UNIVERSITEIT GENT

Faculteit Letteren en Wijsbegeerte

Work without author:

Lockean theories of intellectual property and the product of artificial intelligence

Nathan Pelgrims

Werk zonder auteur: Lockeaanse theorieën van intellectueel eigendom en het product van artificiële intelligentie

Masterproef neergelegd tot het behalen van de graad van Master in de Wijsbegeerte door Nathan Pelgrims

Promotor: prof. dr. Sigrid Sterckx
Commissaris 1: dr. Jonathan Shaheen
Commissaris 2: dr. Kasper Raus

Studentennummer: 0140 8840
Academiejaar: 2020-2021

Word count

Word count 1: **32 453**

This is the sum of the number of words in the introduction, corpus, and conclusion.

Word count 2: **33 349**

This is the sum of the number of words in the total thesis, excluding any appendices and supplemental material.

Contents

Acknowledgements

While writing this thesis, I have received help and support from many people, without whom this would not have been possible. First, I need to thank prof. dr. Sigrid Sterckx for being willing to take up the role of promotor. Her classes have inspired me to think about moral and political philosophy more clearly. I am also hugely indebted to dr. Lisa Diependaele, my original promotor, for guiding me through my early ideas to achieve a clear goal. Her remarks and criticisms have helped me throughout the entire process. Timme Craeye's and Catherine Hale's corrections of my drafts have also been particularly valuable. I am grateful to my parents for their trust and patience. Lastly, there is Pauline Magerman, whose willingness to discuss these weighty subjects has helped me tremendously to avoid putting things too brown. Her continuous support has proven invaluable.

Abbreviations

AI: artificial intelligence

AGI: artificial general intelligence

ANI: artificial narrow/weak intelligence

DFL: desert from labour

LMA: labour-mixing argument

PR: property right

IPR: intellectual property right

TT: John Locke's Two Treatises of Government (followed by I or II)

P-creativity: psychological creativity

H-creativity: historical creativity

Preface

Intuitively, getting a reward for your labour seems fair. But why? What is it that makes labour into an activity that justifies property? I like to think of philosophy as the activity which questions our assumptions, finds problems where they are unexpected, where they should not be. To be fair, these problems are everywhere. Nonetheless, the universally accepted place of work in our society seems to raise the stakes in questioning the apparent.

One of the peculiarities in labour theories of property is the way they deal with automation. When a labouring person is replaced by a machine, all this ‘labour’ seems to have suddenly vanished, while the appropriation continues. With manual labour, this is something inevitable: nearly every form of technology represents automation in some way. However, recent developments have made automation of intellectual labour more than just a far-off prospect. Due to the rise of artificial intelligence it is now a very real thing that is already happening and is unlikely to stop soon.

In this thesis, I have ventured to combine these two things: a labour theory of property and autonomous, creative AI. In this respect, it is not a comprehensive work. Instead, my aim in this work is to provide a sketch of what the outcome of this confrontation may look like by identifying at least some of the issues where it will play be played out. I hope that this will elucidate some of the assumptions present in property theory.

Chapter 1 - Introduction

It was no more than a thought experiment when in 1913 Émile Borel had the idea of a monkey pressing random keys on a typewriter. If left to go at it for long enough, Borel proposed, the monkey would inevitably type every book in the French National Library.[1] Borel did not stop to think what this would imply for the copyright on these books. A little more than a hundred years later, this is becoming a real issue.

Technology has a way of creeping up on people. Artificial intelligence presents no exception. While we are used to hearing the term used either in apocalyptic predictions or in utopian futuristic fictions, the fact is that it is already here. Instead of the metal box with googly eyes and a lamp for a nose, it exists in the form of mathematical algorithms running on ordinary computers. This unexpected banality, however, does not mean that it does not bear huge potential, foreboding changes to all aspects of our lives.

One of these aspects, and perhaps one of the first to be affected, is property. Until recently, property theory and law did not have to mind the possibility of a machine authoring a work. That is, not in the way that is currently becoming possible. An autonomous machine taking the place of author could disturb existing theories of property. Some theories have no problem with this, but a Lockean labour theory of property is not one of them. For John Locke, labour was the justificatory principle for ownership. This is proving to be a difficult premise to maintain.

---

1 Émile Borel, 'La Mécanique Statique et l'irréversibilité', *Journal de Physique Théorique et Appliquée* 3, no. 1 (1913).

Locke could not predict his enormous influence on property theory. Known as ‘Lockean’ theories, countless interpretations of Locke’s work on ownership followed, mostly adopting labour as the foundational origin of property. Today, Lockean theories are still around, but are increasingly under threat from consequentialism. The latter kind of theories has the advantage of generally being more easily adaptable to change. But, having lived in the 17th century, Locke could hardly be blamed for failing to predict that one day machines would become capable of performing processes which could even replace human intellectual labour.

With this thesis, I intend to inquire upon this situation. Specifically, I will try to answer the question of who deserves property rights to the output of artificial intelligence from the perspective of a labour theory of property. This will require some reflection on a number of fundamental questions regarding intellectual property. Concepts like authorship, creativity, labour and moral agency will all need to be considered. I will argue that none of the relevant stakeholders are capable of claiming full authorship of the output of the AI. In other words, this means that a labour theory of property cannot provide a satisfactory answer to the question at hand.

The main difficulties that a labour theory of property is presented with when confronted with an AI-generated output derive from the fact that the AI seems to take up a role normally reserved for humans, namely that of the author. The adoption of this role by a machine could mean that valuable and meaningful labour is being performed without any real labour being expended. While such a conclusion would depend on two different notions of labour, it still manages to point out a fundamental difficulty encountered by labour theories in their confrontation with autonomous machines.

To be clear, this dissertation will not try to answer the question of ownership rights to the AI itself. This is another issue, requiring a fundamentally different approach. The scope of this text does not allow for the treatment of an additional matter like this. Therefore, I will refer to the owner only in an ambiguous sense, leaving the exact determination of this function open to further discussion.

An overview of this text

Before all else, it is necessary to determine which version of a labour theory we should elect to keep in mind throughout the rest of the text. This means, specifically, that I will discuss existing interpretations of Locke's theory, and decide which one is most capable of facing the issue of ownership of an AI-generated work. First (in 2.2), Locke's own theory will be discussed, followed by the problems surrounding this theory. By the next section, it should be clear why Locke's theory by itself is not adequate to tackle the issue of AI. In this part (2.3), more specific Lockean theories of property are also to be specified. Two interpretations stand out: the desert-from-labour theory and the libertarian theory. However, up to this point, only 'physical' or 'material' property has been considered. Intellectual property has some specific characteristics and needs, which I will look at in 2.4. In the same section, it should become clear how desert theories of property could offer some solutions to problems often faced with respect to intellectual property, and are therefore most suited for our purposes.

In chapter 3, artificial intelligence will be introduced. First, in 3.2, some of the features of AI relevant to this discussion are distinguished. Based on these features, I will identify three stakeholders, who all have a certain claim over the output of the AI. This will happen in 3.3. Each of these stakeholders has somehow contributed, and all are therefore potential owners from a labour-desert point of view. This translates into four candidates or options for appointing ownership: (1) the AI itself, (2) the creator of the AI, (3) the owner of the input data, and (4) determining ownership through a contract.

The AI as owner of its output will be discussed first in 3.4. First, I try to determine whether or not an AI is indeed capable of receiving rights. As this will not provide a satisfactory answer, the next matter to be treated is the question of whether the process performed by an AI can in fact be described as labour, and can therefore be susceptible to desert-from-labour. Again, the answer is ambiguous. There seems to be some possibility for ascribing labour to an AI, but this remains a difficult matter. Altogether, ascribing authorship to an AI seems to be impossible, due to the special nature and requirements associated with the concept.

The next candidate, the creator of the AI, is examined in 3.5. Central in this discussion will be the concept of creativity. As creativity is key to ascribing authorship, it is necessary to specify who performs the creativity. In other words: is an AI capable of achieving creativity autonomously? The difficulty in defining creativity itself makes this a challenging matter. I will argue that there are several kinds of creativity, some of which are attainable by a machine. Therefore, the creator of the AI does not present a valid claim over full ownership.

Then, in 3.6, it is time to consider the owner of the input data as a potential candidate. To clarify the role of this data, a technical sketch of the necessity and use of external input data is first required. The importance of this input data is indeed large, but problems with this option are soon encountered. The claim of the data-owner rests on an outdated romantic conception of originality and authorship, which is impossible to examine properly due to its mystifying definition, or rather lack thereof. Attempts to further define this notion of authorship and creativity are susceptible to logical contradictions. Additionally, to be consistent, this romantic concept would need to be applied to human authors too, leaving many human creators to lose their right to authorship.

I will look at the final option, entailing a contract to define authorship over the output of the AI, in 3.7. To begin with, the possibility of free transfers and alienation of property, both necessary for a contract, are discussed. It will be argued that in most labour theories, this possibility is present. Taking Nozick as an example, I arrive at consent as a requirement for justified transfer and alienation. However, it is impossible to get this consent in our case regarding AI, as there is no single party which has a full claim of authorship or ownership. Dropping the requirement of an author would be necessary to formulate a contract, but this would make it impossible to delineate the claim to ownership, as contributions from labour are spread along nearly infinite chains of causality.

I hope it will be clear that my aim in these chapters is not to formulate decisive conclusions regarding the questions to be treated in my arguments. Most of these questions have already supplied philosophers with years of debate, to which I do not pretend to have the answers. Instead, I tried to draw from these existing debates to reflect on a rather peculiar confrontation between a mature theory dealing with a timeless issue and a budding field of technology that may shake the world. This may tell us something about our existing beliefs and assumptions, and how they may

change in the face of new developments. Most of all, I hope this text will enhance the readers' understandings of property theory, and illuminate some of the challenges artificial intelligence will present in the future.

Chapter 2 - Labour theories of intellectual property

2.1 Introduction

There are many ways of arguing for property and intellectual property. Two large categories that take up a majority in this aggregation of theories are utilitarian or consequentialist theories on the one hand, and theories of labour and desert on the other. The differences between these two groups stem roughly from which element authors in both camps take as the justificatory element for property, which methods they adopt in their reasoning, and which forms of knowledge-production they find to be relevant.

While other thinkers such as Marx and Hegel have theorised about property too, formulating their own justifications and premises which diverge significantly from what will be discussed here (but are no doubt worth a discussion themselves), the conflict between consequentialist and labour theories is interesting when discussing creation by an autonomous AI, as it uncovers some of our basic intuitions about property and appropriation. The introduction of creative AI generates some problems which these theories may not be prepared for.

Consequentialist theories look at the consequences of property, and try to base property on the question of which distribution of goods has the best general outcome. One of the most popular and well known varieties of this kind of theories is utilitarianism. Generally, utilitarianism will produce the similar in relation to property theory. Examples of consequentialist justifications for property are

nearly ubiquitous in texts from authors hailing from a legal background.[2] These consequentialist theories stand in stark contrast to theories of labour and desert. In this second category, the consequences of distribution are not considered to be relevant. Instead, the principles that are fundamental to how a distribution came to be are considered. They follow a more deontological method: the consequences for those affected could be worse in a certain distribution than in others, but if this distribution was installed in a correct manner, following the right principles of property, it is the only right distribution.

Another way to illustrate this distinction is to define it in terms of pre-political and political rights. For consequentialist authors, a distribution of goods, or a rule according to which this distribution is made, is always subject to legitimate change if it turns out that there is a possibility of achieving better results. For them, property is part of a social contract: the method used to achieve a distribution of property is an agreement, not an inalterable principle. In other words: property is political, it is a social contract.

From a labour- or desert principle, property is often considered to be pre-political: it is not an agreement, but a right that applied already before societies and institutions were formed. While property is pre-political in this view, rights that are political are secondary. Thus, this static right of property can always be used as a counterargument to consequentialist objections to a certain distribution or principle of distribution. This is the reasoning used in theories of natural law, as in Locke's theory, for whom this may have been a very conscious decision.[3]

---

2 see for example Colin R. Davies, 'An Evolutionary Step in Intellectual Property Rights – Artificial Intelligence and Intellectual Property', *Computer Law & Security Review* 27, no. 6 (December 2011): 601–619, doi:10.1016/j.clsr.2011.09.006; Randall Davis, 'Intellectual Property and Software: The Assumptions Are Broken.' (Massachusetts Institute of Technology Artificial Intelligence Laboratory, 1991); Gönenç Gürkaynak et al., 'Questions of Intellectual Property in the Artificial Intelligence Realm', *Robotics Law Journal*, 2017; Elizabeth Rocha, 'Sophia: Exploring the Ways AI May Change Intellectual Property Protections Case Notes & Comments', *DePaul Journal of Art, Technology and Intellectual Property Law* 28, no. 2 (2018 2017): 126–146.

3 P. J. Kelly, *Locke's Second Treatise of Government: A Reader's Guide* (Continuum, 2007), 25.

Because of this decision, Locke needs to start his argument from a presupposed state of nature.[4] Another example of this position is Nozick, who states that property is a pre-political right, and that every form of state or government is to stay within the boundaries of this right:

> "Individuals have rights, and there are things no person or group may do to them (without violating their rights). So strong and far-reaching are these rights that they raise the question of what, if anything, the state and its officials may do."[5]

However, in the category of labour and desert theories of property, there exists a large variety of theories. The diversity is in fact so substantial that it may be problematic to put them in one category. There are nonetheless some good reasons to consider them together in this discussion.

First, the arguments used by the two kinds of theories – those of desert and those of labour – overlap in many aspects. Arguments of desert are sometimes used for theories of labour when the concept of labour as a value-producing activity is used for the justification of property rights. For this reason, there are authors, as for example Annis and Bohanon, who try to formulate theories of labour in terms of theories of desert.[6] I will come back to this in section 2.3.

Secondly, the creation of added value itself plays a large role in the definition of labour and the definition of desert. Theories of labour and desert cannot be reduced to one another completely, but they are therefore difficult to contemplate in total isolation of each other.[7]

---

4 Kelly, 26.

5 Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 17.

6 David B. Annis and Cecil E. Bohanon, 'Desert and Property Rights', *The Journal of Value Inquiry* 26, no. 4 (1992): 537–546.

7 Annis and Bohanon.

In this chapter, I will examine what is often considered to be the standard for labour-based justifications of private property, namely John Locke's theory. As this theory on its own is subject to many problems, and was never explicitly intended for intellectual property, it will mainly serve as a foundation for other theories. I will give an overview of some of the most important problems and solutions offered by other scholars. Subsequently, I will discuss the application of these theories to intellectual property.

2.2 Locke and natural rights

Ownership is a complicated cluster of claims, privileges and powers, and it is difficult to claim all of these rights to an object. A labour theory of property, in its basic form, may offer an answer to what Thomson calls the Acquisition Schema: "If a thing is unowned, then if X does alpha to it, X thereby comes to own it."[8] In any discussion of labour theories of property, it is difficult to avoid including John Locke. His version of the Acquisition Schema would involve labour as the activity that justifies property. Even though he was not the first to link property to labour, he has been by far the most influential advocate.[9] This despite his relatively brief treatment of the subject, which is limited to a part of Chapter V from the second book of his Two Treatises of Government.[10] It is this passage which has inspired all the following 'Lockean' theories of property.

The *Two Treatises*, as a whole, are an attack on the absolute monarchy, defended by Robert Filmer.[11] As a contemporary of Locke, he argued for a political system led by an absolute ruler. The territory

8 Judith Jarvis Thomson, *The Realm of Rights* (Harvard University Press, 1990), 324.

9 Peter Drahos, *A Philosophy of Intellectual Property* (Routledge, 2016), 47.

10 John Locke, 'The Second Treatise: An Essay Concerning the True Original, Extent, and End of Civil Government', in *Two Treatises of Government and A Letter Concerning Toleration*, ed. Ian Shapiro (Yale University Press, 1689), chap. 5, https://www.jstor.org/stable/j.ctt1npw0d.7.

11 James Tully, *A Discourse on Property: John Locke and His Adversaries* (Cambridge University Press, 1980), 54–55.

of a state would be his personal property. Any property rights enjoyed by his subjects would effectively be only a kind of ‘lease’ from the king, who could change the distribution of property at will.[12] The uncerlying political situation in 17th century England preceded the theoretical disagreement between Locke and Filmer. At the root of the discussion laid the very practical political issue of taxes, which the king wanted to raise without appealing to the parliament. It was Filmer who came to defend the king, while Locke opposed the taxes.

According to Filmer, every form of property originally emerged from the inheritance of Adam, who received ownership over the whole world from God. A king, according to Filmer, inherits the territory of a state from Adam indirectly. This means that the king owns everything inside this territory. Every distribution of property inside this territory is therefore based on nothing more than convention, which can be overruled at any point by the genuine claim of the monarch.

Hence, the primary intellectual opponent of Locke was Filmer. The motivation behind Locke’s arguments was at least partly political. The main issue in the *Two Treatises* is consequently not the justification of a certain theory of property, but rather the defence of democratic self-government in the face of a monarchical absolutism.[13] The discussion on property is only a side issue.

Locke reacts to Filmer by arguing that property is a pre-political right of every individual. A modification of the distribution of goods and land without agreement from the other parties is thereby a violation of this fundamental right.[14] To justify this claim, Locke claims God presents ownership of the world not to Adam, but to all people collectively.[15] From this, Locke infers that the world is the collective property of the whole world. However, this premise immediately presents

12 Kelly, *Locke’s Second Treatise of Government*, 61.

13 Gregory S. Alexander and Eduardo M. Peñalver, *An Introduction to Property Theory* (Cambridge University Press, 2012), 36.

14 Kelly, *Locke’s Second Treatise of Government*, 61–62.

15 Tully, *A Discourse on Property*, 60.

Locke with his greatest obstacle: how can personal property exist if the world is collectively owned by all people? The step from commons to personal property is what Locke needs to justify.

To achieve this task, Locke looks to the natural law theory. According to this theory, popular in Locke's time, God has built into creation a system of rights and duties. These rights and duties are rational, and thereby both accessible and binding to every person.[16] God is in a position to impose such laws, duties and rights, because he created mankind, and therefore owns it.[17] According to Locke, God only gave the world to mankind collectively with the single explicit goal of allowing them to adhere to these so called 'natural laws'. Because we are the property of God, we are not allowed to kill ourselves, we must keep ourselves alive, and we must strive to keep others alive if this is not in conflict with our own self-preservation.[18] The world and everything in it only serve this self-preservation: nature offers the food, shelter, resources, etc., that people need to stay alive. For Locke, this is the origin of the commons, which are the shared property of humankind.

For the next step, from commons to individual property, Locke applies his famous labour-mixing argument:

> "Though the earth, and all inferior creatures, be common to all men, yet every man has a property in his own person: this nobody has any right to but himself. The labour of his body, and the work of his hands, we may say, are properly his. Whatsoever then he removes out of the state that nature hath provided, and left it in, he hath mixed his labour with, and joined to it something that is his own, and thereby makes it his property. It being by him removed from the common state nature hath placed it in, it hath by this labour something annexed to it that excludes the common right of other men. For this labour being the unquestionable property of the labourer, no man but he

16 Alexander and Peñalver, *An Introduction to Property Theory*, 65.

17 Tully, *A Discourse on Property*, 40.

18 Kelly, *Locke's Second Treatise of Government*, 65.

> can have a right to what that is once joined to, at least where there is enough, and as good, left in common for others."[19]

He argues that in order to appropriate something from the commons, one must "mix" his labour with it. This only works because we are our own property, which entails that our labour is also our property, which, when mixed with something unappropriated produces individual private property.[20] In the same way that God's creation leads us to be His property, the product of our labour becomes our property. It is important to view this in the light of natural law; God included these rules into creation, Locke argues, because they are necessary for our survival. The idea behind this divine law seems to be that people can only thrive when they can have property.

This labour-mixing argument, crucial in Locke's theory of property, is the reason why Lockean theories are commonly called labour theories of property. The reasoning seems clear: labour takes a central role in Locke's concept of fair appropriation. Property can only come into existence through labour mixed with the commons.[21]

Provisos

Unmistakably, this theory in the form that I have described it so far would lead to much unfairness in practice. It is not difficult to imagine a few people ultimately appropriating all of the commons, leaving nothing for others. Or someone could appropriate all the food, and then let it go to waste. In order to prevent this, Locke adds some conditions that need to be fulfilled in order for an

---

19 Locke, 'The Second Treatise' § 27.

20 Drahos, *A Philosophy of Intellectual Property*, 50.

21 Taking this further, one could argue that the existence of the commons, too, is dependent upon labour. The commons only exist because they are necessary to fulfil the duty of self-preservation installed by God. This duty could have only been installed through God's creation. If the process behind creation could be described as God's 'labour', a difficult but not entirely inconceivable assertion about a being that is all-powerful, the existence of the commons would be dependent upon this labour.

appropriation to be justified. The first of these provisos requires everyone to leave ‘enough and as good’ for others. This renders appropriation of all resources by a few impossible. The second proviso requires that appropriation happens without spoilage. People cannot appropriate more than they will need, as this would result in part of it going to waste. God gave the world to mankind in common so that they could use it for their self-preservation, not for them to squander the resources.

