing to constraints on the imagination. In this view, constraints are typically cashed out as guides to the imagination: constraints are "explicit, intentionally-followed 'algorithms,' so to speak," that guide the imagination in answering practical or theoretical questions (Kinberg & Levy, 11). Some constraints on the imagination include (but are not limited to): empirical and theoretical statements, prior beliefs, perceptual experiences, etc., all of which constrain the inputs and content of the imaginative unfolding. For example, in using my imagination to answer whether a couch will fit through a doorway, constraining my imagination involves picturing the couch and doorway with particular dimensions, that gravity is at work, and so on. In this case, empirical statements about objects in the world, and the world itself, constrain my imagining.

Kind, a proponent of the constraints-based approach, claims that constrained imagination not only serves “as an impetus for new ideas but rather in a justificatory capacity with respect to those ideas” (2018, 229). Conveniently, Kind leverages her account by an analogy to computer simulations: “[a] computer simulation contains only the facts that are put into it, but it can nonetheless provide us with information about the world,” so too, imaginative simulations can provide us with information about the world insofar as they are appropriately constrained (2018, 241). Since we typically consider ourselves justified in relying on the outputs of constrained computer simulations, we should consider ourselves justified in believing and relying on the outputs of constrained imaginative simulations.

Without constraints, the imagination is typically chaotic, random, and largely unhelpful in problem-solving heuristics (e.g. daydreaming). The thought is that by applying constraints to the imagination, we rein in the chaos and focus the imagination to solve the problem at hand. Constraints factor into the imagination in that they “causally affect how the imagination unfolds…[c]ognitive uses of the imagination often involve constraining my imaginings according to *my beliefs or evidence*” (my emphasis, Myers, 3254). For example, when I use my imagination to learn whether my professor will give my paper a good grade, my imagination is constrained by my beliefs and dispositions about my professor and empirical and theoretical statements about the quality of my work. By applying constraints to the imagination, we can be justified in relying on the outputs of the imagination. Thus, relying on the imagination is epistemically unproblematic in the context of justification insofar as the imagination is appropriately constrained.

Under my reconstruction of the view, the constraints-based approach to the imagination faces the burden of explaining how the imagination obtains its reliability despite its black box features. Though I have not seen this explicitly stated in the literature, I take it that what proponents of the constraints-based approach have in mind is that constraints on the imagination allow us to overcome (but not eliminate) the black box features of the imagination. By 'overcome,' I mean that the black box features of the imagination act as obstacles in our explaining how the imagination obtains its reliability. Constraints overcome the obstacle of black box features of the imagination since we can appeal to constraints in explaining how the imagination obtains its reliability, i.e., we can justifiably rely on the imagination despite its black box features insofar as it is appropriately constrained. So, on the constraints-based approach, our lack of knowledge of all of the epistemically relevant elements of the imagination need not make our reliance on the imagination epistemically problematic in the context of justification. By applying constraints to the imagination, we can explain the reliability of the imagination, thus making it epistemically unproblematic to rely on the imagination in the context of justification.

The constraints-based approach, then, offers an account of the imagination with the following features: (i) it establishes the reliability of the imagination, (ii) it explains how the imagination obtains its reliability via constraints, and (iii) it shows that relying on the imagination is epistemically unproblematic in the context of justification since constraints allow us to overcome (but not eliminate) the black box features of the imagination.

5.0 – Solution Through Constrained DLMs

Now that we have seen that the constraints-based approach offers an account of how our reliance on the imagination is epistemically unproblematic in the context of justification, we are well-suited to extend a similar approach to the case of opaque DLMs. Is scientists' reliance on opaque DLMs epistemically unproblematic in the context of justification insofar as the model is appropriately constrained? If the analogy holds, then the answer is yes, and would result in a philosophically novel approach to the justification for relying on opaque DLMs. The project of the following section is to explain what exactly ‘constraints’ amount to for DLMs and to determine whether constraints make their opacity epistemically unproblematic in the context of justification.

