‘AI COLONIALISM’ IS A CONCEPTUAL METAPHOR

W. J. T. (Thomas) Mollema
6409658

Supervised by:
Dr. Uğur Aytaç
Dr. Dominik Klein

A thesis submitted in fulfillment of the requirements for the degree of MSc. Artificial Intelligence
30 ECTS

Department of Information and Computing Sciences
Graduate School of Natural Sciences
College of Science

Utrecht University
Utrecht, the Netherlands
July 2024

For Maas and Finn, my two little philosophers.

Abstract

Contemporary scholars and civil organizations increasingly associate Artificial intelligence (AI) with *colonialism*. How should these normative comparisons be interpreted? This thesis conducts a philosophical investigation into three different ways of understanding the indictments of 'AI colonialism'. First, AI systems or their development and deployment could *literally* instantiate the injustices constitutive of colonialism. Second, AI colonialism could be a *metaphor* that expresses a similarity between AI and colonialism. Third, it could be a case of *conceptual engineering*, prompting either the amelioration of the existing concept of colonialism or the expansion of the lexicon with the term 'AI colonialism'. In order to evaluate which interpretation is correct, an operational definition of 'colonialism' is pursued. The linguistically normative core of the definition established here is the 'institutional dehumanizing subversion of self-determination', while its descriptive periphery is found to consist in exploitation, the taking of land, cultural imposition and violence. Subsequently, the contours of AI colonialism are investigated. Based upon a technical understanding of machine learning algorithms, six grounds for associating AI with colonialism are found in the literature: extractivism, exploitation, objectification, cultural imposition, epistemic violence and racialization. Based on these grounds and colonialism's operational definition, it is argued that the literal interpretation is invalid because AI systems cannot literally colonize and AI development and deployment doesn't instantiate colonialism's linguistically normative core. The thesis' corresponding positive argument is threefold. First, it is argued that AI colonialism is a *conceptual metaphor* that doesn't warrant colonialism's conceptual amelioration, but is a candidate for lexical expansion. Second, it is argued that AI colonialism, as conceptual metaphor, expresses insightful interrelations between AI and colonialism's contingent properties: its projection 'stretches and twists' colonialism to disclose a meaningful coherence in AI injustices. This claim stands in between the metaphorical and conceptual engineering interpretations because it is currently a metaphor but could develop into a full neologism over time. Finally, the political metaphor that is AI colonialism is found to be epistemically successful because it is inference preserving and invites taking a theoretically fruitful political perspective.

Acknowledgements

The text facing you wouldn’t have been possible without the collaboration and support of other people. First, I would like to thank Arthur Gwagwa and prof. dr. Joel Anderson, who recognized the project’s potential and were generous with their thoughts during a brainstorm in the project proposal phase. Arthur also read one of the final versions and supplied me with feedback, which was very helpful. Secondly, my thanks go out to dr. Dominik Klein. Dominik’s sharp comments and supportive perspective on the project proved valuable to me; I want to thank him for his time. Thirdly, I am greatly indebted to my supervisor dr. Uğur Aytaç. Uğur supplied me with many research (re)directions, thorough *and* constructive criticism, and thoughts of his own, so that the project could never have become what it is without him. Uğur’s uplifting smile not only made the supervision meetings fun, but also communicated that I was on the right track.

Finally, I would like to express my gratitude to my wife Aranea. She supports me through thick and thin and has shown to possess an enormous amount of patience with me throughout the past three years of my enrollment in the AI master’s program. I thank her for her emotional and theoretical support and I apologize for always testing my philosophical ideas out on her first.

W. J. T. (Thomas) Mollema, Berkel en Rodenrijs, July 2024.

Table of contents

1
Introduction

Artificial intelligence's (AI) impact on the digital economy is significant and harms and injustices associated with these socio-technical systems and their development and deployment have emerged. Theorists hailing from media, science and technology studies, sociology, ethics and political philosophy have provided problematizations of AI. In the past five years, a particularly interesting charge has emerged: the association of AI with *colonialism*. Especially the decolonial movement in technology studies has compared AI systems and their development and deployment with colonialism (Ricaurte, 2022; Couldry & Mejias, 2019; Mbembe, 2022). Colonialism is a complex historical injustice, involving domination, exploitation, cultural imposition, the taking of land and the invention of race. Therefore, any comparison of a present-day phenomenon with colonialism is a grave indictment.