Some commentators hold that Locke describes a third proviso: the ‘charity proviso’.[22] The speculations of this extra proviso originate from a passage in the First Treatise, that according to some is often overlooked:

> “But we know God hath not left one man so to the mercy of another, that he may starve him if he please: God, the Lord and Father of all, has given no one of his children such a property in his peculiar portion of the things of this world, but that he has given his needy brother a right to the surplusage of his goods; so that it cannot justly be denied him, when his pressing wants call for it: and therefore no man could ever have a just power over the life of another by right of property in land or possessions; since it would always be a sin, in any man of estate, to let his brother perish for want of affording him relief out of his plenty. As justice gives every man a title to the product of his honest industry, and the fair acquisitions of his ancestors descended to him; so charity gives every man a title to so much out of another’s plenty as will keep him from extreme want, where he has no means to subsist otherwise [...]”[23]

This fragment seems to imply that Locke deems it necessary for any person with private property to provide aid to others when they are in need. Yet, in the formulation of his theory in the *Second*

22 Drahos, *A Philosophy of Intellectual Property*, 40; Robert P. Merges, ‘Locke’, in *Justifying Intellectual Property* (Harvard University Press, 2011), 31–67, doi:10.4159/harvard.9780674061125.

23 John Locke, ‘The First Treatise: The False Principles and Foundation of Sir Robert Filmer’, in *Two Treatises of Government and A Letter Concerning Toleration*, ed. Ian Shapiro (Yale University Press, 1689), 7–99, https://www.jstor.org/stable/j.ctt1npw0d.6 § 43.

Treatise, Locke does not repeat this condition, nor does he explicitly acknowledge it as a proviso, leaving the meaning of this passage open to interpretation to some degree. If it was earnestly written, and not meant as a rhetorical addition, there seems to be no other option than to consider it a third proviso. However, as most commentators have disregarded this passage, and hence only a few interpretations include it, I cannot regard it as a necessary component in a labour theory of property.

To summarise, let me list Locke's six core propositions, as formulated by Drahos[24]:

1. God has given the world to people in common.
2. Every person has a property in his own person.
3. A person's labour belongs to him.
4. Whenever a person mixes his labour with something in the commons he thereby makes it his property.
5. The right of property is conditional upon a person leaving in the commons enough and as good for the other commoners.
6. A person cannot take more out of the commons than they can use to advantage.

And a contested seventh proposition:

7. Any person in need has a right to a share of the property of another who has enough.

24 Drahos, *A Philosophy of Intellectual Property*, 50.

2.2.1 Shortcomings of Locke's theory

Locke's theory of property has been as popular as it has been controversial. Almost every part of it has been subject to criticism by contemporary philosophers.[25] This has not led to the complete abandonment of the theory, but instead has provoked many authors to propose new interpretations in its defence. In this section, I will go over some of the more important problems that have been identified.

Some of the most famous arguments made against Locke's theory of property concern the 'mixing labour' metaphor central to his argument. Fundamentally, Waldron argues, the metaphor of mixing labour with a physical part of the unowned commons is simply a category mistake.[26] Taking the relation of land, for example, to the labour that is performed upon it, 'mixing' is not a valid description. Land and labour belong in different categories, and cannot be mixed. Another problem with the metaphor is that it begs the question how much labour is necessary for appropriation, and how much property results from it.[27] While for Hettinger a property right based on a minimal amount of labour is justified[28], this is not sufficient for Nozick.[29] He asks why mixing one's labour with something unowned creates property instead of being no more than a loss of labour:

> "But why isn't mixing what I own with what I don't own a way of losing what I own rather than a way of gaining what I don't? If I own a can of tomato juice and spill it in the sea so that its molecules (made radioactive, so I can check this) mingle evenly throughout the sea, do I thereby come to own the sea, or have I foolishly dissipated my tomato juice?"[30]

---

25 Adam Mossoff, 'Saving Locke from Marx: The Labor Theory of Value in Intellectual Property Theory', *Social Philosophy & Policy; Oxford* 29, no. 2 (July 2012): 283–317, doi:http://dx.doi.org/10.1017/S0265052511000288.

26 Jeremy Waldron, *The Right to Private Property* (Clarendon Press, 1990), 184–189.

27 Benjamin G. Damstedt, 'Limiting Locke: A Natural Law Justification for the Fair Use Doctrine Note', *Yale Law Journal* 112, no. 5 (2003 2002): 1179–1222.

28 Edwin C. Hettinger, 'Justifying Intellectual Property', *Philosophy & Public Affairs* 18, no. 1 (1989): 31–52.

29 Nozick, *Anarchy, State, and Utopia*, 174–175.

30 Nozick, 174–175.

Not only does Nozick ask how much labour is required for property to result from it, but he also asks how this property could be measured, and therefore appropriately sized. When a farmer works the land, Locke's favourite example of appropriation by labour, he would be entitled to all of the land. But, as Alan Ryan wants to know, when someone plucks an apple from a tree, does he now own the entire tree? No, Ryan concludes.[31] And yet the two situations do not seem to be very different.

While Damstedt interprets Nozick's argument mainly in terms of the required amount of labour and the size of the resulting property[32], it can be taken to imply that the connection between labour and property is completely absent. As Judith Jarvis Thomson remarks, property is an arbitrary reward for labour: "Why not instead a medal and a handshake from the president?"[33]. Especially in a secular context, where Locke's premises of God's divine right are no longer self-evident, the gap between labour and property widens. The labour-mixing argument seems to take an unwarranted step, and cannot bridge the gap by itself. As Becker puts it:

> "Insofar as one's labor is inseparable (by way of ownership rights) from one's body, it is understandable how the first 'extension' - from ownership of the body to ownership of the labor - is warranted. But the same can hardly be said for the second extension - from ownership of the labor to ownership of labor's products."[34]

Even with the religious arguments left in place, Waldron argues, they are not enough: "at most that explains only why men labour in the first place […]; it does not explain why he continues to be

31 Alan Ryan, *Property and Political Theory* (Oxford: Basil Blackwell, 1985), 32–35.

32 Damstedt, 'Limiting Locke'.

33 Thomson, *The Realm of Rights*, 324–325.

34 Lawrence C. Becker, 'The Labor Theory of Property Acquisition', *The Journal of Philosophy* 73, no. 18 (1976): 659, doi:10.2307/2025823.

entitled to what he has laid out. […] The mere fact of His command does not indicate that obedience is not a way of losing one's labour."[35] An alternative justification needs to be found.

Before I begin to discuss some interpretations scholars have formulated to work around these fundamental problems, it is worth mentioning that the difficulties do not end here. Another point that has received much attention, much of it criticism, is the addition of Locke's provisos. The 'enough and as good' proviso, in a strict reading, would make appropriation altogether impossible, even in a state of nature: "Even if we consider that in a hypothetical initial position there are very few people and a very large world, the takings of one person, even though leaving plenty for others to acquire, would always leave the others with less than they could have acquired in the first instance."[36] Even if we were to interpret this constraint in egalitarian terms, demanding that every person receives an equal share, we would face a seemingly unsolvable problem of indeterminacy. How can we measure equal shares of land and resources? Furthermore, it would be impossible to uphold the same requirement for future generations.[37] An interpretation that could work, according to Kelly, would require from an appropriator not to leave enough for others to appropriate something of exactly the same value, but instead that others would still be able to appropriate according to their needs of subsistence.

2.3 Labour theories after Locke

As we have seen, Locke's theory of property by itself entails many problems, most of all a justificatory gap between labour and property. Still, intuitively, a strong connection exists between the two. And even though when taken in isolation it has been heavily discredited ever since its formulation, Locke's theory offers plenty of room for interpretations and adjustments to serve as the basis for many other derivative theories of property.

35 Waldron, *The Right to Private Property*, 190.

36 Kelly, *Locke's Second Treatise of Government*, 75.

37 Kelly, 75.

2.3.1 Libertarian theories of property

Nozick, despite his disagreements with Locke, supports a labour-based theory of property. In a way, Nozick's treatment of the subject can be seen as a next step in the same tradition. Note that I do not call it a theory explicitly. This is due to the fact that Nozick's writings on property, although extensive, are difficult to condense into a coherent theory. As Waldron puts it: "On his own account he does not offer a full theory of private property, only a sketch of what the form of such a theory might be."[38] Perhaps the title alone of Barbara Fried's paper "Does Nozick have a theory of property rights?" presents an ample demonstration of the confusion about the matter.[39]

Leaving aside this problem, Nozick's writings have surely had a large influence on property theory, and can with no doubt be placed within a libertarian doctrine. In fact, as Fried notes, "he has at least three mutually inconsistent theories: utilitarianism; Lockean libertarianism; and anything goes".[40] I will try to summarise the Lockean libertarian interpretation of Nozick here. As a starting point for property, Nozick rejects Locke's natural rights doctrine, supplanting it instead with a strictly negative community in which people merely have duties of non-interference to each other. Instead of the entire world being common property of humanity, in this view everything is originally unowned.[41] Thus, the starting principle can be identified as the rights of individuals, and most of all the right to liberty in terms of non-interference.[42] Explicit individual consent is necessary for anyone to limit someone else in their liberty. Crucial for a justified appropriation is the question of whether or not others are harmed by it. If the freedom of others is respected, Nozick argues, one cannot take

38 Waldron, *The Right to Private Property*, 15.

39 Barbara Fried, 'Does Nozick Have a Theory of Property Rights?', in *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*, ed. Ralf M. Bader and John Meadowcroft (Cambridge University Press, 2011), 230–252.

40 Fried.

41 Alexander and Peñalver, *An Introduction to Property Theory*, 53.

42 Waldron, *The Right to Private Property*, 128.

away the freedom of people to appropriate things, and to pass that property on to others as they wish. This follows the classic libertarian case for property, which argues that a just state will only regulate property by protecting the owner's rights.

Additionally, Nozick holds that appropriation will not make non-owners worse off.[43] In fact, in many cases, appropriating something will lead to the situation of non-owners to be improved, as it allows for the exploitation of resources, necessary for production of valuable and useful goods. This, according to Nozick, could bridge the gap between labour and property.

2.3.2 Desert theories of property

Another interpretation, which by many authors is seen as one bearing a lot of intuitive appeal, is based on desert.[44] It also finds a lot of backing in Locke's original writings:

> "He that had as good left for his improvement as was already taken up, needed not complain, ought not to meddle with what was already improved by another's Labour; If he did, 'tis plain he desired the benefit of another's Pains, which he had no right to..."[45]

In this approach, which I will call the desert theory from here onwards, labour is seen as a morally good action, so it deserves a reward. This is desert in what Waldron calls a 'strong moralistic sense',

---

43 Alexander and Peñalver, *An Introduction to Property Theory*, 54.

44 "Desert" here refers to the worthiness of a reward due to morally good behaviours or characteristics. See Daniel Haas, 'Merit, Fit, and Basic Desert', *Philosophical Explorations* 16, no. 2 (June 2013): 226–239, doi:10.1080/13869795.2013.786971. for a more thorough discussion and definition of the concept. He identifies two common uses of the term: a compatibilist version, defining desert "as a 'fit' between our moral responsibility judgments and properties of the agent and her action such that it would be fitting, or appropriate, to claim that the agent is blame/praise-worthy.", and an alternative version linked to "manipulation arguments", understanding desert "as a kind of 'merit' such that it is fair to blame/praise the agent precisely because they are worthy of blame or praise given that they performed the act in question and were sensitive to its moral status." (p. 226)

45 Locke, 'The Second Treatise', II, Ch V, §34.

meaning that ‘A deserves X’ is not just equal to the statement ‘A is entitled to X’, but instead serves as an argument for the latter statement.[46] Labour, then, would be the activity that creates desert. To fit in a Lockean framework, labour would have to be necessary and sufficient for a desert-based claim to property. This is not difficult to imagine, as Locke describes labouring both as an act of direct obedience to God’s command[47], and as an activity which produces a useful contribution to the wealth and prosperity of mankind[48], which seems plenty to argue for its moral value.[49] Becker summarises the desert theory in the following way[50]:

> (1) When the labour is beyond what morality requires, deserving of some benefit,
> (2) When the benefits deserved are proportional to the values produced;
> (3) When nothing but property rights in the things produced can be considered a fitting benefit;
> (4) Then the property rights are deserved.

This formulation of the theory, however, leaves little to be justifiably appropriated, as Annis and Bohanon note[51], with Becker himself quickly offering the same criticism, noting that it might be applicable to a small rock collection, but not much more.[52] The big problem in fulfilling these demands stems from the fittingness requirement (3): arguing that property is the only proper reward can be difficult. A less strict version comes from John Stuart Mill, who replaces this fittingness requirement by a “why not?” argument. The full reasoning, again summarised by Becker, is summarised as follows[53]:

---

46 Waldron, *The Right to Private Property*, 201–202.

47 see Locke, ‘The First Treatise’ § 45-46 and § 32-35.

48 see Locke, ‘The Second Treatise’ § 36-37.

49 Waldron, *The Right to Private Property*, 202–203.

50 Lawrence C. Becker, ‘Property Rights: Philosophic Foundations’, *Books by Hollins Faculty and Staff*, January 1977, 53–54, https://digitalcommons.hollins.edu/facbooks/76.

51 Annis and Bohanon, ‘Desert and Property Rights’.

52 Becker, ‘Property Rights’, 56.

53 Becker, ‘The Labor Theory of Property Acquisition’.

(1) When the labour is beyond what morality requires,
(2) When it produces something that would not have existed otherwise,
(3) When its product is something that others lose nothing by being excluded from,
(4) Then it is not wrong for producers to exclude others from the product of their labour

Either of these formulations seem sufficient to offer at least some property rights to original appropriators. Annis and Bohanon attempt to expand on the first formulation by proposing a theory that considers both desert and social welfare factors.[54] They argue that recognition desert, the things we deserve based only on the recognition of our personhood, warrants a very weak ownership claim. Adding appraisal desert, which is applicable to cases that go beyond what is morally required, justifies according to them a reward for labour. They conclude that this reward must be property, depending on the situation and the relevant social welfare factors, which include "items as whether others are denied significant opportunities, whether their freedom is restricted, and general welfare issues."[55] In short, this seems to be a combination of both formulations of the Desert Theory, with some libertarian values mixed in.

Having finished the considerations of some of the main justifications of property based on Locke's labour theory, we can move on to applying these to intellectual property. There are of course many more versions of a labour-based argument, but I consider those mentioned to be the most important ones, as they not only bear a lot of intuitive appeal, but they are also often discussed by authors considering the subject of property.[56]

---

54 Annis and Bohanon, 'Desert and Property Rights'.

55 Annis and Bohanon.

56 As an example, see Alexander and Peñalver, *An Introduction to Property Theory* part I, chapter 2; Lawrence C. Becker, 'Too Much Property', ed. Jeremy Waldron and Stephen A. Munzer, *Philosophy & Public Affairs* 21, no. 2 (1992): 196–206; Bryan Cwik, 'Labor as the Basis for Intellectual Property Rights', *Ethical Theory and Moral Practice* 17, no. 4 (August 2014): 681–695, doi:10.1007/s10677-013-9471-y; Waldron, *The Right to Private Property*. While there is no full consensus about which interpretations are the most valuable, the ones I have

2.4 Labour theories of intellectual property

Applying the concept of property to ideas and abstract objects is not what Locke, his contemporaries, or his predecessors usually had in mind when they wrote on the subject. Yet, since then, it has become increasingly common to do just that. Locke's theory, in particular, has been a popular one to adapt to intellectual property. Before I can discuss the application of a labour theory of property to ideas, there are some preliminaries that need to be settled. First of all, it is necessary to consider what the subject of intellectual property includes.

2.4.1 The subject of intellectual property

While the term 'intellectual property' is common in contemporary debate, it encompasses a wide range of subjects. In law, the distinction is made between copyrights, patents, trademarks, trade secrets, industrial design rights, trade dress, plant variety rights, etc. Their validity depends on the laws of each country, but there is generally a lot of continuity to be found.[57] The two most important and most widely recognized categories are copyrights and patents. A copyright offers protection only against copying, whereas a patent offers protection against both copying and independent creation of the same work.[58] In short, copyright is seen as applicable to traditional cultural works, such as literature and artworks, while patents apply to useful inventions.

discussed here are present in the majority of contemporary debate, and are seen as interpretations with a strong intuitive appeal.

57 Justin Hughes, 'The Philosophy of Intellectual Property', *Georgetown Law Journal* 77, no. 2 (1989 1988): 287–366.

58 Drahos, *A Philosophy of Intellectual Property*, 9.

Even though in a law context it is impossible to fully equate these different kinds of intellectual property, from a philosophical perspective they share an important fundamental character: they are all rights to abstract objects[59], what some would call 'ideas'[60].

In law, intellectual property claims must fulfil some requirements. Aside from requirements such as the presence of an "invention" pertaining to some field of technology, the possibility of industrial application, and the involvement of an inventive step, European patent law, to take an example, requires from inventors a proof of novelty and non-obviousness.[61] In other words, the invention must be at least sufficiently new and not a trivial achievement in consideration of already existing knowledge.[62] While some philosophical theories of intellectual property will try to include these legal requirements, others will reject them. Lockean theories run into some problems with these legal requirements, as it can be difficult to argue that they follow from pre-political property rights.

2.4.2 The case for Lockean theories of intellectual property

Special features of intellectual property

Substituting material objects for abstract objects and reapplying the same theory of property seems simple enough. But there are some key differences between the two kinds of objects that make this problematic. First, information consumption is non-rivalrous.[63] In other words, information cannot be overconsumed. It can be copied either through writing, or simply through education. Abstract objects are not limited by physical constraints at all. Contemporary digital means of transferring and

59 Making a distinction between material and abstract objects can lead to many metaphysical problems, the treatment of which is beyond the scope of this text.

60 Drahos, *A Philosophy of Intellectual Property*, 9.

61 European Patent Office, 'Guidelines for Examination: Patentability', European Patent Office, accessed 23 December 2020, https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_i_1.htm.

62 Alexander and Peñalver, *An Introduction to Property Theory*, 186.

63 Peter Lewin, 'Creativity or Coercion: Alternative Perspectives on Rights to Intellectual Property', *Journal of Business Ethics* 71, no. 4 (April 2007): 446–447, doi:10.1007/s10551-006-9150-1.

copying information can magnify this with higher speed, greater ease, further reach, and lower costs. Once an idea, in the form of text, a design, a formula, digital information, or other forms, is initially conceived, it can be nearly infinitely distributed with little to no cost, and without losing the original abstract object 'held' by the original creator.[64]

Considering the possibility of infinitely copying abstract objects, a theory of intellectual property will have to be argued for in a manner different from the theories of material property. The point of intellectual property rights (IPR's) is not to protect an owner from losing an object, but most of all to protect people from losing the advantage – usually in an economic sense – that their original idea has given them. Intellectual property rights introduce artificial scarcity for the benefit of the owners.[65] So it is this artificial scarcity that needs to be justified in a theory of intellectual property.[66]

A second feature that puts intellectual property apart from material property is the difference in rights they both claim. While a normal property right mainly requires rights to exclusivity, in Lockean terms based on the fruits of one's own labour, Hettinger argues that intellectual property in most cases also requires the right for an inventor or writer to "make her work public, sell it in a market, and then prevent others from making copies".[67] While the right to the benefit of the fruits of one's labour are already argued for in a Lockean theory of property, the same does not go for these

64 Lewin, 445–446.

65 Consequentialist theories have a second reason to introduce scarcity, namely to provide a long term benefit to society in general by providing incentive for creativity and invention. As labour theories cannot use this argument, they have to rely on the benefit for the owners themselves. For more on this, see Lawrence C. Becker, 'Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory', *Chicago-Kent Law Review* 68, no. 2 (1993 1992): 615–617.

66 Becker, 616.

67 Hettinger, 'Justifying Intellectual Property', 40 Note that this is not always the case. Some patented technologies are not sold, or are not allowed to be sold.

extra rights. This, according to Hettinger, creates an argumentative "gap", requiring additional justification.[68]

(Dis)advantages of a labour theory

A classic utilitarian reasoning for intellectual property is based on incentive.[69] Developing a drug, for example, could require large investments. If the company that initially developed it has no protection from other companies copying their formula, the initial investments would be unlikely to be earned back, and thus no one would be willing to make the investments in the first place. This would constitute a net loss for society at a whole. The requirements of novelty, non-obviousness and originality in patent and copyright law make a lot of sense when following this type of justification. As the goal is to create an incentive for invention and creativity that would otherwise not occur, law can be limited to works that require some significant effort.[70] Similarly, the limits to IPR protection can be justified very easily using a utilitarian theory. The cost to society that a monopoly resulting from IPR's generates must be balanced. Therefore, copyrights and patents are limited in time. Patents expire after 20 years, while copyright survives anywhere from 70 to 95 years after the death of the author. And the fair use doctrine allows copyrighted works to be copied more easily for purposes such as education, news reporting, parody, and other similar uses.[71] Considering all this, the current way law deals with intellectual property, and many of our intuitions surrounding it, all seem to speak in favour of a utilitarian theory of intellectual property, Becker argues. Utilitarian arguments can better explain and justify the details of existing property rights practices.[72]

---

68 This is a different "gap" in reasoning from the one described earlier, in 2.2. While the first gap was about the general difficulty of connecting property to labour, this gap concerns the justificatory space between ownership rights in general, and special intellectual property rights that allow creators or inventors to prevent others from copying it while making it public and 'selling' the abstract object themselves. Combining the two to make an even bigger gap seems possible, but treating them separately allows for more clarity.