Recall that in the case of imagination, there is a collection of constraints on the inputs and content of the imaginative processes, such as empirical and theoretical statements, background beliefs and dispositions, prior evidence, etc. This is what allows us to narrow the focus of our imagination to solve the task at hand. Thus, for the analogy to hold, we ought to expect that constraints on DLMs constrain the inputs and content of DLMs according to empirical and theoretical statements, background beliefs, prior evidence, etc. If it turns out that there are cases of opaque DLMs on which we impose constraints, then (depending on the constraints) a plausible extension of the analogy will be that the application of constraints on DLMs allows us to overcome their black box features such that opacity in DLMs is epistemically unproblematic in the context of justification.

The epistemological concern about opaque DLMs, and many ML approaches in general, is that we are barred from extracting interpretable information and obtaining complete knowledge of what high-level logical rules the model follows. Some ML learning approaches, however, attempt to overcome these difficulties by imposing constraints on models. There are at least two physical constraints on ML models: there are “soft constraints, which are enforced by adding extra penalties to the loss function; and hard constraints, which refer to conditions that must be satisfied when generating the model” (Pan 2021, 244). Physics-informed neural networks (PINNs), for example, are deep networks on which we impose soft and hard constraints. Developers train PINNs by “integrating fundamental physical laws and domain knowledge [into the model] by ‘teaching’ ML models about governing physical rules, which can, in turn, provide ‘informative priors’ — that is, strong theoretical constraints and inductive biases on top of the observational ones…physics-informed learning is needed, hereby defined as the process by which prior knowledge stemming from our observational, empirical, physical or mathematical understanding of the world can be leveraged to improve the performance of a learning algorithm” (Karniadakis et. al., 423). Further, PINNs “are constrained to respect any symmetries, invariances, or conservation principles originating from the physical laws that govern the observed data” (Raissi et. al., 687). As such, constraints on PINNs amount to ‘injecting’ the model with prior knowledge which “can act as a regularization agent that constrains the space of admissible solutions to a manageable size” (Raissi et. al., 686). Reducing the space of admissible solutions to a manageable size has the beneficial effect of making the models more interpretable: PINNs “utilize *prior knowledge* or *constraints* [to] yield *more interpretable* ML methods that remain robust in the presence of imperfect data (such as missing or noisy values, outliers and so on) and can provide accurate and physically consistent predictions, even for extrapolatory [and] generalization tasks” (my emphasis, Karniadakis et. al., 423). Thus, the merits of using constrained DLMs, such as PINNs, are that the performance of the model improves, and the model is more interpretable insofar as its solution space is reduced.

Given the definition and application of constraints above, it turns out that constraints on DLMs, such as PINNs, are similar to constraints on the imagination. On the one hand, PINNs are constrained (using soft and/or hard constraints) according to “prior knowledge stemming from our observational, empirical, physical or mathematical understanding of the world” (Karniadakis et. al., 423). On the other hand, the imagination is constrained according to prior knowledge stemming from empirical and theoretical statements, our beliefs about the world, prior evidence, etc. In the case of imagination, constraints allow us to overcome the black box features of the imagination such that we can justifiably rely on its outputs in the context of justification. By analogy, then, constraints on opaque DLMs can help us overcome their black box features such that we can justifiably rely on their outputs in the context of justification. This argument by analogy might appear too quick. To buttress the argument, I will show that constraints on opaque DLMs promote three forms of transparency, which leverages the claim that constraints allow us to overcome (but not eliminate) the black box features of opaque DLMs.

6.0 – Constraints Improve Transparency

Constraints on opaque DLMs increase interpretability by reducing the size of admissible solutions space, which, in turn, improves the network's transparency (thus reducing opacity). The application of constraints also contributes to network transparency by improving each of Creel's three forms of transparency. I'll consider each form of transparency in turn.

To have functional transparency, one must know the high-level logical rules that lead to an output, which amounts to knowing the "algorithm that guides the learning process and the algorithm, model, or decision procedure that is learned" (Creel, 575). In the same way that constraints in the case of imagination are cashed out as "algorithms" (in the figurative sense) that guide imaginative processes, constraints in the DLM case can be cashed out as "algorithms" (in the literal sense) that guide the learning and problem-solving processes of DLMs. In this sense, we know what guides the learning process of the model, namely, the constraints on the model.