The grounds for the comparisons between AI and colonialism are as varied as colonialism's aforementioned harms. To hint at the variety of comparisons, AI is associated with 'the coloniality of power', 'digital and data colonialism' or 'algorithmic colonialism' because of…

i. The nature of data extraction by and for AI (Eke, Wakunuma & Akintoye, 2023, 48; Muldoon & Wu, 2023, 2; Lambrechts, Sinha & Mosoetsa, 2022);
ii. The extraction of natural resources by the supply chain for AI systems and the dependency of the AI economy on power imbalances mirroring those of historical colonialism (Muldoon & Wu, 2023, 2; Joler, 2020; Madianou, 2019);
iii. The refraction of "forms of coloniality and the historical construction of race" (Adams, 2021, 2) by AI systems (McQuillan, 2022);
iv. The hegemonic reproduction of Western knowledge and standards of justification by AI (Katz, 2020; Muldoon & Wu, 2023);
v. The usage of AI makes possible forms of algorithmic governmentality (Ricaurte, 2022, 732);
vi. The dependence of the AI economy on invisible and outsourced labor (Mohamed, Png & Isaac, 2020; Crawford, 2021);
vii. The exploitation of the poorest communities in the global South by individuals and corporation from the global North via AI-based markets (Birhane, 2020); and
viii. The imposition of foreign values by means of AI systems (Eke, Wakunuma & Akintoye, 2023, 132; Mhlambi & Tiribelli, 2023; Ugar, 2023).

The list could go on.

In this thesis, I ask how 'AI' and 'colonialism' should be understood together and investigate whether it makes sense to associate these concepts. Doing so is valuable, because it dispels confusions and misconceptions underpinning this adjecting of AI to colonialism. To be concrete: is there really colonialism involved with AI systems or their usage in the same sense as, e.g., the colonialism the Dutch practiced in the Indies? And if so, is the wrong committed in both cases *the same*? In short, I inquire into, firstly, *the properties that colonialism and AI injustices might share* and, secondly, their *possible convergence in terms of what makes them wrong.* Positive results in both categories are valuable because the first will help explain why people compare AI systems/usage to colonialism, while the second yields an important normative result in that the contemporary injustices entangled with AI would require the same normative treatment as colonialism:.

As Hanna Pitkin (1972, 115) wrote: "what we see and what is there for us to see will depend on the concepts we bring to our experience". We should thus take care of our conceptual arsenal and prohibit nonsensical additions while welcoming useful newcomers, as this has direct consequences for interactions with social events, their classification and moral evaluation. If colonialism is used to describe contemporary AI injustices, it shapes our perception of and responses to them. The use of the word *suggests* the presence of certain wrongs to us: colonialism's conceptual shape comes to precede the apprehension of the injustices related to AI. The *applicability* of colonialism to these injustices should be carefully evaluated. The first problem is comparing AI and colonialism. To do so, we need to know what colonialism is and what makes it wrong. And, we need to understand the grounds for comparing AI with colonialism. The second problem to solve is: can a *legitimate conceptual case* be made for 'AI colonialism', or is the concept superfluous – merely a seditious metaphor – because the injustices associated with AI are better conceptualized in terms of other normative concepts, like domination and exploitation?

This thesis' main question thus is: *how should we understand 'colonialism' in AI colonialism?* Does it point to colonialism literally being the case? Does AI colonialism articulate social reality in a distinctive way that sheds new light on contemporary injustices? Or, is it superfluous, contributing nothing to the existing normative vocabulary, or conceptually confused, leading to obfuscations of our understanding of social reality? My running hypothesis is to construe AI colonialism as social "thinking tool" (Dennett, 2017): a *metaphor* that provides an analogical model for political theories. In line with Black's (1977, 445) saying that "every metaphor is the tip of a submerged model", I will develop an account of AI colonialism as a metaphor that connects two conceptual domains and that has irreducible expressive and epistemic value.

But whether literally or metaphorically employed, what are concepts anyway? As Pitkin (1972, 70) writes, concepts are "compounds, assembled out of the variety of cases in which they are characteristically used". Using a concept thus is to employ a more or less fuzzy partitioning of reality or clustering of similar cases. Taking cues from Ludwig Wittgenstein's later logico-grammatical analyses, it can be seen that words signify concepts while playing certain roles. A language-game is language-use (Wittgenstein's *Sprachgebrauch*) that suffices for some purpose, which is the cause of why it is played. The words that figure in the language-game signify concepts that have a *point*, a pragmatic application (Wittgenstein, 1975, §474). The use of words points to their meaning, by showing language-users certain conceptual interrelations. So reasons (distinct from the cause of why the game is played) can be given for why concepts are used as they are. I will delimit the concepts of colonialism and AI colonialism to clarify their points and their utility for addressing our needs. Correspondingly, AI colonialism, on analysis, can prove vindicated or "may also turn out to be *pointless* or even *dysfunctional* for us" (Queloz & Cueni, 2021, 21). But for Pitkin (1972, 61), the boundaries of concepts are fundamentally undrawable. To fruitfully investigate the point of AI colonialism, we should recognize that colonialism's conceptual contours unite similar cases, while leaving room for their dissimilarities. But boundaries can be drawn up, at the cost of an awareness of their approximative nature: a delimitation of a concept yields an 'operational definition'. Doing so for colonialism enables the analysis of AI and colonialism's interrelation. Using the operational definition, its relations to the particular contours of AI colonialism can be evaluated to see whether colonialism literally or metaphorically applies and whether it discloses a political perspective in an epistemically successful way.

Against this background and in response to the foregoing questions, the thesis will develop and defend four diagnostic arguments. First, that for AI systems and AI development and deployment, the association with colonialism can be made *metaphorically,* but never literally. Second, I argue that AI colonialism as

description of AI systems or their development and deployment doesn't warrant the amelioration of the existing concept of colonialism.