69 Alexander and Peñalver, *An Introduction to Property Theory*, 184–185.

70 Alexander and Peñalver, 186.

71 Alexander and Peñalver, 187.

72 Becker, 'Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory', 620.

A labour theory of property, however, will have to find another way to argue for intellectual property. And, if it wishes to do so, it will need to find another way to explain the legal requirements and limits of IPR protection. Nozick, for example, holds that this time limit on patents follows from his concept of appropriation.[73] Other libertarian authors, however, dispute this, claiming that intellectual property rights, enforced by the state, constitute an infringement of liberty and tangible property rights.[74] Tom Bell, for example, argues that "the law of copyrights and patents violates the very rights that Locke defended".[75]

This internal conflict in libertarianism is not unreasonable. The fact that abstract objects are non-rivalrous poses a problem for Lockean theories. As Shiffrin argues, Locke's theory of property only warrants appropriation when private ownership is necessary.[76] An interpretation of Locke starting from labour instead of self-preservation may of course lead to a different conclusion. Altogether, as Alexander and Peñalver argue, "for Locke – the problem of developing a system of private ownership is significantly less urgent than for tangible property".[77]

However, there are also advantages to adapting a Lockean theory to intellectual property. The Lockean proviso are easy to meet for intellectual property, as "enough and as good" and non-spoilage seem to apply automatically to abstract objects.[78] As intellectual products can be infinitely copied, there is always enough left for others. Neither do abstract objects seem to expire. As Becker argues, "scarcity will be a function of artifice: withholding or restricting communication".[79] The

---

73 Nozick, *Anarchy, State, and Utopia*, 181–182.

74 Alexander and Peñalver, *An Introduction to Property Theory*, 192.

75 Tom W. Bell, 'Indelicate Imbalancing in Copyright and Patent Law', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, May 2007), https://papers.ssrn.com/abstract=984085.

76 Seana Valentine Shiffrin, 'Lockean Theories of Intellectual Property', in *New Essays in the Political Theory of Property*, ed. Stephen R. Munzer (Cambridge university press, 2001), 144–154.

77 Alexander and Peñalver, *An Introduction to Property Theory*, 193.

78 Shiffrin, 'Lockean Theories of Intellectual Property', 139–140.

79 Becker, 'Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory', 616.

way property rights are created may also speak in favour of a Lockean theory. An original idea could be described as crafted from intellectual labour mixed with something unowned, formed using nothing but our own minds and memories.[80]

The originality doctrine of copyright could also be justified from a Lockean point of view. As "Lockean labour theory provides support for the copyrightability of intellectual products involving only labour on the part of the author, as opposed to some broader element of creativity" [81]. If artistic creation is seen as labour rather than something that comes naturally, originality would imply labour, and could thus be seen as a legitimate threshold for ownership.

So, the case for a Lockean theory of intellectual property certainly does not seem to be doomed from the start. While there are certain problems to deal with, there are various advantages to be enjoyed as well. Many commentators even consider a Lockean labour theory of property to be more defensible in its application to intellectual property than when applied to physical property[82]. Having settled this, an important question still remains: which interpretation of Locke is most suited to face the challenge of autonomous creative AI?

2.4.3 Deciding on a labour theory of property

Taking Locke as a starting point, one could formulate a theory of intellectual property that would allow for nearly all abstract objects to become individual property. This so-called 'strong' theory,

80 Adam D. Moore, 'A Lockean Theory of Intellectual Property', *Hamline Law Review* 21, no. 1 (1998 1997): 65–108.

81 Carys J. Craig, 'Locke, Labour and Limiting the Author's Right: A Warning against a Lockean Approach to Copyright Law', *Queen's Law Journal* 28, no. 1 (2003 2002): 17.

82 See for example Hughes, 'The Philosophy of Intellectual Property', 296–297., Becker, 'Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory', 609., Shiffrin, 'Lockean Theories of Intellectual Property', 138–139.

Drahos argues, would need to make several assumptions[83]: (i) it would need to ignore Locke's original religious metaphysical justification of self-preservation, (ii) it would need to interpret abstract objects, both creative work and invention, as the product of intellectual labour, and (iii) it would need to decide for or against the existence of intellectual commons.

In other words, assumption (i) means that the theory would be based purely on an argument of labour and self-ownership, omitting the duty of self-preservation that accompanies a religious interpretation. This solves the problem of secularisation, but means that the justificatory gap established in section 2.2 still needs to be argued from another set of premises. The metaphysical assumptions that are set aside in (i) are also swiftly replaced by (ii), namely the proposition that abstract objects are produced by labour, or that human labour is capable of producing them.[84] While an alternative to this, such as platonism, is perhaps even more problematic, the premise that abstract objects are created by labour is not an uncontroversial assumption.[85]

To replace the original argument from self-preservation, there are several interpretations that can be applied. The libertarian and desert interpretations of material property discussed in 2.3 can be reused here, but need some adaptations in order to deal with the special nature of abstract objects.

(a) Argument from freedom

One existing argument for private intellectual property is based on the concept of freedom, an argument associated mainly with libertarian thinkers. Similarly to how the argument would be applied to material property, a libertarian theory would assert that "some measure of acquisitiveness among humans is inevitable", and that restricting this would require unjustifiable infringements of

83 Drahos, *A Philosophy of Intellectual Property*, 57–63.

84 Hughes, 'The Philosophy of Intellectual Property', 310–314.

85 Drahos, *A Philosophy of Intellectual Property*, 59–60.

the right to liberty, while the state should exist only to protect this property and the liberties of everyone.[86]

As an attempt to justify material property, such a theory could hold some sway. As a theory of intellectual property, however, it encounters many problems from the start. The most important problem comes down to different interpretations of the extent of this right to freedom. As intellectual property goes much further in the duties it demands from non-owners towards owners, not only requiring exclusive use of a produced object but also insisting on prohibiting any attempts at copying the original (abstract) object (see 2.4), it might be limiting freedom more than it would be preserving it.[87]

It is a very different kind of limit to freedom than the kind that material property implies, in the sense that it does not stem from a lack of access to certain material objects. Attas compares it to a ban on smoking: "You may have the space, time and tobacco, but you will still be prohibited from using them in the specified ways."[88] A theory of material property can justify a limit on the freedom of non-owners to use an appropriated object by contending that exclusivity is a necessary condition for the effective use of the object. It could even be argued that "property rights in tangible objects do not restrict liberty at all - they simply restrain action. Intellectual property rights, on the other hand, do restrict liberty".[89] Therefore, from a libertarian point of view, intellectual property enforced in the way in which it commonly is today seems hard to justify. Protecting abstract objects in the name of liberty only restricts liberty, much more severely so than other property rights do. All things considered, intellectual property rights seem incompatible with libertarian premises.

---

86 Lawrence C. Becker, 'The Moral Basis of Property Rights', *Nomos* 22 (1980): 193–194.

87 Tom G. Palmer, 'Are Patents and Copyrights Morally Justified - the Philosophy of Property Rights and Ideal Objects', *Harvard Journal of Law & Public Policy* 13, no. 3 (1990): 833–831.

88 Daniel Attas, 'Lockean Justifications of Intellectual Property', in *Intellectual Property and Theories of Justice*, ed. Axel Gosseries, Alain Marciano, and Alain Strowel (London: Palgrave Macmillan UK, 2008), 34, doi:10.1057/978-0-230-58239-2_2.

89 Palmer, 'Are Patents and Copyrights Morally Justified - the Philosophy of Property Rights and Ideal Objects', 831.

(b) Mill and the no hardship argument

Another way to attempt justifying intellectual property would be through the Locke-Mill version of labour theory. This is John Stuart Mill's interpretation of Locke's theory, and, according to Becker, most likely the closest to what Locke had in mind himself.[90] Mill states this without much elaboration: "It is no hardship to any one, to be excluded from what others have produced: they were not bound to produce it for his use, and he loses nothing by not sharing in what otherwise would not have existed at all."[91] In a more expounded manner, the argument is as follows: when someone (i) produces something through labour that would not have been created otherwise, (ii) beyond what is morally required, (iii) without producing a loss for others when disallowing them access to the product, then (iv) appropriation of the product by the labourer, by excluding others from it, is not wrong.[92] This does not amount to desert quite yet, but rather puts forward the claim that those who did not contribute any labour have no right to the product whatsoever, while those who did might as well claim it. In essence, it comes down to an argument from "why not?".[93]

Property in abstract objects, however, does seem to imply some kind of loss. While someone could appropriate a plot of land and leave others to do the same with another plot of land, once an idea is appropriated, it can no longer be freely copied. It could be argued that others suffer no real loss by being excluded from use of an abstract object, as their material situation after the appropriation remains exactly the same as before.[94] Nozick pleads a similar case, contending that "(a)n inventor's patent does not deprive others of an object which would not exist if not for the inventor."[95] Waldron

---

90 Becker, 'Property Rights', 41.

91 John Stuart Mill, *Principles of Political Economy with Some of Their Applications to Social Philosophy*, ed. John M. Robson, The Collected Works of John Stuart Mill, vol. 2, 33 vols (Toronto: University of Toronto Press, Routledge and Kegan Paul, 1965), http://oll.libertyfund.org/title/102 Book II, Ch. 2, §6.

92 Becker, 'Property Rights', 41.

93 Becker, 'The Moral Basis of Property Rights', 193.

94 Jeremy Waldron, 'From Authors to Copiers: Individual Rights and Social Values in Intellectual Property Symposium on Intellectual Property Law Theory', *Chicago-Kent Law Review* 68, no. 2 (1993 1992): 866.

95 Nozick, *Anarchy, State, and Utopia*, 182.

proposes a counterexample of Q suffering a deadly disease, and being excluded from the cure discovered by P, who only shares the drug with friends. But while Q will die either way, "clearly Q will suffer *something* as a result of this".[96] There might be no loss in strictly material terms, but from a moral perspective there is suffering. And other examples are easy to imagine. When someone develops an idea into a business, claiming intellectual property rights over the concept, others are no longer free to walk the same path. Intellectual property *necessarily* implies some loss for others.

While this discussion could clearly be expanded upon, justifying ownership of abstract objects by arguing that it entails no losses or hardships for others seems increasingly implausible. The moral suffering that intellectual property can lead to suggests that this is not the way to proceed.

Desert justifications

A recurring argument for intellectual property rights has been the argument from desert, also called the desert-for-labour argument [97]. Even though in essence the force of the argument in every instance comes from moral desert, there are in fact many ways in which an argument pertaining to this category can be construed. These different versions of the argument depend on the source of the desert. There are three kinds of desert that are commonly used in the justification of intellectual property rights: desert from excellence, desert for value, and desert from need.

96 Waldron, 'From Authors to Copiers', 866.

97 Richard A. Spinello, 'The Future of Intellectual Property', *Ethics and Information Technology* 5, no. 1 (March 2003): 1–16, doi:10.1023/A:1024976203396.

(c) Desert from excellence

Arguments supported by desert from excellence start from the claim that the labour necessary for developing abstract objects suitable for intellectual property rights goes beyond what is morally required from the labourer. Accordingly, this moral surplus needs to be rewarded, and the suitable reward for it consists of property rights to the product.

The moral excellence of labour can be derived from an avoidance principle: labour being work, people generally are keen to avoid it as much as possible.[98] This view of labour, seen as an unpleasant and undesirable activity, echoes in Locke's description of labour as "pains".[99] Connecting this sentiment to desert seems simple: "The proposal is that labor is something unpleasant enough so that people do it only in the expectation of benefits (and since unlaboured-on things are of little or no value anyway), it would be unjust not to let people have the benefits they take pains to get."[100] This description of the argument, however, could be interpreted as a consequentialist or pragmatic formulation, if it is interpreted to be based on incentive. A strict labour theory cannot be fully based on incentive, as it would therefore be concerned with consequences rather than principles of appropriation. Hence, the process of deriving desert from labour should be interpreted without a dependence on the creation of incentive.

(d) Desert for value

Intellectual labour, similarly to material labour, is often considered to be creating valuable things. Take for example the maxim in British copyright law, stating that "what is worth copying is prima facie worth protecting".[101] By virtue of this intuitive appeal, creation of value is regularly used as an argument for the justification of intellectual property.

---

98 Hughes, 'The Philosophy of Intellectual Property', 302.

99 Locke, 'The Second Treatise', §34.

100 Becker, 'The Labor Theory of Property Acquisition', 655.

101 University of London Press, Ltd. v. University Tutorial Press, Ltd. (2 Ch D 601. 1916).

The argument is similar to other desert arguments: “People deserve compensation for creative activities that add value, and the most appropriate compensation is the right to take temporary ownership of that creation so that the compensation received is proportionate to the value placed on the creative work by society.”[102] The argument is so similar to a reasoning of desert from excellence, that it could be considered nearly the same. And indeed, value creation could be subsumed under a theory of desert from excellence when value creation is seen as a behaviour that excels moral expectations. When instead the value creation is taken as a justifying property by itself, with no reference to moral excellence, it can be considered a separate category of desert. An argument based on value creation applied to abstract objects seems to have the benefit of offering a more persuasive answer to problems of delineation than when applied to tangible property. It is more plausible to say in a context of intellectual property that labour has created most of the value in the final product.[103]

(e) Desert from need

A final version of the desert argument, based on need, is much less discussed in literature on the subject, despite its widespread intuitive appeal. This version is based on the possible special needs that (intellectual property) labourers could develop, which could be accommodated by awarding property rights to the product of their labour.

It is again Becker who describes this kind of desert, and who links it to an argument of “identity-dependence”.[104] This identity-dependence entails that the integrity or welfare of a labourer is sustained by their identity, which has become dependent on the product of their labour. When this dependence creates a basic personal need that is sustained by social norms, either additional social norms to help the labourer in this need should be created, or the original social norms sustaining this

102 Spinello, ‘The Future of Intellectual Property’, 11.

103 Alexander and Peñalver, *An Introduction to Property Theory*, 194.

104 Becker, ‘Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory’, 626–627.

need should be changed.[105] Ownership could be a social norm fit to help meet the needs of the labourer.

The argument from need by itself seems insufficient to serve as a Lockean labour justification of ownership. Considered in isolation, this reasoning seems to amount to no more than another pragmatic consequentialist scheme to maximise welfare. Combined with one of the other versions of desert-arguments, however, it could serve an important role. The main problem for desert justifications of property is the problem of appropriate reward. While the claim that labour deserves a reward is by itself very defensible, asserting that this reward should be property rights, and that property rights are thus the only appropriate reward, is far more problematic. The argument of desert from need could provide an answer to this problem by providing justification for the appropriateness of property rights as a reward for intellectual labour.

2.5 Conclusion

Finding a suitable labour theory of property is a precarious task. In this chapter I discussed the basis of most labour theories, namely John Locke's theory. Locke's original formulation leaves many questions to be asked, and much room for interpretation. It also gives life to many problems. Mainly, the argumentative connection between labour and property is missing. By itself, the theory is therefore insufficient.

Luckily, many interpretations have been developed. Two of them have stood out: the libertarian theories, and those based on desert. These two offer some perspectives to closing the gap between labour and property.

The difficulties mount when looking at intellectual property, however. The very different nature of abstract objects is prone to breaking some core mechanisms of existing theories, and demands a

105 Becker, 626–627.

different framework. The fact that abstract objects are non-rivalrous poses a difficulty for justification here: abstract objects, unlike physical objects, can be copied and handed over to others infinitely without the original being lost. An intellectual property right also lays claim on very different rights for the owner. Instead of only requiring that others do not take away an object, it is also stipulated that they do not copy it without permission.

Consequentialism can deal with these problems nearly effortlessly. Labour theories of property, on the other hand, cannot appeal to incentive, and require a significant effort to make the case that these IPR's are pre-political, and not resulting from some type of social contract. While libertarian theories do not seem to be capable of this, desert theories offer a solution. Desert can be situated on several different levels pertaining to property acquisition, and through this variety it is capable of answering to many challenges.

It is now time to add the element of artificial intelligence to this discussion. In the next chapter, I will discuss its relevant properties, and assess its impact on justifications of intellectual property.

Chapter 3 - AI, its product, and its stakeholders

3.1 Introduction

In the coming section I will address the delineation of the artificial system involved in the creative or inventive activity that concerns this text. Joanna Bryson offers a useful basic definition of artificial intelligence: “Intelligence is the ability to do the right thing at the right time given a dynamic environment (that is, a shifting landscape of ‘right times’ requiring more ‘right things’). AI is intelligence constructed deliberately as an artefact of a culture.”[106] This definition is not overly technical, and may not be sufficient in some very specific situations. For the purpose of this text, however, it will do.

The emphasis this definition puts on the relativity of intelligence is important: what is right or ‘intelligent’ in one situation is wrong in another. A computer that proves its intelligence by being able to defeat a human chess champion is entirely unintelligent when it comes to the task of designing tablecloth prints. This creates space for several different kinds of AI: on the one hand there is AI specially designed for one specific task, also called ‘weak’ or ‘narrow’ AI, and on the other hand there is also artificial general intelligence (hereinafter ‘AGI’), which is capable of

106 Dagmar Monett and Colin WP Lewis, ‘Getting Clarity by Defining Artificial Intelligence—a Survey’, in *3rd Conference on" Philosophy and Theory of Artificial Intelligence* (Springer, 2017), 212–214.

behaving intelligently in a general sense, by meeting or even surpassing human intelligence on most levels or every level. Most relevant in contemporary debate is weak AI. Although AGI is seen by some as a very realistic or nearly certain outcome of future development in the field of computing[107], it will remain exactly that (a future prospect) for at least years to come. Thus, the primary point of attention for this text is weak AI.

3.2 Relevant features of AI

Artificial intelligence is a notoriously difficult concept to define, resulting in plenty of disagreements.[108] The definition I offered above could certainly be criticised.[109] For this thesis, there is no need to fully define AI, as I will focus on the difficulty that AI presents for a theory of ownership based on labour and desert, namely creation without a direct human creator. Although the technology may still develop in several very different directions, and each one of those would yield distinct results, some further fundamental considerations regarding the features of AI will suffice. Bryson's definition offers a base that is wide enough to incorporate most examples of AI. In the next paragraphs I will explain some additional features of AI that are both realistic and relevant for a labour theory of property.

---

107 see for example Peter M. Asaro, 'What Should We Want from a Robot Ethic?', *The International Review of Information Ethics* 6 (2006): 9–16; Joanna J. Bryson, 'Robots Should Be Slaves', *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, 2010, 63–74; Kate Darling, 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects', in *Robot Law*, by Ryan Calo, A. Froomkin, and Ian Kerr (Edward Elgar Publishing, 2016), 213–232, doi:10.4337/9781783476732.00017; Ben Goertzel, 'Thoughts on AI Morality', *Dynamical Psychology: An International, Interdisciplinary Journal of Complex Mental Processes*, 2002; David J. Gunkel, 'The Other Question: Can and Should Robots Have Rights?', *Ethics and Information Technology* 20, no. 2 (June 2018): 87–99, doi:10.1007/s10676-017-9442-4.

108 Monett and Lewis, 'Getting Clarity by Defining Artificial Intelligence—a Survey'.

A first feature that is relevant for any theory of labour or desert is the autonomy of the AI. From a labour point of view one may at any point doubt the autonomy of the AI, therefore making the question of ownership of the product less problematic. If it is apparent that the user or creator bears a large enough role in the creation of the product, thereby passing beyond a certain threshold of dependency, there could be no more question about the ownership of the product.

It is conceivable that this autonomy could come in many different degrees. On one side of the spectrum there could be a program that takes all its instructions directly from its creator, while on the other side there could be a computer generating artworks without any instructions with an output that the programmer could not have predicted or made himself. For the sake of simplicity and relevancy, in this text I will consider creation by a fully autonomous and independent artificial intelligence.

Not only does this decision produces a more interesting discussion, but an independent AI is also becoming increasingly realistic. Arguably, in some forms, it is already in use. In news media, for example, the use of "automated journalism" is already widespread. This involves news articles generated by computers using complicated algorithms, thus often completely omitting every need for human intervention.[110] This is a particularly relevant example, as it shows the necessary dependency on existing works of intellectual property. News is never created in isolation. It is, almost by definition, derived from existing news, which is subject to property rights.

109 For example, one could ask what a "right thing" is, or how to define "artefact". Neither of these concepts are unproblematic. See Pei Wang, 'What Do You Mean by"AI"?', in *AGI*, vol. 171, 2008, 362–373. for more on this subject.

110 Javier Díaz-Noci, 'Artificial Intelligence Systems-Aided News and Copyright: Assessing Legal Implications for Journalism Practices', *Future Internet* 12, no. 5 (May 2020): 85, doi:10.3390/fi12050085.