We can make use of the analogy to the imagination in the discussion of functional transparency. Duede argues that opacity in DLMs is epistemically problematic in the context of justification because we do not know the high-level logical rules that the model follows. I have argued against this in the case of imagination; one need not know all of the epistemically relevant elements of the imaginative process (including whatever high-level logical rules are being followed) to be justified in relying on the imagination in the context of justification. That is, constraints on the imagination allow us to overcome its black box features, thus making it epistemically unproblematic to rely on the imagination in the context of justification. By analogy, then, we need not know all of the high-level logical rules that opaque DLMs follow to be justified in relying on the outputs of these models in the context of justification insofar as they are appropriately constrained. Constraints on opaque DLMs allow us to overcome their black box features, thus making opacity in DLMs epistemically unproblematic in the context of justification.

Constraints on opaque DLMs also improve their structural transparency. What is required for structural transparency is "knowledge of the code as it instantiates the algorithm" (Creel, 577). Given that one has "access to the input data [and] the current weights of the network…at least for smaller neural networks, it is possible to know how the algorithm is instantiated in code" (Creel, 579). So, it is in principle possible to have structural transparency if one has access to the input data and the current weights of the network. In constrained DLMs, we do have access to the input data since the input data itself is being constrained by the developer. For example, one might constrain the input data of a weather model such that the model trains only on local data, say, data for a specific city. This input data is accessible to the developer of the weather model. Similarly, we can access the current weights of the system in PINNs, for example, since the weights themselves can be “user-defined or tuned automatically, and play an important role in improving the trainability of PINNs” (Karniadakis, et. al., 425). In cases where weights are user-defined, developers can constrain the model's weights, i.e., choose one value over another, and tailor these weights to whatever purpose the model is designed to serve. Thus, by applying constraints to an opaque DLM, we improve structural transparency since the input data and current weights of the model are themselves among the elements being constrained.

It might be argued here that even with constraints, the size and complexity of current DLMs make it impossible to know each instantiation of the algorithm in the code. However, I explained that in PINNs, for example, the existence of constraints on the model reduces the size of admissible solutions to one that is manageable. In this sense, constraints might help reduce even large and complex models to a size that allows us to know each instantiation of the algorithm in the code. I take it that the line between a ‘small neural network’ and ‘large neural network’ is fuzzy and with it the difference in being able to know each instantiation of the algorithm in the code. All I wish to suggest for structural transparency is that insofar as constraints only allow us access to input data and the current weights of the model, they improve structural transparency. But, I take it to be plausible that insofar as constraints reduce the size of admissible solutions, interpretability and structural transparency, too, are improved.

Finally, constraints promote run transparency. What is required for run transparency is “knowledge of the program as it was actually run in a particular instance, including the hardware and input data used” (Creel, 569) I take it that knowledge of the specific hardware of models is easy to come by, and not particularly important concerning the constraints that we impose on models. For example, developers might “constrain” a model by installing cheap hardware that increases the time it takes for models to produce outputs; insofar as this is a constraint, it is straightforward that developers know whatever hardware the model has. We know the input data used because, again, the input data itself is being constrained by the model's developers. Recall that, in the example of the weather model, a developer may constrain the input data to local weather data only (and not, say, weather data for the whole state). Thus, constraints improve run transparency in that we know the input data and hardware of the model.

We now have a complete picture of how constraints can improve three forms of transparency in opaque DLMs. Constraints can improve functional transparency in that they are "algorithms" (in the literal sense) that guide the model's learning and problem-solving processes. The discussion of functional transparency gave way to my argument by analogy to the imagination. In the case of imagination, we need not know the high-level logic involved in the imaginative process to justifiably rely on its outputs. Constraints on the imagination allow us to overcome its black box features such that we can rely on its outputs in the context of justification. Analogously, we need not know the high-level logical rules that opaque DLMs follow to be justified in relying on their outputs. Constraints on opaque DLMs allow us to overcome their black box features such that we can rely on their outputs in the context of justification. Constraints improve structural transparency in at least one of two ways: first, constraints allow us access to the input data and current weights of the model, and, second, constraints decrease the size of admissible solutions, which plausibly makes the model more interpretable and more likely (at least) that we know each instantiation of the algorithm in code. Finally, constraints improve run transparency because the input data and hardware of the model are being constrained, i.e., in constrained models, the input data and hardware used are user-defined.