Necessary dependency

Even though the AI that I will consider in this text will be on the autonomous side of the spectrum, there are some necessary limits to this independence. For one, the AI was made by someone. To a certain extent, at the basis of this autonomous system, there is always some human input. One could imagine, not quite unrealistically so, an artificial intelligence which is created by another artificial intelligence. But, even if that AI was created by another AI, and so on, at some point the chain of creation ends with a person, or several, who produced this artificial system. The necessary human input, however small it may be, would shape the results at the other end of the chain in a very significant manner. Therefore, the autonomy of any AI is limited by the fact that it is created – directly or indirectly – by a person.

For example, someone could design an AI that generates pictures independently. Another option would be to design an AI for generating text. The output of both systems is entirely different, and thus the initial creator of the systems has a large influence on the final product. Of course, in the case of an AGI, all these options would be available, and the choice could be made independently by the artificial system itself. This would limit the input of the initial creator, but it would still not do away with it completely. The choices made by the creator concerning mechanisms of learning, structure of thinking, processing power, sensory inputs, deep-learning input data, and so forth. would all have a large influence on whatever the system could (independently) ‘decide’ to produce.

A second limitation of any creative system is the input. Much like a human artist, an AI will need some inspiration, for lack of a better term. Most contemporary artificial systems need an input to generate a new one. Perhaps, some day, some AI’s will be able to produce creative or inventive works without any external input of existing works. Yet, in these cases, the limit will be situated elsewhere. Both in art and in invention, no work is created in full isolation from existing works.[111]

111 Carys J. Craig and Ian R. Kerr, ‘The Death of the AI Author’, *SSRN Electronic Journal*, 2019, 27, doi:10.2139/ssrn.3374951.

3.3 Three stakeholders

Given these considerations, it is possible to discern three parties that have a possible claim to property rights of the product of an autonomous AI: (i) the AI itself, (ii) the producer of the AI, and (iii) the owner of the data used in the creation by the autonomous system.[112] These three parties are the basic stakeholders in this situation, as seen from a labour-desert perspective. Each of these parties contributed, and the contribution of each party is necessary to achieve the final product. Yet, none of these stakeholders can make a full claim to the final product. The labour of one party is in no case enough to obtain the final product.

In current practice surrounding property, this situation of shared labour is common. The process of the publishing of a book could serve as an example here. Not only does the author put labour into the final product, but usually a whole array of people shares a part in the labour, such as publishers, editors, graphic designers, lay-outers, and others. The initial text itself could even be written by multiple authors in close co-operation. The way this is commonly treated is through a contract, deciding who gets which share, which party retains the intellectual property rights under which terms, and so forth. Through agreement, the initial property right can be completely modified. In the same was as physical property rights, intellectual property rights can be handed to someone else, traded, shared, or otherwise distributed. I will discuss these possibilities as one option, summarised as (iv) the contractual ownership of the output of AI.

3.4 Option 1: the AI as owner of its output

The most obvious choice, although arguably also the most controversial one, would be to attribute the property rights to the AI itself. After all, the creative process and the direct labour both originate from it. The focus on human authorship could be seen as an unwarranted assumption:

112 One could also identify the end-user as a fourth party here. But, in the case of an autonomous AI, this role is supposed to be unnecessary. Therefore I will not discuss it further in this text.

> "A core assumption in today's law is that it is uniquely human ingenuity and creativity that is an essential element for copyrights and inventions. Yet, today's Al systems often exhibit expression and independent creativity that we might otherwise attribute to humans."[113]

Following this view, while someone may have created the AI, the distance of that person's labour to the final product could be considered too great for there to be a relation of direct ownership. One could argue that this relation is instead comparable to that of parents and their children. While the children were 'created' by the parents, they are themselves capable of their own labour, and thus of creating their own property.

The problems with this option, however, are manifold. 'AI' in these cases most realistically will still refer not to a humanoid AGI, but to a very specialised system with a very narrow kind of intelligence. It certainly seems questionable to claim that this type of system, which presumably has no feelings, intentions, or use for any possible rights, would have to receive the property rights to its own product. So, while this option could make sense in the case of an AGI, the claim of a weak AI is much more insubstantial. Even in the case of a more advanced AGI, the basic problem remains: while the labour, in a way, was conducted autonomously by the machine, it seems wrong to ascribe it personhood or the rights that come with it.

There are two ways in which personhood can play a role here. First of all, the definition of labour might preclude the productive activity of an AI to be described as labour, as it was not executed by a person. Secondly, the possibility of receiving or deserving rights could be dependent on the recipient bearing the status of personhood or agency. I will begin by discussing whether or not AI could be considered a person, or in other words, if it could be given the rights and moral consideration that are enjoyed by people.

---

113 Russ Pearlman, 'Recognizing Artificial Intelligence (AI) as Authors and Investors under U.S. Intellectual Property Law', *Richmond Journal of Law & Technology* 24, no. 2 (2018 2017): 1–2.

3.4.1 AI, moral consideration, and rights

The problem of ascribing rights to intelligent machines could come down to the opposition between different views on the role of robots in society and technology. One could look at them from an instrumentalist point of view: robots can be categorised as just any other technology, which, as Gunkel has proposed, is often seen as "only a means to an end; it is not and does not have an end in its own right".[114] Continuing in this line of thinking, robot rights would be impossible. Even the "notion of robots having rights", in the eyes of David Levy, is "unthinkable".[115] While this instrumentalist view has a strong intuitive appeal, it does run into problems. It seems overly simplistic, condensing all forms of technology into the category of tools. This denies the existence and possibility of autonomous artificial systems altogether. On top of that, it cannot account for the special status of social robots either. "In other words", Gunkel argues, "practical experiences with socially interactive machines push against the explanatory capabilities of the instrumental theory, if not forcing a break with it altogether".[116]

Bryson, on the other hand, has argued against giving robots ethical consideration, without declaring robot rights totally impossible. In her aptly titled essay *"Robots should be slaves"*, she argues for a complete moral disregard for intelligent machines. She imagines it to be technically possible to produce AI within the requirements for agency or patiency, but denies the desirability or need to actually do so.[117] Robots, she argues, are property, and should only serve their owners, as slaves. Giving them rights, or producing artificial systems with a viable claim to rights, would cause "misassignations of responsibility or misappropriations of resources".[118]

---

114 Gunkel, 'The Other Question'.

115 David Levy, *Robots Unlimited : Life in a Virtual Age* (A K Peters/CRC Press, 2005), 393, doi:10.1201/b10697.

116 Gunkel, 'The Other Question'.

117 Bryson, 'Robots Should Be Slaves'.

118 Bryson, 65.

Such an approach would require us to refrain from producing certain kinds of robots, but also from behaving in a manner that would treat robots as more than mere tools. This, according to Gunkel, simply does not correspond to the empirical facts concerning human-machine relations.[119] There is plenty of evidence, he argues, to show that humans will very quickly form a strong bond even with unintelligent machines. Social robots, in particular, will often cause a very strong emotional response, similar to one that would be caused by a human being. Yet, as the argument of Darling proves, this does not mean that these machines deserve moral consideration.[120] Our strong feelings towards certain artificial entities only testify of our tendency to anthropomorphise things, and ascribe to them the ability to be conscious, feel or think, while they are not deserving of such expectations. Darling concludes that we should acknowledge our emotional responses to these robots, without giving them any further moral consideration.

There seems to be a consensus emerging about contemporary AI. In its current state, these machines are considered by most authors to lack the necessary requirements for moral consideration. Even in today's most advanced systems, AI's offer only a very narrow, specialised kind of intelligence. An AI that is only capable of playing chess, however proficient it may be, does not seem to have any real claim to moral agency. But, as Goertzel argues, this could change soon.[121] While today the question of AI rights might not be pressing, very soon, he believes, it will be. Asaro goes beyond this, stating that "At some point in the future, robots might simply demand their rights".[122] Perhaps because morally intelligent robots might achieve some form of moral self-recognition. This turning point will be met once AI technology has advanced to the point where some feature is present which will be sufficient for moral agency.

Following this "properties-approach", which is by far the most popular approach in the debate on the moral standing of non-human entities, the argument for or against moral agency is made from a

119 Gunkel, 'The Other Question'.

120 Darling, 'Extending Legal Protection to Social Robots'.

121 Goertzel, 'Thoughts on AI Morality'.

122 Asaro, 'What Should We Want from a Robot Ethic?', 12.

determination of the ontological condition of the entity at hand. In other words, whether something deserves moral consideration depends on what it is or which features it shows evidence of possessing.[123] This approach is by no means self-evident, but it is difficult to come up with another explanation. Even Coeckelbergh, who attempts to offer an alternative to the property-approach by arguing for a relational theory, must admit that it too can be reduced back to a discussion about properties: "… given the importance of appearance, one could still speak of properties as long as it is understood that we mean properties-as-they-appear-to-us within a social-relational, social-ecological context."[124] If being human is sufficient to be undeniably deserving of moral recognition, it must come down to some property that is part of this "being human" - whether it be relational, cognitive, or otherwise.

Giving rights to an entity has strong moral implications: it ascribes to this entity inherent moral value, which demands a fitting treatment. This is closely linked to deontological and natural rights traditions in moral theory, which themselves can be interpreted as a secular continuation of the religious idea that a person has an inherent worth as a creature of God.[125]

However, other theories of ethics, as for example consequentialism, have expanded moral consideration to non-human entities as well. Peter Singer defends animal rights on the grounds that suffering should be minimised in general, which holds for every sentient being.[126] From the deontological side, a similar argument has been made. According to Regan, some animals must be granted rights because their welfare affects them, as they are capable of having wants, beliefs, feelings and memories. These arguments are sometimes coupled to an emancipatory claim. In this

123 Gunkel, 'The Other Question'.

124 Mark Coeckelbergh, 'Robot Rights? Towards a Social-Relational Justification of Moral Consideration', *Ethics and Information Technology* 12, no. 3 (September 2010): 209–221, doi:10.1007/s10676-010-9235-5.

125 Coeckelbergh.

126 Peter Singer, 'Animal Liberation', in *Animal Rights: The Changing Debate*, ed. Robert Garner (London: Palgrave Macmillan UK, 1996), 7–18, doi:10.1007/978-1-349-25176-6_1.

view, moral consideration is seen as developing towards more inclusion: after having added slaves, women, gays, and animals, robots might be added as well.[127]

An important premise to both these justifications, consequentialist as well as deontological, lies in their use of ontological claims: for both, moral consideration is dependent on the ontological features of the entities at hand. These features can be sentience, intelligence, the capacity to have wants, feelings, and so forth. These features, or some of them, are taken to be sufficient for moral consideration. Because of this shared fundamental strategy, both justifications suffer from similar problems. Firstly, this puts the threshold very high. No existing AI would be considered sentient or conscious. One might compare this situation to environmental ethics: trees are neither sentient, nor conscious, but this does not seem to imply that they cannot receive moral consideration. This problem is echoed in the so called argument for marginal cases: not all humans share the necessary features required for moral consideration. Take for example young infants, people in a vegetative state, or simply anyone who is asleep. Following the adherence to ontological properties, one could argue that therefore these humans should not receive moral attention, or should not always receive it.[128] Such a conclusion would clearly raise some serious moral concerns.

In addition to these problems, the use of ontological properties is vulnerable to some more general epistemological problems: agreement on the primacy of ontological features relevant for moral consideration is difficult. According to Goertzel the deciding factor would be intelligence[129], while Singer claims that any being that is sentient deserves rights[130], and others link it to “primitive self-consciousness”[131]. It could even be impossible to achieve certainty over the possession of these features. This goes for robots as much as it does for humans themselves. The theoretical possibility

127 Coeckelbergh, ‘Robot Rights?’

128 Coeckelbergh.

129 Goertzel, ‘Thoughts on AI Morality’.

130 Singer, ‘Animal Liberation’.

131 Shaun Gallagher, ‘The Moral Significance of Primitive Self-Consciousness: A Response to Bermúdez’, *Ethics* 107, no. 1 (1996): 129–140.

of so-called ‘philosophical zombies’, or beings who look and act exactly like humans, but who lack consciousness or sentience, seems to eliminate any certainty that might be achieved in this matter.[132] From an external point of view, it is impossible to determine the presence of these ontological features.

To work around these difficulties, some authors have proposed to lower the thresholds for moral considerations, or to place them in a different conceptual framework. Coeckelbergh, for example, proposes to eliminate the requirement of certain knowledge about real ontological features, and replace it by the requirement of appearance, being satisfied when the possession of these features is merely apparent.[133] As a pragmatic solution, this works well, eliminating the epistemological problems that a requirement of certainty about ontological features generate. However, new problems quickly take their place. One might ask, to begin with, what would justify ascribing moral consideration to entities which only apparently seem to possess the relevant ontological features. Why would a philosophical zombie, incapable of feeling or thought, deserve rights? The impossibility of confirming actual presence of ontological features such as consciousness or sentience may not be enough to justify reducing the requirement to appearance. As Darling argues, our tendency to anthropomorphise non-human entities does not prove a moral duty to these entities.[134] Furthermore, the need to agree on which ontological features would be relevant still remains, and is no easier to fulfil with this solution. People would still need to decide between the appearance of sentience, the appearance of consciousness, and so forth. Additionally, agreement on which behaviour is sufficient to speak of ‘apparent’ consciousness or sentience seems nearly as unlikely as agreement on the conditions for certainty of these ontological features.

So, while the property approach seems to be the only approach with a possible answer to the question of whether AI could and should have rights, many problems must still be overcome before

132 Robert Kirk, ‘Decision, Control and Integration’, in *Zombies and Consciousness*, 2005.

133 Coeckelbergh, ‘Robot Rights?’

134 Darling, ‘Extending Legal Protection to Social Robots’.

any such answer could be found. For some of these problems, a solution seems impossible. The complete lack of access to knowledge of the relevant ontological properties of other entities is central here. Yet, despite these problems, there is one thing that can be agreed on: contemporary AI has not yet satisfied the necessary requirements for moral consideration or rights, but future AI probably will if we – their producers – eventually decide that it should. Even though this is exactly what Gunkel laments as "less a solution and more of a decision not to decide"[135], it seems that this pragmatic conclusion will be part of the cards that we have been dealt. For the purpose of this text, it will have to be sufficient. The question is: can it warrant the further conclusion that AI cannot own rights to its own product?

3.4.2 AI, labour and production

To answer the question of AI ownership over its product, the question itself must be critically examined. What does it mean for something to be a 'product'? Can we describe the relation of an AI to its output in terms of production, implying thereby that the AI is indeed the producer? And how should we understand the role of labour in this context?

Lockean productive labour

Once more, let us take Locke as a starting point. Similarly to his theory of property in general his use of the term labour is overall vaguely defined, and leaves much room for interpretation. One way in which it is often understood is in the sense of productive labour[136], meaning the kind of labour which produces something. This sort of labour can be found in many of Locke's examples of "labour mixing", such as husbandry, cooking, production of clothing and other essential goods.[137]

The basic concept of productive labour allows for a variety of activities to be included in its use. An AI producing something new autonomously through the use of algorithms and machine learning

135 Gunkel, 'The Other Question', 92.

136 Mossoff, 'Saving Locke from Marx', 290–291.

137 Locke, 'The Second Treatise', chap. V; Mossoff, 'Saving Locke from Marx', 295.

might fall under this concept too. In the end, something new is coming into being in a way that certainly seems to fall under the concept of production. Taking production as sufficient for labour does not clear things up much, however. A "neutral" use of the term, as proposed by Daniel Gervais who defines production simply as something that did not exist yet, coming into being[138], may be too broad. Would this include smoke rising from a fire, or leaves growing on a tree? While both could be described as production in a certain sense, for example as the production of smoke or greenery, they hardly seem sufficient for a description of labour.

In Locke's view, labour, as the kind of action that gives rise to property rights, is an action that is performed by a person, which for Locke means an intelligent free agent.[139] A child, the way Locke sees things, "has not Understanding of his own to direct his Will, he is not to have any Will of his own to follow".[140] I will come back to the requirement of a free will later. Moreover, the use of the term 'labour' by Locke and his contemporaries places labour opposite of recreation: any action that is not recreation is seen as labour. For them, "[l]abour is coterminous with non-recreational actions".[141] This distinction is based on the idea that labour consists of acting according to duty.[142] Applying this distinction to the actions of an AI seems hopeless: what would recreation even mean for an artificial system? Can an AI act according to duty? Does this in turn mean that an AI is always labouring, or that it is incapable of labour altogether?

Clearly, Locke's concepts are unprepared for this type of problem. Adapting them to a debate on AI is arduous to the point of pointlessness. And, in the end, they are based on the religious metaphysics underlying his arguments of divine law. Following this reasoning, an AI would never be qualified to own property in any case, as it was not created by God with the duty to preserve itself. This kind of

138 Daniel J. Gervais, 'The Machine As Author', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, March 2019), 4, https://papers.ssrn.com/abstract=3359524.

139 Tully, *A Discourse on Property*, 106–107.

140 Locke, 'The Second Treatise', § 58.

141 Tully, *A Discourse on Property*, 109.

142 Tully, 109.

religious metaphysics is best avoided in contemporary philosophical debate[143], and is likewise abandoned by labour-desert theories. It may help to look at other conceptions of productive labour.

Marxist productive labour

The concept of productive labour is often associated with Karl Marx. In his view, labour is seen "as 'formative' activity, as activity through which human beings give form to materials and thus objectify themselves in the world".[144] This type of concept seems very much compatible with a labour mixing argument, and with Locke's examples. It also adds an important element to the definition which is present in Locke, but rarely explicitly mentioned. This element is intentionality: "The simple elements of the labour processes are (1) purposeful activity, that is work itself, (2) the object on which that work is performed, and (3) the instruments of that work."[145] That feature of labour which Locke described as acting according to duty is translated here as purposefulness: "Marx's category of labor is understood in terms of a goal directed social activity that mediates between humans and nature, creating specific products in order to satisfy determinate human needs. Labor, when understood in this way, is considered to lie at the heart of all social life."[146] In other words, intentionality plays a large role here.

Marx goes much further than just intentionality, however. For classical Marxists, labour is understood as the activity of production and creation, both of the product and of the self as a fully human being. Labour is seen as the most human of activities:

> "In laboring, man not only produces objects to satisfy his needs, thereby changing and subjugating nature, but also makes himself into what he is; through laboring, through

143 See chapter 2

144 Sean Sayers, 'The Concept of Labor: Marx and His Critics', *Science & Society* 71, no. 4 (2007): 432.

145 Karl Marx, *Capital: A Critique of Political Economy - Volume 1* (London: Penguin, 1883), 284.

146 Moishe Postone, *Time, Labor, and Social Domination: A Reinterpretation of Marx's Critical Theory* (Cambridge University Press, 1995), 7–8.

> developing his productive powers, man comes to educate, civilize, and indeed, create himself as a fully 'human' being."[147]

Marx's account, in its turn, is based on G.W.F. Hegel's concept of labour, for whom labour is a uniquely human and formative activity.[148] This focus on self-development and labour as a social activity takes any relatively primitive AI out of the picture, or even AI in general. In classical Marxism, labour is seen as a human activity; so much so that it is often considered the defining trait of humans as a species.[149] Unless the definition of a human were thoroughly adapted, there seems to be no room for AI in this discussion.

Intentionality, free will, and consciousness

The Marxist-Hegelian view on labour of course carries a lot of metaphysical luggage from Hegel's and Marx's vast and complicated philosophies. One element that some authors propose to take away from it is intentionality.[150] It can serve as a condition for labour. Often, it is described as "aboutness" or "directedness".[151] Claiming that an AI is directed in an action towards a certain object or goal does not seem far fetched at all. When an AI is generating text, it might be directed towards writing a news article, for example. This applies the terms intentionality and directedness in a very loose and everyday sense, but perhaps that is all that can be expected from such a definition.

While some authors consider intentionality to be a necessary component for any AI to ever receive rights[152], it is unclear if this kind of intentionality is to be regarded as the same kind that a Marxist

147 R. N. Berki, 'On the Nature and Origins of Marx's Concept of Labor', *Political Theory* 7, no. 1 (1979): 36.

148 Sayers, 'The Concept of Labor', 433.

149 Sayers, 434.

150 see Cwik, 'Labor as the Basis for Intellectual Property Rights'.

151 Angela Mendelovici, *The Phenomenal Basis of Intentionality* (New York, USA: Oxford University Press, 2018), 4.

152 see John P. Sullins, 'When Is a Robot a Moral Agent', *International Review of Information Ethics* 6, no. 12 (2006): 28.

conception of labour requires. Making moral agency a requirement for intentionality, and thus labour, is appealing, but may set the bar too high. Does an entity really need to be a full moral agent to be able to perform intentional acts? As several philosophers have argued that young children and certain animals are not full moral agents because they lack self-consciousness[153], would this mean that these children and animals are incapable of intentional acts, and thereby are ineligible for labour? Locke would certainly think so[154], but it does remain a harsh conclusion.

In a similar vein, intentionality could also be linked to free will, as proposed by Pearl and Mackenzie[155], or to consciousness as proposed by Mendelovici.[156] This would mean that, in order to be capable of performing intentional acts, an entity or agent would first have to possess free will or consciousness. With this requirement, Locke's condition of an "intelligent free agent" could also be met. However, the problem of free will and consciousness has been notoriously insurmountable, with John Searle noting that: "(t)he persistence of the traditional free will problem in philosophy seems to me something of a scandal. After all these centuries of writing about free will, it does not seem to me that we have made very much progress."[157] A test for consciousness or free will does not exist in any practical or theoretical way, let alone an agreed-upon definition. In fact, if we are to doubt the consciousness of a machine exhibiting behaviour that could be seen as the result of consciousness, we are left with no reason to trust the possession of consciousness by other human beings and should therefore doubt it just the same.[158]

This seems to be where property theory, law, and practice follow a double standard. For humans, the appearance of possessing free will and consciousness is enough to make a valid property claim;

153 see Gallagher, 'The Moral Significance of Primitive Self-Consciousness'.

154 See Locke's remarks on children lacking an understanding of their will above

155 Judea Pearl and Dana Mackenzie, *The Book of Why: The New Science of Cause and Effect*, 1st ed. (USA: Basic Books, Inc., 2018), 326–328.

156 Mendelovici, *The Phenomenal Basis of Intentionality*, 4.

157 John Searle, *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power* (Columbia University Press, 2006), 37.

158 Kirk, 'Decision, Control and Integration'.

nobody will require any further proof of eligibility. For a machine, on the other hand, no proof is ever taken to be sufficient. In fact, if the free will requirement were to be consistently maintained, no one would be eligible, neither machines nor humans, as no one would be able to provide proof of their possession of free will. Moreover, if we take a strong determinist position[159], free will would not exist at all, and therefore (intellectual) property would become an impossibility.

From a neurological point of view, free will and intentionality could be altogether seen as an illusion: "Granted that free will is (or may be) an illusion, why is it so important to us as humans to have this illusion? […] Why did evolution labor to endow us with this conception?"[160], Pearl and Mackenzie ask, concluding that "[i]f neural signals from the brain trigger all our actions, then our brains must be fairly busy decorating some actions with the title 'willed' or 'intentional' and others with 'unintentional.'"[161] If free will and intentionality are illusions, the results of millions of years of irrational and chaotic evolution favouring a contingent criterium of survival and reproduction, how can they be used to justify appropriation and identify an author?

3.4.3 The remaining case for the AI as owner

It is clear: the case for AI ownership over its own product is difficult to make, and goes against nearly all conventional conceptions of moral agency, labour, and intentionality. Yet, despite the seemingly insurmountable common intuitive resistance, the possibility seems to be extant. The very existence of advanced, autonomous AI questions our use of these fundamental terms, and by redefining them or putting them aside altogether, ascribing labour-desert to an AI over its own product may be tenable.

On the other hand, a decision such as this would require several serious concessions, and would have numerous consequences that are difficult to oversee. After all, the AI we have defined earlier is

159 see for example Searle, *Freedom and Neurobiology*, 44–45.

160 Pearl and Mackenzie, *The Book of Why*, 18.

161 Pearl and Mackenzie, 18.

not of the conscious and feeling kind. Attributing it authorship and therefore ownership would therefore have the effect that these ownership rights are in the hands of an entity which cannot do anything with them. Furthermore, authorship is a deeply social notion, involving countless social interactions between humans over generations. As Craig and Kerr argue:

> "GAN-generated outputs[162] render imperceptible all of these dialogic processes undertaken by prior generations of humans participating over time in the social practice of authorship. Consequently, when we substitute an AI for a human, we are permitting the AI to stand in for significant human expressive activity and relations of communication that occur, invisibly, behind the scenes. Anthropomorphic framings of the work done by a GAN that speak of "deep learning," "emergent creativity," "generative works," "algorithmic authorship" and the like may offer some utility; but such rhetorical flourishes also reinforce the illusion that machines possess a kind of intelligence and sociality that they do not—and cannot—in fact have."[163]

Ascribing full authorship to an AI seems therefore unwarranted. The AI does indeed seem to adopt most the functions regularly ascribed to an author, and fulfils the work that is required from the author[164] – to such an extent that, if the process performed by the AI had been instead performed by a human, this human would probably be deserving of authorship. Yet, because of the social-historical nature of the concept of authorship, and the fact that an AI cannot actually do anything with the rights ascribed to authors, accrediting full authorship to an AI is simply impossible.

---

162 'GAN' is short for 'general adversarial network', a kind of AI method of producing AI algorithms which involves two systems (usually in the form of neural networks) competing to produce the best result. See Josef Drexl et al., 'Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 2019), doi:10.2139/ssrn.3465577.

163 Craig and Kerr, 'The Death of the AI Author', 27–28.

164 I will discuss authorship more exhaustively in 3.5 and 3.6, assessing also more precisely which functions an AI could be capable of taking up.

It seems that appointing the AI as author would require a redefinition of some of our fundamental concepts linked to authorship. Additionally, the other relevant parties, namely the creator, the data-owner, the user and the investor would be required to give up their claim of ownership against a machine that has no real interest in property rights for as long as it cannot be said to feel pleasure or suffering. The question of ownership and authorship goes much further still, and the other parties involved have a strong claim to both from a labour perspective too, as we shall see below.

3.5 Option 2: the creator of the AI as owner of its output

If we take the claim to property rights by the AI itself to be insufficient, as many authors do, the next best option for an owner is the creator (‘developer’, ‘producer’, ‘programmer’ or ‘manufacturer’) of the AI. The producer is the one who laboured on the AI to make it capable of generating an output in the first place. If the output cannot be owned by the AI, it must be the creator of the AI who deserves it. This would solve the problems of a lack of a human creator in legal requirements embedded in the concept of authorship[165] and could meet the possible need for a human creator that a concept of productive labour implies (see the previous section).

Yet, appointing the creator of the AI as the owner of its product is an option that is not without problems from a labour-desert point of view. The final product of the AI seems to be a very indirect result of the labour put into the creation of the AI. For some authors, it is particularly the autonomy of the AI which makes its creator an unfit candidate for ownership of the IPR’s over the final product: “the process of creation does not require the management or even the involvement of

165 see Gonenc Gurkaynak et al., ‘Questions of Intellectual Property in the Artificial Intelligence Realm’, SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 2017), https://papers.ssrn.com/abstract=3295747.

human intelligence. Thus, the developer of AI may not be deemed owner of the work created by the AI on its own."[166]

This reasoning is echoed in the formulation of IP law, for example in the requirements of an inventive step. While the producer of the creative AI may have created a work (both in the case of invention and in copyright) that is commonly protected by IPR's, the human creative or inventive process that has normally led to the final product is in this case missing.[167] While the creator of the AI possesses without a doubt the technical knowledge and skills to produce the artificial system, they do not necessarily need to possess the creative or inventive skills or knowledge that would normally be necessary for the work produced by the AI. This is ultimately the point of the AI: it can process huge amounts of data in very little time, and eventually produce something similar to the input data itself. For Gervais, it is difficult to still attribute ownership to the creator of the AI in these circumstances: "In such a scenario, it is but fiction to see a human author as being responsible for—or the owner of rights in—the creation, because the AI machine uses its own insights to create."[168]

So, a choice presents itself: does the process leading up to the final IPR-protected work matter, or is the only relevant measure the mere existence of a work belonging to some category of works that is normally protected by IP law? In the latter case, certain justifications of intellectual property rights would become incompatible with the conclusion. An argument pertaining to desert from excellence that is based on the excelling skills or knowledge of the creator could not justify this step, as these characteristics would simply be missing. An argument based on freedom such as the one presented by Cwik[169] also seems to be problematic, as it is no longer the freedom to choose a life of artistic

166 Gurkaynak et al.

167 Amir. H. Khoury, 'Intellectual Property Rights for Hubots: On the Legal Implications of Human-like Robots as Innovators and Creators', *Cardozo Arts & Entertainment Law Journal* 35, no. 3 (2017 2016): 651–652.

168 Gervais, 'The Machine As Author', 7.

169 see Cwik, 'Labor as the Basis for Intellectual Property Rights'. Cwik argues for intellectual property rights based on labour from a principle of liberty. He contends that intellectual property is justified because it allows people to choose how they themselves labour: "An ideal set of IP institutions is, thus, at least in this one big way, conducive

creation or invention that is protected by IP rights, but rather the choice to create an AI. The same goes for a justification based on desert from need, as IP laws would no longer serve to protect the special needs of the people who identify as inventors or artists.

So, the options come down to a choice between process-based justifications (such as desert from excellence, need, and freedom), and value-based justifications (or desert-for-value). In labour terms, this distinction could be rephrased as the decision whether the labour is justifying property rights because of some special characteristic of the labour, or because of the value this labour creates. While physical and intellectual property rights are generally distinguished by referring to their object, namely material and abstract objects, it is necessary here to go further and look at the act of labour that creates the object. In the first option of process-based justifications, creativity is by many considered to be the distinguishing element that puts this kind of labour apart from other labour.[170]

3.5.1 The creative process

To begin, let me consider the creative process that is at work during creation by an artificial machine.[171] Discussing creativity is challenging, and a consensus about how to define it is not yet on

to the goal of creating the economic conditions in which individuals have control over how, and under what circumstances, they labor, and can use their labor as a means to meet their needs and desires, and pursue the ends they find most valuable." (p. 694) In reality, this reasoning seems to be dependent on several kinds of labour-justifications, and could be rephrased as a purely desert-based theory. Arguing for labour as a means of meeting special needs specific to work with abstract objects hinges on desert from need.

170 see for example Jani Ihalainen, 'Computer Creativity: Artificial Intelligence and Copyright', *Journal of Intellectual Property Law & Practice*, 2018.

171 By this "creative process" I mean to indicate both creation of artistic works fit for copyright as well as inventive works that are to be protected by patents. This use of the term not limiting creativity just to artistic works is inspired by, among others, Pat Langley and Randolph Jones, 'A Computational Model of Scientific Insight', December 1986, https://escholarship.org/uc/item/54x8v354., who discuss scientific insight as a form of creativity. While

the horizon.[172] Discussing creativity related to AI is arguably even more difficult, as the previously discussed problems of consciousness, free will, and the impossibility of judging mental states from the outside[173] create underlying uncertainty in this discussion.[174]

Unpredictability

A first and common way to denote this autonomous creativity in machines is to look at the predictability of their output.[175] The idea behind this is that creativity is opposed to strict, rule-based, and thus predictable thinking. So, the idea is that, if the output of an AI cannot be predicted by its creator, the AI possesses real creativity. In law, this is echoed by the requirements of non-obviousness and novelty. Ultimately, this view can be reduced to a conception of the machine as a mere tool.[176] If the product is predictable, the AI is a tool like any other, and the person using it is the one performing the labour, and therefore the one who deserves property rights over the product.

A simple way to avoid a predictable product is by introducing randomness. If the output of a machine is random, it becomes by definition impossible to predict. According to some, randomness is necessary to simulate creativity in artificial intelligence.[177] For an AI, adding randomness could be a simple matter of editing some code. Allowing such a trivial alteration to make an entity eligible to be ascribed real "human" creativity seems wrong. And while deep learning allows for machines

---

scientific insight is perhaps a broader concept than inventiveness, the inventiveness demanded by copyright does imply the possession of knowledge and skills that are not commonplace amongst non-experts. Inventiveness in such a technical manner does seem to amount to creativity.

172 Annemarie Bridy, 'Coding Creativity: Copyright and the Artificially Intelligent Author', *Stanford Technology Law Review* 2012 (2012): 10.

173 see Kirk, 'Decision, Control and Integration'.

174 See my discussion of AI as the potential owner of the IPR, earlier in this chapter, for an overview of the arguments concerning consciousness and free will.

175 Bridy, 'Coding Creativity', 10.

176 Gervais, 'The Machine As Author', 19.

177 David Cope, *Computer Models of Musical Creativity* (MIT Press Cambridge, 2005), 12.

to make unpredictable choices without refraining to randomness[178], this still makes unpredictability an unreliable criterium for creativity.

More than unpredictability

What seems to be expected from an entity for it to be called creative is more than unpredictability. Concerning the evaluation of internal processes in a machine, John Searle offers a thought experiment that plays to the sceptical side of the debate. It is widely known as the 'Chinese Room Argument'.[179] It imagines someone sitting in a sealed room, accompanied by the documentation of rules necessary to translate English text to Chinese, either in the form of books or a computer program giving instructions. When a piece of paper containing a string of English words is slipped under the door, Searle imagines the person could translate it to the appropriate Chinese word-combination following these set rules. While the person could have no understanding of Chinese, it may seem as if there is indeed a Chinese speaker present in the room.

This thought experiment has mainly been used as an argument against the Turing test, by showing that the outward appearance of language comprehension is no proof for any kind of real intelligence, but instead can be the result of following a simple set of rules with no understanding of the language.[180] Applying this argument to creative behaviour of an AI is simple. While an artificial system's output may seem similar or could even be indiscernible from human creative output, Searle's argument points out that this does not signify that the AI possesses true creativity. This would mean that an AI generating works that are generally supposed to fall under IP protection may have made no real creative contribution.

---

178 Gervais, 'The Machine As Author', 19.

179 John Searle, 'Minds, Brains, and Programs', *Behavioral and Brain Sciences* 3, no. 3 (1980): 417, doi:10.1017/s0140525x00005756.

180 David Cole, 'The Chinese Room Argument', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2020 (Metaphysics Research Lab, Stanford University, 2020), https://plato.stanford.edu/archives/win2020/entries/chinese-room/.

The Chinese Room Argument seems to suggest that there is more to mimicking human behaviour than simply generating the same output. What Searle demands from his Chinese speaker is understanding. If the person translating the text does not understand what is happening, and is simply following rules to generate an outcome, she is no real Chinese speaker. Similarly, one could argue that an AI generating an output that is supposed to be creative or mimic creative works without understanding what it is doing, is not exhibiting creative behaviour at all, but is instead merely imitating it – regardless of whether or not the output is different from that of a human author.

A requirement for understanding could be interpreted as a demand for consciousness. If consciousness were to be taken as a constituting element of creativity, this would entail that a creative AI is unlikely to ever exist, and according to some even impossible: "If creativity is defined in terms of human consciousness [...] then machines *ex vi termini* will never be able to achieve it, no matter how sophisticated they become."[181]

However, if creativity is based on the potentiality of consciousness, all the problems concerning the impossibility of any kind of test of its presence, and the difficulties surrounding its definition return.[182] Consciousness is virtually impossible to define, and difficult to manage as a practical criterium due to the lack of a test. The debate on whether or not consciousness will ever be ascribed to an AI is also not settled.[183] Maybe it is safe to assume that contemporary AI and weak or narrow AI of the kind that is to be developed in the coming years are not yet ready to be considered conscious, although this again would depend on the definition of conscience.

It is perhaps easier to rephrase the need for more than just randomness as a general requirement of resemblance to the process of human creativity in a broader sense. In other words, an entity could

181 Bridy, 'Coding Creativity', 9.

182 See note 159

183 Steve Torrance, 'Ethics and Consciousness in Artificial Agents', *AI & Society* 22, no. 4 (2008): 498.

be ascribed creativity if the method by which it generates its output is similar to the way human creativity is practised and expressed. An answer to this criterium depends on two things: the nature of human creativity, and the mechanisms of artificial reasoning. If the nature of human creativity is structured in a such a manner that it cannot be sufficiently formalised in order for it to be understood, extracted and adapted to programming in any way, then a creative AI is impossible to achieve. The same goes for the mechanisms of artificial reasoning: if they are too rigid or unfit in any other way to adapt to the necessities of creativity, the possibility of AI creativity is equally ill-fated.

Scepticism concerning the possibility of translating human creativity to a formalized model is not uncommon.[184] Liane Gabora is doubtful about the feasibility of such a model, due to the nature of creative ideas. In her opinion, these ideas cannot be reduced to a mere combination of other, constituent ideas:

> "Is it possible to mathematically model the creative process? One big stumbling block is that a creative idea often possesses features which are said to be emergent: not true of the constituent ideas of which it was composed. For example, the concept snowman has as a feature or property 'carrot nose', though neither snow nor man does."[185]

For Erden, this difficulty is sufficient to conclude that human-like creativity altogether cannot be attained by any AI. In Wittgenstein's terminology, he concludes:

184 see for example Y. J. Erden, 'Could a Created Being Ever Be Creative? Some Philosophical Remarks on Creativity and AI Development', *Minds and Machines* 20, no. 3 (August 2010): 349–362, doi:10.1007/s11023-010-9202-2; and Peter Kassan, 'Al Gone Awry: The Futile Quest for Artificial Intelligence', *Skeptic (Altadena, CA)* 12, no. 2 (June 2005): 12.

185 Liane Gabora, 'Cognitive Mechanisms Underlying the Creative Process', in *Proceedings of the 4th Conference on Creativity & Cognition*, C&C '02 (New York, NY, USA: Association for Computing Machinery, 2002), 130, doi:10.1145/581710.581730.

> "When we call someone or something creative we use a term which holds for those language-users who share our games. [...] even if a machine could mimic most of our creative abilities, these abilities would not be creative in the way in which this word applies to human beings. This is because the machine […] does not share our form of life, and therefore cannot share our language-games."[186]

So, for Erden, an AI cannot attain what we express when we talk about creativity in a human context. However, he does leave open the possibility of artificial creativity in other meanings of the word, suggesting that "focus can instead turn to consideration of what it would mean for an AI program to be creative in AI terms".[187] This offers some perspective by introducing the possibility of multiple kinds of creativity.[188]

However, whereas in Erden's conclusion the kind of creativity that is accessible for AI is placed outside of any original, human type of creativity, others do not always share this sentiment. Bridy, for example, identifies a kind of creativity in human art that could be very well suited for an artificial adaptation.[189] It is a kind of rule-bound creativity, exemplified by the French literary movement of Oulipo.[190] The corresponding "Oulipian method of writing requires writers to compose under self-imposed external constraints, often based on mathematical equations".[191] If a rule-based method of writing can be considered as a kind of human creativity, and its output

186 Erden, 'Could a Created Being Ever Be Creative?', 361.

187 Erden, 361.

188 Erden's use of Wittgenstein's maxim "meaning is use" may be taking the principle out of context, and therefore be unwarranted. By analysing the definition of creativity to evaluate if the activity by AI fits, he is committing to a definition-based approach, which a consistent application of the principle of "meaning is use" may not allow. This is open to interpretation, however, and does not have any impact on the relevant argument for this text, so can be safely disregarded here.

189 Bridy, 'Coding Creativity', 11.

190 Oulipo is an acronym for "Ouvroir de Litterature Potentielle", or "Workshop for Potential Literature". It was embodied by a writing group founded in 1960 by Raymond Queneau and Francois Le Lionnais. See Loss Pequeno Glazier, *Digital Poetics: The Making of E-Poetries* (University of Alabama Press, 2001), 120.

191 Bridy, 'Coding Creativity', 11.

susceptible to copyright, then the case for AI creativity is getting stronger, Bridy argues. Essentially, the existence of this kind of creativity shows how flexible the common conception of human creativity actually is.

In fact, Bridy's argument is even more far-reaching. She maintains not only that some form of formal or rule-based human creativity exists, but also that it encompasses a lot of the artistic work that was already being created at any rate.[192] She points towards W.A. Mozart's *Musikalisches Würfelspiel*, a method of manually "generating" new musical works based on a combination of both the general rules of composition and pure chance. In fact, many of the mediums to which creative works are necessarily limited are already by definition rule-based: "Oulipo's embrace of rules and constraints [...] can productively be understood as a means of making a virtue of necessity; it isn't as if writers (or any other kind of artist, for that matter) can ever really break free of rules - of grammar, of syntax, of diction, etc."[193]

According to Margaret Boden, the common notion that creativity is the opposite of rules and constraints is on the whole unwarranted: "[p]eople often think that talk of 'rules' and 'constraints' - especially in the context of computer programs - must be irrelevant to creativity, which is an expression of human freedom. But far from being the antithesis of creativity, constraints on thinking are what make it possible."[194] Indeed, instead of limiting creativity, rules are vital to its manifestation.

3.5.2 Kinds of creativity

So, much of our human creativity is already limited by or even produced by rules and constraints, which makes the affected kinds of creativity much more suitable for an artificial refitting. But

192 Bridy, 12–13.

193 Bridy, 12.

194 Margaret A. Boden, *The Creative Mind : Myths and Mechanisms*, 2nd ed. (Routledge, 2004), 95, doi:10.4324/9780203508527.

which kinds do we mean by this? A sceptic of artificial creativity could still insist that these kinds of creativity, although associated with the concept, are not of the legitimate kind that authorship requires.

P-creativity and H-creativity

Boden identifies two important kinds of creativity.[195] The first, which she dubs 'psychological creativity' ("P-creativity" for short), "involves coming up with a surprising, valuable idea that's new to the person who comes up with it. It doesn't matter how many people have had that idea before." The second kind of creativity she identifies is 'historical creativity' (H-creativity), which refers to coming up with an idea that is historically new, and has never arisen before in human history.[196] According to Boden, "H-creativity is the more glamorous notion, and is what people usually have in mind when they speak of 'real' creativity".[197] This is good news for AI, as it has already proven its capacity of generating H-creative output in coming up with designs of technical components in a fully autonomous manner.[198]

P-creativity is another matter. While the capacity to produce H-creativity seems to imply a necessary ability to also generate P-creativity, this will depend on its interpretation. Opposing the notion of P-creativity as merely individual novelty, "one might prefer to define a P-creative idea as a fundamental novelty (with respect to the person's previous ideas) *whose significance is*

195 Boden, 43–44.

196 Bridy proposes that these two kinds accompany existing categories of IP-law. "H-creativity aligns with the standard of novelty in patent law", while P-creativity "aligns with the originality standard in copyright law and with copyright law's requirement of independent creation as opposed to absolute novelty" see Bridy, 'Coding Creativity', 13.