This detour through transparency was meant to buttress the claim that constraints on opaque DLMs allow us to overcome their black box features such that we can rely on their outputs in the context of justification. Since transparency and opacity are opposite one another, the more transparent the model, the less opaque the model. Because constraints improve three different forms of transparency, they are responsible for decreasing opacity in DLMs; thus, constraints on opaque DLMs allow us to overcome the black box features of the model.

7.0 – Conclusion

In this paper, I have aimed to show three things. First, that there exists an analogy between the epistemological concern in the case of opaque DLMs and the imagination, which is that our reliance on each is epistemically problematic in the context of justification, albeit epistemically unproblematic in the context of discovery. I explained that the constraints-based approach answers the concern in the imagination case, i.e., constraints allow us to overcome black box features of the imagination such that we can justifiably rely on its outputs in the context of justification. I suggested that by extending a similar approach, we can answer the concern in the opaque DLM case. Indeed, it turned out that constraints on ML models, such as the soft and hard constraints imposed on PINNs, are similar to constraints on the imagination. I argued by analogy, contra Duede, that opacity in DLMs is epistemically unproblematic in the context of justification insofar as the model is appropriately constrained. Finally, I showed that the claim that constraints allow us to overcome the black box features of opaque DLMs has traction in that constraints can improve three forms of transparency.

Works Cited

1. Creel, Kathleen A. “Transparency in Complex Computational Systems.” *Philosophy of Science*, vol. 87, no. 4, Oct. 2020, pp. 568–589, https://doi.org/10.1086/709729.
2. Duede, Eamon. “Deep Learning Opacity in Scientific Discovery.” *Philosophy of Science*, vol. 90, no. 5, 16 Feb. 2023, pp. 1089–1099, https://doi.org/10.1017/psa.2023.8.
3. Duede, Eamon. “Instruments, Agents, and Artificial Intelligence: Novel Epistemic Categories of Reliability.” *Synthese*, vol. 200, no. 6, 19 Nov. 2022, https://doi.org/10.1007/s11229-022-03975-6.
4. Humphreys, Paul. “The Philosophical Novelty of Computer Simulation Methods.” *Synthese*, vol. 169, no. 3, 28 Nov. 2009, pp. 615–626, https://doi.org/10.1007/s11229-008-9435-2.
5. Karniadakis, George Em, et al. “Physics-Informed Machine Learning.” *Nature Reviews Physics*, vol. 3, no. 6, 24 May 2021, pp. 422–440, https://doi.org/10.1038/s42254-021-00314-5.
6. Kinberg, Ori, and Arnon Levy. “The Epistemic Imagination Revisited.” *Philosophy and Phenomenological Research*, vol. 107, no. 2, 22 Aug. 2022, pp. 319–336, https://doi.org/10.1111/phpr.12909.
7. Kind, Amy, and Peter Kung. “Imagining Under Constraints.” *Knowledge Through Imagination*, 1 Mar. 2016, pp. 145–159, https://doi.org/10.1093/acprof:oso/9780198716808.003.0007.
8. Kind, Amy. “How imagination gives rise to knowledge.” *Oxford Scholarship Online*, 21 June 2018, https://doi.org/10.1093/oso/9780198717881.003.0011.

9. Myers, Joshua. “The Epistemic Status of the Imagination.” *Philosophical Studies*, vol. 178, no. 10, 9 Jan. 2021, pp. 3251–3270, https://doi.org/10.1007/s11098-020-01600-1.
10. Pan, Jie. “Can Machines Learn with Hard Constraints?” *Nature Computational Science*, vol. 1, no. 4, 19 Apr. 2021, pp. 244–244, https://doi.org/10.1038/s43588-021-00067-2.
11. Raissi, M., et al. “Physics-informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations.” *Journal of Computational Physics*, vol. 378, Feb. 2019, pp. 686–707, https://doi.org/10.1016/j.jcp.2018.10.045.
12. Zednik, Carlos. “Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence.” *Philosophy & Technology*, vol. 34, no. 2, 20 Dec. 2019, pp. 265–288, https://doi.org/10.1007/s13347-019-00382-7.