197 Boden, *The Creative Mind*, 43.

198 See for example John Koza's fully computer-generated design for general-purpose PID and non-PID controllers in Robert Plotkin, *The Genie in the Machine: How Computer-Automated Inventing Is Revolutionizing Law and Business* (USA: Stanford University Press, 2009), 1–3; and Bridy, 'Coding Creativity', 14.

recognized by the person concerned".[199] This is the kind of creativity that can be linked to the aforementioned requirement of consciousness proposed by other authors. The difference here is that this is not seen as an absolute and only kind of creativity, but merely as one of several kinds. All of the kinds of creativity described so far can also be identified in human behaviour, as this is what they are derived from, and the same kinds are sufficient for humans to claim authorship over a work. When one of these kinds of creativity is outsourced to an AI, the property claim of the creator over the output is consequently severely threatened.

Another level of creativity

Aside from the level of P- and H-creativity, which revolves around individuality, Boden describes a second level, at which more types of creativity can be identified.[200] This level presents itself in terms of the nature of the ideas that are subject to creativity, and the different ways in which these ideas are processed and assimilated.

Boden establishes three different kinds of creativity at this level. The first is combinational creativity, which involves new combinations of existing ideas that lie within the same inherent conceptual structure. This is at play in the creation of things such as poetic imagery and analogy. The second type she calls exploratory creativity, which "involves the generation of novel ideas by the exploration of structured conceptual spaces"[201], which produces ideas that are not only novel, but are also unexpected. It could be described as "taking what is already there and exploring its outer edges, extending the limits of what is possible while remaining bound by the rules".[202] Finally, there is transformational creativity. As the name suggests, this kind deals with "the transformation

199 Boden, *The Creative Mind*, 44.

200 Margaret A. Boden, 'Creativity and Artificial Intelligence', *Artificial Intelligence*, Artificial Intelligence 40 years later, 103, no. 1 (August 1998): 348–349, doi:10.1016/S0004-3702(98)00055-1.

201 Boden, 348.

202 Marcus du Sautoy, *The Creativity Code: Art and Innovation in the Age of AI*, *The Creativity Code* (Harvard University Press, 2019), 8, https://www.degruyter.com/hup/view/title/565282.

of some (one or more) dimension of the space, so that new structures can be generated which could not have arisen before".[203] These last two, Boden notes, are closely connected, but although distinguishing them can be difficult when the conceptual space is ill-defined, they are in fact separate.

One could point out, as Boden does, a hierarchical order present in this distinction. Creativity of the first kind seems to be the simplest, followed by the second and the third respectively. The second kind seems to be the most prevalent one in human creative works, but the third, rare kind has a great impact. The more fundamental the transformation is, the more it surprises and amazes, as long as the connection to the old conceptual space is maintained.[204] However, this hierarchy does not translate itself into a parallel hierarchy concerning the ease of emulation by an AI. In fact, the second kind "is also the sort of creativity at which computers excel. Pushing a pattern or set of rules to an extreme is a perfect exercise for a computational mechanism that can perform many more calculations than the human brain can".[205]

At first sight, the third kind of creativity seems impossible for an AI to attain. Its disposition to think outside of the existing conceptual space seems like a kind of resourcefulness or originality that a rigid rule-based machine cannot hope to achieve. This is the argument that original sceptics of artificial creativity such as Ada Lovelace would posit.[206] Nonetheless, much has changed since the 19th century, not least of all machine intelligence. In the recent decades, artificial intelligence has moved from a rule-based top-down approach to a bottom-up approach, meaning that the output of an AI is no longer limited by the code of the creator.[207] Through this new approach, involving deep

203 Boden, 'Creativity and Artificial Intelligence', 348.

204 Boden, 348–349.

205 du Sautoy, *The Creativity Code*, 8.

206 A. Lovelace, 'Faster Than Thought', in *Notes on Menabrea's Sketch of the Analytical Engine Invented by Charles Babbage*, ed. B.V. Bowden (London: Pitman, 1953), 398.

207 Jean-Marc Deltorn, 'Deep Creations: Intellectual Property and the Automata', *Frontiers in Digital Humanities* 4 (2017): 3–4, doi:10.3389/fdigh.2017.00003.

learning and neural networks, AI has been capable of achieving transformational creativity.[208] Examples of this are EURISKO and AM, programs used to "synthesize new and different programs, or which modify and improve themselves."[209]

3.5.3 Concluding remarks on the creator's claim

What is left of the original scepticism concerning artificial intelligence and creativity that was contained in the adapted version of the Chinese Room Argument? It has become clear that the argument faces some problems. First, it misinterprets human creativity. Trying to subsume all of human creativity under one definition, as we have seen, is a project that is bound to fail. A single unitary definition cannot grasp the plurality of ways in which creativity can be practised and expressed.[210] 'Real' human creativity can be rule-bound too, which makes the prospect of a machine adopting it very real. But even outside of this rule-bound kind of creativity, it is apparent that AI can already demonstrate creativity of a fundamental kind, not just following rules, but also changing them, and transforming the original conceptual structures. The fact that AI's have often become less transparent in recent years[211] also means that creativity does not have to be fully understood to be replicated, and can hold on to its mysterious image.

208 Boden, 'Creativity and Artificial Intelligence', 351.

209 Douglas B. Lenat, '9 - The Role of Heuristics in Learning By Discovery: Three Case Studies', in *Machine Learning*, ed. Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell (San Francisco (CA): Morgan Kaufmann, 1983), 286, doi:10.1016/B978-0-08-051054-5.50013-3.

210 In fact, this misinterpretation can already be identified in the original argument. It is very difficult to define Chinese as a single language, as only 2/3 of the Chinese-speaking population speaks Mandarin, or "common Chinese", within which there are already several varieties. The other part of the population "are divided into several groups, such as the Cantonese, Hakka, and Min, with forms of speech so distinctive that they are mutually unintelligible." John DeFrancis, *The Chinese Language: Fact and Fantasy* (University of Hawaii Press, 1986), 39.

211 Andrés Guadamuz, 'Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, June

This leads us to a second problem with the adapted Chinese Room Argument. The claim of AI creativity goes further than simply noticing a similarity of output. Contemporary ways of creating AI programs, including deep learning and neural networks, point towards strong analogies to human creativity in the process that precedes the output. In fact, AI's that are created through artificial neural networks are modelled after the structure of the human brain.[212] Implicit in this enterprise is a common mechanical view of the human brain, in which the brain is ascribed a strong semblance to a machine that can be understood and reproduced.[213]

Counter-arguments from the sceptic

Sceptics are left with two options here. First, they can assert that of the different kinds of creativity, only a part pertains to what can be called "real" creativity, which AI cannot replicate. Since contemporary machines have not yet been able to simulate all kinds, this would be an effective approach. However, it would also mean that whenever humans create something utilizing these kinds of creativity, they should be prevented from claiming authorship over the work. In fact, some of the greatest artists in human cultural history, such as Bach and Mozart, created works which fall into categories of creativity that are very well susceptible to artificial reproduction[214]. Should we strip them from their statute of creative genius and author? It cannot be denied that they showed great creativity and were the authors of their works. From a labour-desert theory, too, it ought to be admitted that they must be given property rights over their works.

2017), 4–5, https://papers.ssrn.com/abstract=2981304.

212 Drexl et al., 'Technical Aspects of Artificial Intelligence', 5.

213 Pearl and Mackenzie, *The Book of Why*, 14–15.

214 In fact, Bach's musical style has already often been the subject of AI reproduction, and the results have been reliably accurate. See for example Cheng-Zhi Anna Huang et al., 'The Bach Doodle: Approachable Music Composition with Machine Learning at Scale', *ArXiv Preprint ArXiv:1907.06637*, 2019; and Feynman T. Liang et al., 'Automatic Stylistic Composition of Bach Chorales with Deep LSTM.', in *ISMIR*, 2017, 449–456.

The second possible argument of the sceptic is similar, but more fundamental in its doubt about AI: the sceptic could always claim that consciousness is necessary. Awareness of the value of its work and of the process itself is what an AI will likely never attain. But this would create a double standard. From human creators and inventors, this requirement is not usually met. Take Kepler, for example, who, upon creating a model of elliptical planetary orbits, decided that it was a worthless idea, calling it "a cartload of dung".[215] Imagine someone else walking in, realizing that there is some potential in the model, and claiming it as their own invention. Most people, Kepler included, would have some trouble with this claim. Especially from the perspective of a labour theory of property, it would be a grave injustice. While the recognition of a work's potential value is of great importance, recognizing an idea as valuable often does not at all require a similar amount of work to developing the idea in the first place. Whoever would have taken Kepler's model as his own would be cancelling years of careful observation and record-keeping that were necessary to produce the model.

The sceptic's position is therefore a difficult one to maintain. This does not mean that we should therefore simply give all credit to the AI. What I have tried to show here is that some of the fundamental requirements for authorship are missing from the claim of the creator of the AI. Despite the strong intuitive claim, and the clearly large amounts of work and effort that are necessary to produce a functioning AI, a claim of full ownership by the creator over the output is very problematic.

The relation of a creator to the output of its AI system may be better understood in comparison to the relationship of a parent to their child. While it is true that parents have put a lot of effort into putting their child on the world, and the child can properly be called 'theirs' in a serious sense, the child is an autonomous creator too. Even at a young age, when the child is not fully developed and could perhaps be called unaware or unconscious of their actions and the meanings of these actions, they exhibit creative behaviour, being capable of authoring things themselves beyond what they

215 Boden, *The Creative Mind*, 96.

were taught by their parents. Parents showing others a drawing by their child, describing it as their own work, would raise some eyebrows.[216]

The analogy of children and their parents is not perfect, of course. Children are conscious or on their way to becoming so; they have feelings and desires. This is the fundamental difference with an AI in this context: a machine has no use for authorship. This matter of fact is also where the difficulty of an autonomous creative AI is situated. While it does generate a novel and creative output, these machines remain in the end an amalgamation of silicone, metal and code, and are not the kind of entity that needs incentives.[217] But, even so, the full appropriation of this output by the creator of the AI is not warranted from the view of a labour-desert theory of property.

3.6 Option 3: the data-owner as owner of the output of AI

A party that is not usually mentioned in the IP-related debate on human-created works, is the data owner[218]. To recognize the role and thus the claim of the data owner in the discussion on AI-generated works, some background information about the technical process of AI development needs to be mentioned first.

3.6.1 The role of data in the development of AI

A technique that has recently been very popular in AI research is machine learning. It has quickly become the dominant AI technique, representing “89 percent of patent families related to an AI

216 Some of these eyebrows could be raised out of jealousy by Dadaists. I mention this only to point out that even a child’s early drawings can bear significant creative value.

217 Pamela Samuelson, ‘Allocating Ownership Rights in Computer-Generated Works Symposium: The Future of Software Protection’, *University of Pittsburgh Law Review* 47, no. 4 (1986 1985): 1199.

218 I use the term “data owner” in a very broad sense, referring to whoever holds the intellectual property in the data used as input to generate an output. This can be inventive works, creative works, or any other form of abstract objects subject to intellectual property protection.

technique".[219] Due to its significance in current forms of AI, it is the technique that I will focus on here. Other techniques, such as certain forms of evolutionary algorithms[220], may not require data input the way machine learning does. In these cases, the data-owners will either be partially inhibited in their property claim, or lose the claim completely.

Machine learning, as generally defined, consists in teaching an AI to identify patterns in sets of data, which can afterwards be used on new data that the AI has not encountered before.[221] One type of machine learning generates models through training that are called "artificial neural networks", built up from different layers of algorithms in a way that is inspired by the biological structure of the human brain.[222] Deep neural networks, used in so called 'deep learning'[223], is the name given to networks with a high number of layers "whereby the scientific community does not agree upon a clear delineation".[224] According to some authors, this advanced layered structure "allows the machine to learn and make decisions on its own".[225]

These networks start out without any code that tells them how to perform the tasks that they are meant to handle in the end; they are not rule-bound. Instead, they are made to train on data, which is 'teaching' them how to complete their tasks, and improving their abilities over time.[226] In order to be able to train, the system is dependent only on a previously established architecture, consisting of

219 World Intellectual Property Organization, *WIPO Technology Trends 2019: Artificial Intelligence.* (Geneva: World Intellectual Property Organization, 2019), 41, https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5982426.

220 This is another subfield of AI that has been gaining popularity. These evolutionary algorithms use Darwinian evolutionary principles of artificial selection to identify the most suitable solution to a problem. They select from a set of autonomously generated possible answers. These are further improved by applying mechanisms of selection. See Drexl et al., 'Technical Aspects of Artificial Intelligence', 3, 11.

221 Drexl et al., 3.

222 Drexl et al., 5.

223 European Commission, 'Artificial Intelligence for Europe' (European Commission, April 2018), 10, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN.

224 Drexl et al., 'Technical Aspects of Artificial Intelligence', 6.

neurons and weights.[227] As mentioned in the previous section, these kinds of AI are therefore not bound by pre-established rules programmed by their creators, and can therefore output more than what is 'put in'. However, this does mean that they are dependent on data in order to produce an output.

The method of machine learning leads to AI's that are not transparent: "These networks do not follow the rules of probability; they do not deal with uncertainty in a rigorous or transparent way. Still less do they incorporate any explicit representation of the environment in which they operate. Instead, the architecture of the network is left free to evolve on its own. When finished training a new network, the programmer has no idea what computations it is performing or why they work."[228] This lack of transparency combined with the fact that the training happens automatically and is commonly detached from any direct human input, puts a barrier between the programmer and the output of the AI[229]: the programmers are no longer able to predict the outcome of the algorithm, and are unable to tweak it as they see fit. This is the autonomy referred to in the previous section, fundamentally undermining the creator's claim to the product of AI.

The data-owner's claim is of a different nature. The claim depends on the fact that the AI is produced using only their data, interpreted by a pre-made architecture. Craig and Kerr note that "it is crucial to understand that these machines are not islands. Their outputs depend upon, and are inextricably linked to, a vast sea of texts authored by human actions, interactions and creative

225 Gervais, 'The Machine As Author', 6.

226 Pearlman, 'Recognizing Artificial Intelligence (AI) as Authors and Investors under U.S. Intellectual Property Law', 28.

227 "Each neuron is a mathematical function which transforms inputs (the numeric value of the upstream weights) into an output (the numeric value of the downstream weights). The model is composed of the sum of all the functions entailed in the neurons. […] The weights are trainable parameters. They are numeric values that are first randomly allocated and then optimised during the training process." Drexl et al., 'Technical Aspects of Artificial Intelligence', 5–6.

228 Pearl and Mackenzie, *The Book of Why*, 14.

229 Gervais, 'The Machine As Author', 6.

processes".[230] While in the future the human-made architecture may also be generated by artificial intelligence[231], the role of externally provided data seems difficult to replace. Even though the possibility of artificial systems which work with their own randomly generated input exists, for example in evolutionary algorithms[232], systems which train on existing human-made works have a clear advantage, which may be accountable for their successes.[233] The dependence of new artistic or inventive works on existing works seems to make artificial creativity without a link to human creativity challenging. This makes sense: in order to be novel or useful, reference to works made before seems indispensable[234], for AI as well as for humans. To return to Boden's terminology, in order to achieve historical creativity, knowledge of previous works is not imperative, but it certainly is of great value.[235]

Contrary to humans, however, the AI's ability to produce valuable work can seem like it is entirely dependent upon the input of human creative works. As Gudkov postulates, "AI does not create work but merely observes existing content and rules to construct new content under the principles of composition".[236] In other words, the output of the AI could be seen as no more than an intricate 'assemblage' or 'collage' of existing works.

3.6.2 Problems with the data-owner's claim

The data-owner makes a strong case. If all an AI does is taking parts of the thousands of works that it has trained on and re-ordering them according to some mathematical formula, the creators of the original works can be seen as the ones who have laboured for the final value.

230 Craig and Kerr, 'The Death of the AI Author', 26.

231 Drexl et al., 'Technical Aspects of Artificial Intelligence', 6.

232 Drexl et al., 11.

233 European Commission, 'Artificial Intelligence for Europe', 10.

234 Aleksei Gudkov, 'Robot on the Shoulders of Humans', *The Journal of World Intellectual Property* 23, no. 5–6 (2020): 764–765, doi:https://doi.org/10.1111/jwip.12172.

235 See 3.5.2 above or Boden, *The Creative Mind*, 43–44.

236 Gudkov, 'Robot on the Shoulders of Humans', 765.

Several arguments can be made against this claim. First and foremost, comparing this judgment of AI's to the judgment of human creators, this may seem unfair, and another instance of a double standard. Humans are allowed to take inspiration from existing works when creating their own work. The result of this process of inspiration, no matter how many existing works the creator has consulted, is undoubtedly the property of the creator. If not, intellectual property would become an impossibility, as an author could always be said to draw inspiration from works she has seen before. As an author, you cannot decide to forget about the work of another author that you saw before, simply because it is copyrighted or patented. All that is required is that the final work does not copy another that has come before.

This is where the data-owner could frame a rebuttal: while human authors are capable of creating something completely new, artificial systems may not be.[237] But what is this claim based on? Is the gap between human and machine intelligence really so big that it cannot be bridged? Of course, we cannot lay bare the functioning of human creativity and simply compare it to the inner workings of machine creativity. Human creativity is not sufficiently understood, does not bear a real single definition, and artificial intelligence is, in the case of machine learning, not transparent.[238] What we can do, however, is to look at this required novelty, and see if an AI can provide it too. In 3.5, we established that AI does seem capable of creativity. This, however, is supposing that what the AI generated was not a simple copy from its training input. Here, we'll need to look further into the concepts of novelty and originality. Where does the border between a copied work and a new work lie, and on which side of this border can we locate artificially generated works?

237 See Timothy L. Butler, 'Can a Computer Be an Author? Copyright Aspects of Artificial Intelligence', *(Comm/Ent), A Journal of Communications and Entertainment Law* 4, no. 4 (1982 1981): 735.

238 W. Nicholson II Price, 'Big Data, Patents, and the Future of Medicine', *Cardozo Law Review* 37, no. 4 (2016 2015): 1404.

Novelty and originality

Looking at the legal cases, the requirements of novelty and originality are often only vaguely defined. As Bridy notes, U.S. law took a sharp turn regarding copyright law in the early 20th century, before dropping the requirements significantly, setting a new precedent with *Bleistein v. Donaldson Lithographing Co.*[239], where "the standard for originality in copyright law reached a low watermark", meaning that ascribing originality only meant that a work was not copied from a work by another author.[240] Before this, in cases like *Burrow-Giles Lithographic Co. v. Sarony*, terms like "genius", "intellect", and "imagination"[241] were common in cases concerning IP law.[242] Even later, in 1991 with *Feist Publications, Inc. v. Rural Telephone Service Co.*, almost every trace of these terms was gone, with a statement by the U.S. Supreme Court that "copyright protects only those constituent elements of a work that possess more than a *de minimis* quantum of creativity".[243] At every point in this progression in IP law, however, central terms like creativity, originality, novelty, and authorship were ambiguously handled and left undefined.

Yet, despite these vague notions, a clear evolution can be identified in the concept of authorship.[244] Where at first a romantic notion of authorship ruled, it was gradually broken down, until it encompassed only the minimal requirement for originality. If we interpret this minimal requirement of the contemporary legal notion of authorship as the demand that the new work does not completely resemble the works it was inspired by – or, in terms of machine learning, the works that the AI trained on – then the prospect of novelty by artificial creation regains its plausibility.[245] AI

239 Bleistein v. Donaldson Lithographing Co., 188 U.S. 239 (U.S. Supreme Court 1903).

240 Bridy, 'Coding Creativity', 5–6.

241 Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (U.S. Supreme Court 1884).

242 Bridy, 'Coding Creativity', 7.

243 Feist Publications, Inc. v. Rural Telephone Service Co. (U.S. Supreme Court 1991) at 363.

244 Bridy, 'Coding Creativity', 8–9.

245 I refer to legal works to make this point not because I want to define the legal requirements for intellectual property; that is not the goal of this text. Rather, I use legal discussions as an example of a broader concept of authorship, which is present in both everyday intuitive discussions and theoretical debates.

systems have proven their ability to create works which are fundamentally different from their input, to the point where the influence from individual works trained on by the AI has become unrecognizable.[246]

Even of AI systems which are meant to imitate a single artist, often at most we can say that they imitate their style. Take for example “The Next Rembrandt”, a project by researchers from Microsoft and Delft University of Technology, who tried replicating Rembrandt’s style of painting.[247] They were careful to make sure that the result would not just be an average of the features present in Rembrandt’s original paintings[248], and presented their AI with a new subject.[249] This is important: using machine learning, one can create an algorithm which does not simply fuse the data from its training input into a new order. Much like human creativity, the process foregoing the generation of an output is non-transparent, and can be described as a complicated form of association[250], combining inspiration from previously observed works in a new and unique way.[251] The end result of “The Next Rembrandt” bears a startling similarity in style to Rembrandt’s real works, but is not simply a copy or a collage of those paintings. Claiming that this output is directly derivative would therefore seem unjust.

For some types of creative works, being taken apart and put together again is about as much as can be expected. Endless variations of the same narrative parts in different orders may exist in different stories, but this doesn’t make each variation equal to all the others, let alone a copy. For music, too,

246 See Gervais, ‘The Machine As Author’, 7–8; and Butler, ‘Can a Computer Be an Author?’, 743–744.

247 See ‘The Next Rembrandt’, The Next Rembrandt, accessed 3 August 2020, https://www.nextrembrandt.com.

248 du Sautoy, *The Creativity Code*, 119.

249 Guadamuz, ‘Do Androids Dream of Electric Copyright?’, 2.

250 Boden, ‘Creativity and Artificial Intelligence’, 349.

251 Butler, ‘Can a Computer Be an Author?’, 743–744.

the order of existing notes is exactly what makes a piece new.[252] One cannot accuse an AI of simply assembling this order randomly: the algorithm, although opaque, is not random.

The romantic notion of authorship

One could, despite these similarities between human and artificial novelty, still cling to the romantic notion of authorship, and demand more. After all, ascribing "genius", "imagination", etc. to an AI is much more challenging than recognizing basic originality or novelty in its output. For any artificial system, consciousness will likely remain impossible to reach, which makes creativity in this sense difficult to attain.[253] But before coming to any conclusions, let us take a look at what this romantic sense of authorship and creativity really entails.

Margaret Boden splits the romantic sense of authorship in two, describing two widespread views which she calls the 'inspirational' and the 'romantic' views of authorship. These two views "assume that creativity, being humanity's crowning glory, is not to be sullied by the reductionist tentacles of scientific explanation. In its unintelligibility it is splendour."[254] A use of the terms authorship and creativity in this sense echo in the legal cases mentioned earlier. Creativity is often ascribed or required, but rarely really defined.[255]

Attempts to make the romantic view less obscure are in danger of exposing its emptiness: as Gudkov describes it, "[o]riginality takes root in creativity, which, in turn, is based on imagination. Imagination is the act or power of forming a mental image of something not present to the senses.

252 See Deltorn and Macrez's discussion of musical composition by artificial intelligence. Jean-Marc Deltorn and Franck Macrez, 'Authorship in the Age of Machine Learning and Artificial Intelligence', *SSRN Electronic Journal*, 2018, 14–15, doi:10.2139/ssrn.3261329.

253 See 3.4.1 for a more comprehensive discussion on AI and consciousness.

254 Boden, *The Creative Mind*, 14.

255 See for example Gervais, 'The Machine As Author', 43.

Imagination is based on novelty."[256] This definition of creativity can be brought back to a requirement of simple novelty (by lack of definition by the author taken to mean simply the uniqueness of an idea), but with an added requirement of having been processed first by imagination. But what is the point of this extra step involving imagination? And why does it justify authorship? Can it even be applied to all forms of abstract objects, outside of only the visual kind?

Additionally, this particular example of a definition could be exhibiting circularity: if creativity and originality are based on novelty, how does novelty come to be? Understanding creativity as something that creates something new out of nothing, as it is often understood, produces a paradox: "[e]ither what preceded it was similar, in which case there is no real novelty. Or it was not, in which case one cannot possibly understand how the novelty could arise from it."[257] Either this novelty becomes impossible, or it remains mysterious, explainable only by ascribing it to some sort of "magic".[258]

If these arguments against the romantic conception of authorship are rejected, then there still remains the possibility for an AI to train only on works from the public domain. In Lockean terms, the new work would be taking only from the commons, and therefore affecting anyone else's property claims could be prevented.

If we cast aside the romantic view, however, the AI may even have an advantage over humans when it comes to generating original or novel works, through its ability to process thousands of works. That is, if we ascribe creativity not to "any special power, but greater knowledge (in the form of practised expertise) and the motivation to acquire and use it."[259]

256 Gudkov, 'Robot on the Shoulders of Humans', 764.

257 Boden, *The Creative Mind*, 11–12.

258 Boden, 12.

259 Boden, 35.

3.6.3 The remaining case of the data-ower

The claim of the data owner rests on a mystical conception of originality, based on a romantic – and equally mystical – conception of authorship. Ascribing this obscure notion not only makes accurately ascribing authorship impossible, leaving many human creators to possibly not make the threshold, but it also creates justificatory problems and possible circularity.[260]

In labour terms, the data-owner's claim is no more defensible either. While the data used as input for the training of an AI is essential in order to produce the final output, the owner of the input data cannot be said to be the one generating the extra value produced by the AI, nor is the owner of the input data expending any extra labour in order to produce the output of the AI. With the romantic notion of authorship failing, and the data owner not being able to claim a position of authorship over the artificially generated work, even desert from need cannot apply here, as it depends on the identity of an author.[261]

So, thinking of creativity as some cryptic and unconscious faculty in the mind which allows those who possess it to create new works out of nothing, without dependency on works that came before, is unrealistic and difficult to maintain after an analysis. But, this conclusion also aligns with existing psychological research on the matter. Robert W. Weisberg and Lauretta M. Reeves, for example, have a similar assessment of the matter, arguing that a romantic or 'genius' notion of creativity, "the idea that creative thinking depends on extraordinary processes" finds little support in research.[262] In other words: it does not seem to be how human creativity functions. Instead, they conclude "that there is nothing extraordinary about the cognitive processes underlying creative thinking (although those thought processes may result in extraordinary products)"[263]. Weisberg also

260 This depends on the interpretation of the romantic notion of authorship, which, due to its obscure nature, is a challenging matter.

261 For an overview of the arguments concerning labour theories of intellectual property, see chapter 2.

262 Robert W. Weisberg and Lauretta M. Reeves, *Cognition: From Memory to Creativity* (John Wiley & Sons, 2013), 603–604.

263 Weisberg and Reeves, 603.

proposes that creativity, from a psychological perspective, depends on input from existing works: "The new must begin as a variation on old themes."[264] Or, as David Cope concludes: "'New' art, then, consists of a reassembly of already existing art."[265]

With the loss of plausibility of a romantic notion of a creative genius, involving the creation of new and original works from nothing, the claim of the data-owner becomes obsolete. If human authorship depends so heavily on inspiration from existing works itself, the data-owner can no longer argue that the AI's dependency on their input signifies that authorship (and ownership) of the output belongs to whoever created the input. If one were to handle this notion consistently, applying it to AI's as well as humans, human authorship in most forms would become impossible, as it is dependent on previous input too.[266] In other words: the argument in support of the data-owner's claim does depend on a double standard.

This also translates into a more explicit labour-desert framework. The labour performed by the creator of the input is not enough to account for the output of the AI system. If the labour performed by a human artist can consist of creativity in a non-romantic sense, and if this labour is enough and necessary to justify intellectual property rights in the case of a human-created work, then the fact that this labour was not performed by the data-owner makes any their appropriation of the output unsound.[267]

Combining the conclusions of this discussion with those offered in 3.5, we come to a view of AI creativity that characterizes an AI-generated output as potentially independent from both the creator of the AI and the creator of the input used in the training of the AI. Or, at least, this output is marked as being sufficiently independent and autonomously creative to make an authorship claim by the

264 Robert Weisberg, *Creativity: Beyond the Myth of Genius* (New York, NY, USA: WH Freeman, 1993), 21.

265 Cope, *Computer Models of Musical Creativity*, 28.

266 See Craig and Kerr, 'The Death of the AI Author', 27–28.

267 See 2.4.3 for a more elaborate discussion of this labour-desert terminology.

creator and the data-owner problematic from the perspective of a labour-desert justification of intellectual property.

A last problem with the data-owner's claim that I haven't mentioned above derives from the ability of an AI system to train on hundreds of thousands of existing works. This could mean that the owners of these thousands of works would all have a potential claim on the final output (ignoring, for the sake of the argument, the conclusion I made here). Is such a shared claim over the output of an AI credible from a labour-desert point of view? I will discuss these arguments in the next section.

3.7 Option 4: contract-based ownership of the output of AI

The final option I will consider here is a contract-based approach. This option entails the use of a contract to define who get a share of the ownership over the output of the AI, or a share over the gain thereby produced. The use of this approach could imply that the ownership of the final product is either shared in some way between different parties, or is traded or handed over from one party to another (or, of course, several). This introduces some new potential stakeholders to the discussion, namely buyers, investors, end-users, and many more. In effect, the group of possible parties concerned with the property rights becomes essentially boundless, and is not limited to people who are directly involved in the production of the output. The possibility of free transfers of property rights makes anyone a candidate to receive property rights, if they are freely given.

This option could avoid or solve many issues encountered by applying any of the other three options. The labour performed by all three parties is in some way necessary to produce the artificially generated output.[268] Through a contract, all three stakeholders could be ascribed joint

268 In case of the data-owner, the labour may have been performed without the explicit goal of contributing to the AI output. Still, it is necessary to make the output possible.

ownership over the final product, and as such all their performed labour could be acknowledged and rewarded.

3.7.1 Alienation and free exchange of property

The right to transfer one's property to another freely, and therefore alienate it from oneself, is present in many western conceptions of property, but is not a necessary characteristic. Much like rules about inheritance and bequest, societies may put limits on the free exchange of goods. So, a theory of property does not need to include free exchange or alienation of property in the rights associated with property, and there may be theories which do not include it.[269] Despite this, in most cases, a right to free (or mostly free) exchange of property rights is one of the basic rights associated with property. As Waldron notes, "it would be wrong not to recognize that the link between ownership and alienation is somewhat tighter than the connection which ownership has with inheritance and bequest."[270] The free market is a convenient and workable solution to the problem of distribution under changing circumstances of scarcity. While a fair distribution of goods can be theorised, reapplying it each time circumstances change (for example through uneven consumption, loss of value, the addition of new individuals) is difficult to maintain. As Waldron points out:

> "redistribution of that sort has its costs, as many of its critics have pointed out: it disappoints expectations, it undermines security and stability, and it leaves people without the ability to undertake long-term planning of resource use except to the extent that they can prophesy changes in social circumstances and how the society will respond to them."[271]

269 Waldron, *The Right to Private Property*, 53–54.

270 Waldron, 54.

271 Waldron, 54.

For these reasons, the right to alienate and exchange property is usually included in labour theories too.[272] I will look at Robert Nozick as an example of a theory which does include the right to alienation of exchange in property. Due to Nozick's large influence on labour theories and their similarity in the way they handle property as a right[273], arguments related to Nozick's reasoning can often be generalised to other labour theories.

Nozick puts a strong emphasis on the right to free transfer of property, defining it as an essential part of his principle of justice in holdings: "A person who acquires a holding in accordance with the principle of justice in transfer, from someone else entitled to the holding, is entitled to the holding."[274] Just property, in Nozick's view, can only result from either just acquisition or through a just transfer.[275] In terms of labour, this means that property either comes from original appropriation, by labouring on something unowned, or from a transfer of property (originating directly or indirectly from appropriation by labour), preceded by consent from the just owner.

Consent

Following this account of the right to exchange, consent from the just owner is necessary to be able to speak of a just transfer of property. Applying this to our case of AI is difficult. Above, I have argued that neither the AI, nor the creator of the AI or the owner of the input data can justifiably make a full claim over the property rights of the AI from a labour-desert perspective. Who is then to give consent for the exchange by contract in the first place?

---

272 See for example Daniel Stengel, 'Intellectual Property in Philosophy', 2020, 30. and Stephen R. Munzer, *A Theory of Property* (Cambridge University Press, 1990), 276–279. Munzer notes that in a labour theory, the assumption that property rights would not be transferable "is an unrealistic assumption, for an important feature of most legal property rights is transferability. One may, though, drop this unrealistic assumption without precluding transfer so long as certain moral restrictions […] are imposed."

273 See Kelly, *Locke's Second Treatise of Government*, 25.

274 Robert Nozick, 'Distributive Justice', *Philosophy & Public Affairs* 3, no. 1 (1973): 47.

275 Nozick, 47.

To bypass this quandary, the proponent of a desert-from-labour theory could argue to drop the requirement of full authorship, replace it by a determination of contribution in labour or produced value (which of the two is of little matter to the argument), and distribute the property rights proportionately. This approach would be met by another problem: however would one determine the proportions of contributions?[276]

This situation of production involving various interdependent labourers, none of whom are in isolation fully capable of attaining the final product, is not new, nor is it rare.[277] In the simple and common case of the production of a book, for example, several stages of production are necessary. There is the writer, who may be typing a manuscript on a computer, running software developed by programmers, communicating over the Internet running on privately-owned servers with the publisher, who works together with graphical designers, lay-outers, printers, distributors, and many others. The book gets printed by machines developed and built by other people, operated by someone else. One could expand this chain of causality ever further, including the carpenter who made the chair used by the marketeer working for the publisher, or including whoever first taught the writer the alphabet, all the way to the inventors of the alphabet, where it does not end either.

Each of these people's labour contributed to the production of the output of the AI in a causal manner. Their labour was necessary for the AI's output to be possible. As Craig and Kerr note, preceding any creative output by the machine, there is an "entire array of creative efforts and communicative expressions of prior generations of authors and artists engaged in an ongoing

276 Nozick's remark on pouring tomato juice into the sea could illustrate this point (although, clearly, it was not originally meant to do so). After the tomato juice has mixed with the sea, how could one possibly still determine who produced which amount of value? The sea can be valuable in many different ways, and so can tomato juice. The ultimate value produced, and the outcome of the mixing, are both determined by many causal factors. The final picture is therefore virtually impossible to interpret in terms of produced value. For the original argument, see Nozick, *Anarchy, State, and Utopia*, 175.

277 See Drahos, *A Philosophy of Intellectual Property*, 62.

dialogue with others in their communities."[278] This applies beyond just artistic works though, and holds for every form of abstract object. As Alexander and Peñalver argue: "At the most basic level, any intellectual activity depends upon the existence of communicative systems, such as language or mathematics, developed over thousands of years by countless human beings."[279]

A consequentialist theory of property can deal with this situation without any issues. Whatever distribution of rights produces the best outcome (often interpreted in a sense of maximising incentive[280]), is the best option. A delineation of who performed labour in producing the object of property is unneeded, apart from identifying those who need to be given incentive to perform the labour again. A labour theory would normally deal with this situation by identifying one person (or a few) who can be attributed the most essential part of the labour, a part that requires the creativity and knowledge that is required from an author.[281]

In other words, normally, outside of the author(s), no one can present a valid claim to the IPR's. Their labour may have contributed to the production of the final product (as in the example of the inventor of the alphabet), but they are not the authors. Acquiring the function of an author is essential here. If a property right to an abstract object is no longer confined only to its author, every person who had some input in the creation of the final product suddenly gets a valid claim according to a labour theory of property.

When the requirement of authorship is dropped, there remains no way to assess which labour is sufficient for a property claim. In a labour theory, this gap would naturally be filled by others who

278 Craig and Kerr, 'The Death of the AI Author', 27–28.

279 Alexander and Peñalver, *An Introduction to Property Theory*, 194.

280 See for example Moore, 'A Lockean Theory of Intellectual Property', 65; Stanley M. Besen and Leo J. Raskind, 'An Introduction to the Law and Economics of Intellectual Property', *Journal of Economic Perspectives* 5, no. 1 (March 1991): 5, doi:10.1257/jep.5.1.3; Anne Lauber-Rönsberg and Sven Hetmank, 'The Concept of Authorship and Inventorship under Pressure: Does Artificial Intelligence Shift Paradigms?', *Journal of Intellectual Property Law & Practice* 14, no. 7 (July 2019): 576, doi:10.1093/jiplp/jpz061.

281 As you may notice, a desert-from-excellence theory makes this easier than a desert-from-value notion of property.

laboured for the product. However, if we let a vague notion of causal necessity be sufficient for intellectual property rights to an abstract object, then it becomes impossible to delineate who must be included in these rights, and the list of potential claimants becomes endless. Without authorship, "once we admit that all intellectual creation draws at least to some extent on the prior intellectual labor of others, the Lockean case for intellectual property finds itself in the same murky domain of shared credit as the Lockean case for ownership of tangible property."[282]

This leads us to a paradox: while the notion of an author is not strictly necessary for a labour-desert theory, as it can in theory rely on the different contributions of labour to define the appropriate desert, omitting the concept of an author leads to the desert being spread almost infinitely over an immeasurable amount of contributors, making it impossible to assign a single appropriator, let alone a small group of owners. Claiming (partial) responsibility in the authorship of a creative or inventive work is therefore dependent upon a complicated structure of presuppositions about responsibility, creativity, value and causality.

3.7.2 Conclusions on the contract-based approach

The solutions offered by a contract are promising, but short-lived. A contract does not require a single author to be selected. This is convenient, especially in the situation of the output of an AI, where no single contributor can be ascribed full authorship. Yet, in dropping the requirement of an author, the conditions are created so that anyone whose labour somehow contributed to the final result now has a valid property claim. This, I have argued, creates a paradox: labour-desert can do without authorship in theory; but when an author is not identified, private intellectual property becomes an impossibility in a labour theory.

282 Alexander and Peñalver, *An Introduction to Property Theory*, 194.

This paradox is indicative of a larger problem with labour-desert theories, and labour theories in general. Labour is typically difficult or impossible to measure and accredit.[283] Focussing on the value it produces does not clear things up, as Alexander and Peñalver conclude. They refer to Nozick's famous Wilt Chamberlain example. Nozick argues that Chamberlain, a famous basketball player of his time, may sign a contract with a team, stating that he gets a percentage of the income made from ticket sales for the basketball games that he will play at. If people are willing to pay to see him play, Chamberlain seems to be entitled to this income.[284] Alexander and Peñalver argue that even though this seems fair at first, it is impossible to discern the value produced by Chamberlain from the value produced by everyone else whose labour was necessary to produce this value, including the inventors of basketball itself.[285]

With these considerations in mind, a contract approach does not seem capable of solving the issues AI presents for labour-desert theories of intellectual property. Overall, a labour-desert theory seems to be lacking the necessary concepts to tackle these problems.

283 See Peter Jaszi, 'Toward a Theory of Copyright: The Metamorphoses of Authorship', *Duke Law Journal* 1991 (1991): 455.

284 Nozick, *Anarchy, State, and Utopia*, 161.

285 Alexander and Peñalver, *An Introduction to Property Theory*, 195.

Chapter 4 - Conclusion

The act of labouring seems to carry a kind of profound value which has earned it a central part in many philosophies, ranging from Marxist to libertarian. Locke's attempt to derive from it the justifying principle of property has influenced generations of thinkers, who have in turn tried to influence property theory by developing their own interpretations. For 300 years, this project has persevered. Throughout this time, the labour theory of property has adapted to take many forms. Now, it seems, a new challenge is on the horizon, which could either require it to evolve once more, or prove to be another argument against Locke's original ambition.

This thesis has aimed to provide an answer to the question of who should get the intellectual property rights to the output of an autonomous artificial intelligence from the perspective of a Lockean theory of property. Similar questions have been asked before by other authors, but mostly their answers focussed on a consequentialist theory of property.

Labour theories of property

The result of this attempt can be divided into two parts. In the first part, I looked at Lockean theories of property, which problems presented themselves, and which interpretations were most capable of avoiding these problems. Locke's theory by itself if insufficient to tackle the problem of AI authorship. However, an interpretation from a desert-from-labour perspective can account for the special features of abstract objects, while also providing an answer to the justificatory gap between labour and ownership left by Locke. A dependence on incentive could be omitted by taking desert as a basis, while such a decision can also account for the right to keep others from copying non-rivalrous objects. Additionally, being able to refer to three sources of desert (excellence, value, and

need) provides the necessary flexibility to justify political or contractual adjustments to property rights.

AI, authorship, and creativity

In the second part, autonomous artificial intelligence was introduced to the equation. Its necessary dependencies lead to three parties having contributed to the final output of the AI: the AI itself, the creator of the AI, and the owner of the input data. These three parties can therefore present some claim to the property rights. A fourth option comes from the possibility of defining and sharing ownership through a contract.

The conditions necessary for an AI to receive ownership, either by it attaining consciousness or through a redefinition of labour, I have argued, are highly implausible. Most of all, the profoundly social-historical nature of the concept of authorship would be denied by ascribing it to an AI, and the concept would lose its meaning.

Granting the creator of the AI ownership (and therefore authorship) would equally hollow out the notion of an author. With autonomous AI creation, it would be possible for the creator of the AI to have had no creative contribution to the output, granting that this contribution could instead be provided by the AI itself. As a singular definition of creativity is unsuccessful, I have argued that creativity exists in several forms, some of which an AI is capable of attaining and reproducing. So, while the creator is capable of creating a creative AI system, she does not have a valid claim to its creative output.

While data clearly plays a fundamental role in the development of autonomous AI, the owner of this data does not present a justified claim either, as it depends on an outdated romantic notion of originality and authorship, entailing unrealistic connotations of geniuses creating *ex nihilo*. This vague notion, when analysed or attempted to be grasped in a definition, creates justificatory and logical problems. Humans, like AI's, depend on an input in order to create an original and creative work. This dependency does not imply that the produced works are directly derivative from the input. If it did, human creativity and authorship would become impossible too. Therefore, a double

standard is present. To be consistent, it must be recognized that the contribution of a machine makes ascribing full authorship to the owner of the input data impossible.

A contract-based approach to determining ownership, while providing some solutions to the previously mentioned problems, fails too. As none of the other stakeholdes can claim full authorship, no one can provide the consent necessary to make an exchange through contract possible. If authorship is dropped as a requirement for intellectual property, there is no way to discern the contributions (through labour) of the three main stakeholders from those made by all the other people whose labour was causally necessary to make the output possible.

Concluding remarks

Some of the issues I have touched upon in this thesis deserve a more thorough examination. Machine agency, creativity, authorship, novelty and originality are only some of the concepts which could each have filled a thesis by themselves. Although my discussion was necessarily limited in many ways, I do believe some modest conclusions can be derived from it.

It has become clear that a labour theory of property seems to lack some of the concepts necessary to speak clearly about intellectual property. This lack of concepts is also present in the discussion on artificial intelligence. Increasing automation will likely prove to be a great challenge for labour theorists.

While the problems I have described are perhaps not insurmountable for a labour theory, they do present another set of issues for authors who defend it. Many of these are related to the way a labour theory deals with intellectual property. The sometimes seemingly arbitrary step from labour to property (see my discussion of this justificatory “gap” in chapter 2) has led several authors to argue

that a labour theory of property may be better interpreted in consequentialist terms.[286] While this is certainly too big of a judgment to be inferred from the arguments presented here, they do add strength to it.

Émile Borel did not stop to think about the consequences that his typing monkey would produce for property rights. Admittedly, he did not have any reason to do so. Now, technological advancements have lead us far beyond the point of a simple monkey with a typewriter. It is time to consider how we will meet these consequences as a society. I hope this thesis has been able to elucidate this matter. It seems that, despite its long history and its intuitive appeal, a labour theory of property is unlikely to be the most suitable candidate to provide us with answers.

286 See Waldron's comparison of a desert theory to utilitarian theories in Waldron, *The Right to Private Property*, 202., Drahos' comment that Locke's right of property is primarily an instrumental right Drahos, *A Philosophy of Intellectual Property*, 63–64. and Epstein's explicit dependency on consequentialist premises, see Alexander and Peñalver, *An Introduction to Property Theory*, 53.

Bibliography

Alexander, Gregory S., and Eduardo M. Peñalver. *An Introduction to Property Theory* (Cambridge University Press, 2012).

Annis, David B., and Cecil E. Bohanon. 'Desert and Property Rights'. *The Journal of Value Inquiry* 26, no. 4 (1992): 537–546.

Asaro, Peter M. 'What Should We Want from a Robot Ethic?' *The International Review of Information Ethics* 6 (2006): 9–16.

Attas, Daniel. 'Lockean Justifications of Intellectual Property'. In *Intellectual Property and Theories of Justice*, edited by Axel Gosseries, Alain Marciano, and Alain Strowel (London: Palgrave Macmillan UK, 2008), 29–56. https://doi.org/10.1057/978-0-230-58239-2_2.

Becker, Lawrence C. 'Deserving to Own Intellectual Property Symposium on Intellectual Property Law Theory'. *Chicago-Kent Law Review* 68, no. 2 (1993 1992): 609–630.

———. 'Property Rights: Philosophic Foundations'. *Books by Hollins Faculty and Staff*, January 1977. https://digitalcommons.hollins.edu/facbooks/76.

———. 'The Labor Theory of Property Acquisition'. *The Journal of Philosophy* 73, no. 18 (1976): 653–664. https://doi.org/10.2307/2025823.

———. 'The Moral Basis of Property Rights'. *Nomos* 22 (1980): 187–220.

———. 'Too Much Property'. Edited by Jeremy Waldron and Stephen A. Munzer. *Philosophy & Public Affairs* 21, no. 2 (1992): 196–206.

Bell, Tom W. 'Indelicate Imbalancing in Copyright and Patent Law'. SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, May 2007). https://papers.ssrn.com/abstract=984085.

Berki, R. N. 'On the Nature and Origins of Marx's Concept of Labor'. *Political Theory* 7, no. 1 (1979): 35–56.

Besen, Stanley M., and Leo J. Raskind. 'An Introduction to the Law and Economics of Intellectual Property'. *Journal of Economic Perspectives* 5, no. 1 (March 1991): 3–27. https://doi.org/10.1257/jep.5.1.3.

Bleistein v. Donaldson Lithographing Co., 188 U.S. 239 (U.S. Supreme Court 1903).

Boden, Margaret A. 'Creativity and Artificial Intelligence'. *Artificial Intelligence*, Artificial Intelligence 40 years later, 103, no. 1 (August 1998): 347–356. https://doi.org/10.1016/S0004-3702(98)00055-1.

———. *The Creative Mind : Myths and Mechanisms*. 2nd ed. (Routledge, 2004). https://doi.org/10.4324/9780203508527.

Borel, Émile. 'La Mécanique Statique et l'irréversibilité'. *Journal de Physique Théorique et Appliquée* 3, no. 1 (1913).

Bridy, Annemarie. 'Coding Creativity: Copyright and the Artificially Intelligent Author'. *Stanford Technology Law Review* 2012 (2012): 2–28.

Bryson, Joanna J. 'Robots Should Be Slaves'. *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, 2010, 63–74.

Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (U.S. Supreme Court 1884).

Butler, Timothy L. 'Can a Computer Be an Author? Copyright Aspects of Artificial Intelligence'. *(Comm/Ent), A Journal of Communications and Entertainment Law* 4, no. 4 (1982 1981): 707–748.

Coeckelbergh, Mark. 'Robot Rights? Towards a Social-Relational Justification of Moral Consideration'. *Ethics and Information Technology* 12, no. 3 (September 2010): 209–221. https://doi.org/10.1007/s10676-010-9235-5.

Cole, David. 'The Chinese Room Argument'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2020. (Metaphysics Research Lab, Stanford University, 2020). https://plato.stanford.edu/archives/win2020/entries/chinese-room/.

Cope, David. *Computer Models of Musical Creativity* (MIT Press Cambridge, 2005).

Craig, Carys J. 'Locke, Labour and Limiting the Author's Right: A Warning against a Lockean Approach to Copyright Law'. *Queen's Law Journal* 28, no. 1 (2003 2002): 1–60.

Craig, Carys J., and Ian R. Kerr. 'The Death of the AI Author'. *SSRN Electronic Journal*, 2019. https://doi.org/10.2139/ssrn.3374951.

Cwik, Bryan. 'Labor as the Basis for Intellectual Property Rights'. *Ethical Theory and Moral Practice* 17, no. 4 (August 2014): 681–695. https://doi.org/10.1007/s10677-013-9471-y.

Damstedt, Benjamin G. 'Limiting Locke: A Natural Law Justification for the Fair Use Doctrine Note'. *Yale Law Journal* 112, no. 5 (2003 2002): 1179–1222.

Darling, Kate. 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects'. In *Robot Law*, by Ryan Calo, A. Froomkin, and Ian Kerr, 213–232 (Edward Elgar Publishing, 2016). https://doi.org/10.4337/9781783476732.00017.

Davies, Colin R. 'An Evolutionary Step in Intellectual Property Rights – Artificial Intelligence and Intellectual Property'. *Computer Law & Security Review* 27, no. 6 (December 2011): 601–619. https://doi.org/10.1016/j.clsr.2011.09.006.

Davis, Randall. 'Intellectual Property and Software: The Assumptions Are Broken.' (Massachusetts Institute of Technology Artificial Intelligence Laboratory, 1991).

DeFrancis, John. *The Chinese Language: Fact and Fantasy* (University of Hawaii Press, 1986).

Deltorn, Jean-Marc. 'Deep Creations: Intellectual Property and the Automata'. *Frontiers in Digital Humanities* 4 (2017). https://doi.org/10.3389/fdigh.2017.00003.

Deltorn, Jean-Marc, and Franck Macrez. ‘Authorship in the Age of Machine Learning and Artificial Intelligence’. *SSRN Electronic Journal*, 2018. https://doi.org/10.2139/ssrn.3261329.

Díaz-Noci, Javier. ‘Artificial Intelligence Systems-Aided News and Copyright: Assessing Legal Implications for Journalism Practices’. *Future Internet* 12, no. 5 (May 2020): 85. https://doi.org/10.3390/fi12050085.

Drahos, Peter. *A Philosophy of Intellectual Property* (Routledge, 2016).

Drexl, Josef, Reto Hilty, Francisco Beneke, Luc Desaunettes-Barbero, Michèle Finck, Jure Globocnik, Begoña Gonzalez Otero, et al. ‘Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective’. SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 2019). https://doi.org/10.2139/ssrn.3465577.

Erden, Y. J. ‘Could a Created Being Ever Be Creative? Some Philosophical Remarks on Creativity and AI Development’. *Minds and Machines* 20, no. 3 (August 2010): 349–362. https://doi.org/10.1007/s11023-010-9202-2.

European Commission. ‘Artificial Intelligence for Europe’ (European Commission, April 2018). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN.

European Patent Office. ‘Guidelines for Examination: Patentability’. European Patent Office. Accessed 23 December 2020. https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_i_1.htm.

Feist Publications, Inc. v. Rural Telephone Service Co. (U.S. Supreme Court 1991).

Fried, Barbara. ‘Does Nozick Have a Theory of Property Rights?’ In *The Cambridge Companion to Nozick’s Anarchy, State, and Utopia*, edited by Ralf M. Bader and John Meadowcroft (Cambridge University Press, 2011), 230–252.

Gabora, Liane. ‘Cognitive Mechanisms Underlying the Creative Process’. In *Proceedings of the 4th Conference on Creativity & Cognition* C&C ’02 (New York, NY, USA: Association for Computing Machinery, 2002), 126–133. https://doi.org/10.1145/581710.581730.

Gallagher, Shaun. ‘The Moral Significance of Primitive Self-Consciousness: A Response to Bermúdez’. *Ethics* 107, no. 1 (1996): 129–140.

Gervais, Daniel J. ‘The Machine As Author’. SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, March 2019). https://papers.ssrn.com/abstract=3359524.

Goertzel, Ben. ‘Thoughts on AI Morality’. *Dynamical Psychology: An International, Interdisciplinary Journal of Complex Mental Processes*, 2002.

Guadamuz, Andrés. ‘Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works’. SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, June 2017). https://papers.ssrn.com/abstract=2981304.

Gudkov, Aleksei. ‘Robot on the Shoulders of Humans’. *The Journal of World Intellectual Property* 23, no. 5–6 (2020): 759–776. https://doi.org/10.1111/jwip.12172.

Gunkel, David J. ‘The Other Question: Can and Should Robots Have Rights?’ *Ethics and Information Technology* 20, no. 2 (June 2018): 87–99. https://doi.org/10.1007/s10676-017-9442-4.

Gurkaynak, Gonenc, İlay Yılmaz, Türker Doygun, and Ekin İnce. 'Questions of Intellectual Property in the Artificial Intelligence Realm'. SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 2017). https://papers.ssrn.com/abstract=3295747.

Gürkaynak, Gönenç, Ilay Yılmaz, Türker Doygun, and Ekin Ince. 'Questions of Intellectual Property in the Artificial Intelligence Realm'. *Robotics Law Journal*, 2017.

Haas, Daniel. 'Merit, Fit, and Basic Desert'. *Philosophical Explorations* 16, no. 2 (June 2013): 226–239. https://doi.org/10.1080/13869795.2013.786971.

Hettinger, Edwin C. 'Justifying Intellectual Property'. *Philosophy & Public Affairs* 18, no. 1 (1989): 31–52.

Huang, Cheng-Zhi Anna, Curtis Hawthorne, Adam Roberts, Monica Dinculescu, James Wexler, Leon Hong, and Jacob Howcroft. 'The Bach Doodle: Approachable Music Composition with Machine Learning at Scale'. *ArXiv Preprint ArXiv:1907.06637*, 2019.

Hughes, Justin. 'The Philosophy of Intellectual Property'. *Georgetown Law Journal* 77, no. 2 (1989 1988): 287–366.

Ihalainen, Jani. 'Computer Creativity: Artificial Intelligence and Copyright'. *Journal of Intellectual Property Law & Practice*, 2018.

Jaszi, Peter. 'Toward a Theory of Copyright: The Metamorphoses of Authorship'. *Duke Law Journal* 1991 (1991): 455.

Kassan, Peter. 'Al Gone Awry: The Futile Quest for Artificial Intelligence'. *Skeptic (Altadena, CA)* 12, no. 2 (June 2005): 30–40.

Kelly, P. J. *Locke's Second Treatise of Government: A Reader's Guide* (Continuum, 2007).

Khoury, Amir. H. 'Intellectual Property Rights for Hubots: On the Legal Implications of Human-like Robots as Innovators and Creators'. *Cardozo Arts & Entertainment Law Journal* 35, no. 3 (2017 2016): 635–668.

Kirk, Robert. 'Decision, Control and Integration'. In *Zombies and Consciousness*, 2005.

Langley, Pat, and Randolph Jones. 'A Computational Model of Scientific Insight', December 1986. https://escholarship.org/uc/item/54x8v354.

Lauber-Rönsberg, Anne, and Sven Hetmank. 'The Concept of Authorship and Inventorship under Pressure: Does Artificial Intelligence Shift Paradigms?' *Journal of Intellectual Property Law & Practice* 14, no. 7 (July 2019): 570–579. https://doi.org/10.1093/jiplp/jpz061.

Lenat, Douglas B. '9 - The Role of Heuristics in Learning By Discovery: Three Case Studies'. In *Machine Learning*, edited by Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell (San Francisco (CA): Morgan Kaufmann, 1983), 243–306. https://doi.org/10.1016/B978-0-08-051054-5.50013-3.

Levy, David. *Robots Unlimited : Life in a Virtual Age* (A K Peters/CRC Press, 2005). https://doi.org/10.1201/b10697.

Lewin, Peter. 'Creativity or Coercion: Alternative Perspectives on Rights to Intellectual Property'. *Journal of Business Ethics* 71, no. 4 (April 2007): 441–455. https://doi.org/10.1007/s10551-006-9150-1.

Liang, Feynman T., Mark Gotham, Matthew Johnson, and Jamie Shotton. 'Automatic Stylistic Composition of Bach Chorales with Deep LSTM.' In *ISMIR*, 2017, 449–456.

Locke, John. ‘The First Treatise: The False Principles and Foundation of Sir Robert Filmer’. In *Two Treatises of Government and A Letter Concerning Toleration*, edited by Ian Shapiro (Yale University Press, 1689), 7–99. https://www.jstor.org/stable/j.ctt1npw0d.6.

———. ‘The Second Treatise: An Essay Concerning the True Original, Extent, and End of Civil Government’. In *Two Treatises of Government and A Letter Concerning Toleration*, edited by Ian Shapiro (Yale University Press, 1689), 100–210. https://www.jstor.org/stable/j.ctt1npw0d.7.

Loss Pequeno Glazier. *Digital Poetics: The Making of E-Poetries* (University of Alabama Press, 2001).

Lovelace, A. ‘Faster Than Thought’. In *Notes on Menabrea’s Sketch of the Analytical Engine Invented by Charles Babbage*, edited by B.V. Bowden (London: Pitman, 1953).

Marx, Karl. *Capital: A Critique of Political Economy - Volume 1* (London: Penguin, 1883).

Mendelovici, Angela. *The Phenomenal Basis of Intentionality* (New York, USA: Oxford University Press, 2018).

Merges, Robert P. ‘Locke’. In *Justifying Intellectual Property* (Harvard University Press, 2011), 31–67. https://doi.org/10.4159/harvard.9780674061125.

Mill, John Stuart. *Principles of Political Economy with Some of Their Applications to Social Philosophy*. Edited by John M. Robson. The Collected Works of John Stuart Mill. Vol. 2. 33 vols (Toronto: University of Toronto Press, Routledge and Kegan Paul, 1965). http://oll.libertyfund.org/title/102.

Monett, Dagmar, and Colin WP Lewis. ‘Getting Clarity by Defining Artificial Intelligence—a Survey’. In *3rd Conference on" Philosophy and Theory of Artificial Intelligence* (Springer, 2017), 212–214.

Moore, Adam D. ‘A Lockean Theory of Intellectual Property’. *Hamline Law Review* 21, no. 1 (1998 1997): 65–108.

Mossoff, Adam. ‘Saving Locke from Marx: The Labor Theory of Value in Intellectual Property Theory’. *Social Philosophy & Policy; Oxford* 29, no. 2 (July 2012): 283–317. http://dx.doi.org/10.1017/S0265052511000288.

Munzer, Stephen R. *A Theory of Property* (Cambridge University Press, 1990).

Nozick, Robert. *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

———. ‘Distributive Justice’. *Philosophy & Public Affairs* 3, no. 1 (1973): 45–126.

Organization, World Intellectual Property. *WIPO Technology Trends 2019: Artificial Intelligence*. (Geneva: World Intellectual Property Organization, 2019). https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5982426.

Palmer, Tom G. ‘Are Patents and Copyrights Morally Justified - the Philosophy of Property Rights and Ideal Objects’. *Harvard Journal of Law & Public Policy* 13, no. 3 (1990): 817–866.

Pearl, Judea, and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. 1st ed. (USA: Basic Books, Inc., 2018).

Pearlman, Russ. ‘Recognizing Artificial Intelligence (AI) as Authors and Investors under U.S. Intellectual Property Law’. *Richmond Journal of Law & Technology* 24, no. 2 (2018 2017): i–38.

Plotkin, Robert. *The Genie in the Machine: How Computer-Automated Inventing Is Revolutionizing Law and Business* (USA: Stanford University Press, 2009).

Postone, Moishe. *Time, Labor, and Social Domination: A Reinterpretation of Marx's Critical Theory* (Cambridge University Press, 1995).

Price, W. Nicholson II. 'Big Data, Patents, and the Future of Medicine'. *Cardozo Law Review* 37, no. 4 (2016 2015): 1401–1454.

Rocha, Elizabeth. 'Sophia: Exploring the Ways AI May Change Intellectual Property Protections Case Notes & Comments'. *DePaul Journal of Art, Technology and Intellectual Property Law* 28, no. 2 (2018 2017): 126–146.

Ryan, Alan. *Property and Political Theory* (Oxford: Basil Blackwell, 1985).

Samuelson, Pamela. 'Allocating Ownership Rights in Computer-Generated Works Symposium: The Future of Software Protection'. *University of Pittsburgh Law Review* 47, no. 4 (1986 1985): 1185–1228.

Sautoy, Marcus du. *The Creativity Code: Art and Innovation in the Age of AI*. *The Creativity Code* (Harvard University Press, 2019). https://www.degruyter.com/hup/view/title/565282.

Sayers, Sean. 'The Concept of Labor: Marx and His Critics'. *Science & Society* 71, no. 4 (2007): 431–454.

Searle, John. *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power* (Columbia University Press, 2006).

———. 'Minds, Brains, and Programs'. *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–457. https://doi.org/10.1017/s0140525x00005756.

Shiffrin, Seana Valentine. 'Lockean Theories of Intellectual Property'. In *New Essays in the Political Theory of Property*, edited by Stephen R. Munzer (Cambridge university press, 2001).

Singer, Peter. 'Animal Liberation'. In *Animal Rights: The Changing Debate*, edited by Robert Garner (London: Palgrave Macmillan UK, 1996), 7–18. https://doi.org/10.1007/978-1-349-25176-6_1.

Spinello, Richard A. 'The Future of Intellectual Property'. *Ethics and Information Technology* 5, no. 1 (March 2003): 1–16. https://doi.org/10.1023/A:1024976203396.

Stengel, Daniel. 'Intellectual Property in Philosophy', 2020, 32.

Sullins, John P. 'When Is a Robot a Moral Agent'. *International Review of Information Ethics* 6, no. 12 (2006): 23–30.

'The Next Rembrandt'. The Next Rembrandt. Accessed 3 August 2020. https://www.nextrembrandt.com.

Thomson, Judith Jarvis. *The Realm of Rights* (Harvard University Press, 1990).

Torrance, Steve. 'Ethics and Consciousness in Artificial Agents'. *AI & Society* 22, no. 4 (2008): 495–521.

Tully, James. *A Discourse on Property: John Locke and His Adversaries* (Cambridge University Press, 1980).

University of London Press, Ltd. v. University Tutorial Press, Ltd. (2 Ch D 601. 1916).

Waldron, Jeremy. 'From Authors to Copiers: Individual Rights and Social Values in Intellectual Property Symposium on Intellectual Property Law Theory'. *Chicago-Kent Law Review* 68, no. 2 (1993 1992): 841–888.

———. *The Right to Private Property* (Clarendon Press, 1990).

Wang, Pei. 'What Do You Mean by"AI"?' In *AGI*, 2008, 171:362–373.

Weisberg, Robert. *Creativity: Beyond the Myth of Genius* (New York, NY, USA: WH Freeman, 1993).

Weisberg, Robert W., and Lauretta M. Reeves. *Cognition: From Memory to Creativity* (John Wiley & Sons, 2013).