Third, I defend the view that AI colonialism qua metaphor retains *explanatory* value because it is a *conceptual* metaphor. The function of a conceptual metaphor is to better understand certain concepts by interrelating them, and not just for some artistic or aesthetic purpose (Lakoff & Johnson, 2003). A conceptual metaphor is a set of correspondences between a source domain and a target domain (Lakoff, 1993, 4-5), where 'domain' amounts to a "coherent organization of human experience" (Kövecses, 2016, 12).[1] AI colonialism is the metaphorical mapping of properties of the source domain of colonialism onto the target domain of AI.

Fourth, AI colonialism, despite not being literally applicable, isn't descriptively superfluous. Rather, I argue that it is an epistemically successful politico-conceptual metaphor that expresses meaningful interrelations between AI injustices and colonialism that are valuable for understanding the former: it "selects" or "filters" aspects from AI (Black, 1955), disclosing its "intelligible structure" (Miller, 1979).

1.1. Machine learning

As a preliminary, some conceptual distinctions are introduced: for starters a specification of the AI systems the thesis is concerned with; afterwards, some methodological distinctions that enable the subsequent conceptual analyses.

What is AI? The definition of AI is widely debated: it is a capacious term that contains the normatively thick concept 'intelligence', which has eluded philosophers ever since Aristotle. However, the Dutch Organization for Scientific Research (NWO)'s (2019) helpfully describes it: "Today's AI is the science and engineering of making machines intelligent and collaborative". Different disciplines within AI aim to realize these goals from different angles, such as machine learning, knowledge representation and reasoning and autonomous agency. I tailor my use of the term to how the authors I discuss use it: the focus is on *machine learning* (ML). Due to its revival and success in the 21st century, ML is the subject of the (decolonial) technology scholars' discourse. I let 'AI systems' refer to algorithmic ML-systems and let 'AI development and deployment (D&D)' denote the broader sphere of designing, using and manufacturing AI systems.

ML can be defined as algorithms that self-improve based on a performance measure. The way towards the goal of the task that performance is measured on isn't programmed into the algorithm itself, since the weights of the algorithm's network have to be learned (Watson, 2023; El Naga & Murphy, 2015). The learning is supervised or unsupervised (more on this below). The relevant type of ML for the following chapters is *deep learning*, which concerns the usage of multilayered neural networks as a flexible model to learn patterns in huge datasets, whether supervised or unsupervised. Central to ML is that instances (data points that are part of a dataset) are repeatedly inputted into the algorithm, after which the algorithm iteratively improves on itself to reproduce the desired output with respect to the inputted instances. This computational self-improvement is termed 'learning' and the processing of a single input is called an 'experience' or 'observation'. In the algorithm's training phase, learning proceeds through an adjustment mechanism of all points where the algorithm performs a computation so that the results of subsequent computations better match the desired output. This iterative improvement is based on the quantified distance between the

[1] In contemporary conceptual metaphor theory, stemming from the work of Lakoff & Johnson (2003), the definition of a conceptual metaphor is as follows: "*A conceptual metaphor is understanding one domain of experience (that is typically abstract) in terms of another (that is typically concrete)*" (Kövecses, 2016, 1; italics in original).

desired output and the actually computed output. In feedforward neural network architectures this is called 'backpropagation'. Through learning, the model's accuracy increases. Although it could always 'learn more', the algorithm's structure, parameters and hyperparameters cap the target accuracy. After training comes the prediction phase: the model is deployed on novel examples and uses its finetuned and learned weights to process these examples.

In supervised learning the "input and output are predefined": the training data contains instances with a category labelling that the algorithm should output for those instances (Sharma, Sharma & Khanna, 2020).[2] Supervised learning involves classification or regression. In classification, the categorical labeling determines the ML-model's output space and learning enables the mapping of novel inputs accurately on one of the predetermined categories. Regression engages in 'curve fitting': the optimal function is sought for capturing the relationship between a dependent variable and an independent variable in the training data (Sharma, Sharma & Khanna, 2020, 589; El Naga & Murphy, 2015, 5-8). Unsupervised learning involves unlabeled data. The algorithm is tasked with finding an optimal partition of inputs into categories: discerning the data's probabilistic structure (Watson, 2023, 3). According to Watson, there are three general forms of unsupervised learning, of which supervised or self-/semi-supervised versions are equally possible. A *clustering algorithm* carves a dataset into similar groups. At different levels of granularity, based on a sequence of observations, unsupervised clustering statistically approximates the nearest centre relative to cluster variance. It determines statistical regularities shared by observed examples at the level of partition. An *abstraction algorithm* learns simplified representations of the training data and outputs a more abstract pattern that embeds the general shared properties of the dataset.[3] Lastly, a *generation algorithm* creates synthetic text, pictures or sounds. These algorithms construct statistical models through learning. By having an input 'pass through' the model, it is transformed into the desired output (Watson, 2023, 4; 9-10; 13; 16).[4]

The underpinnings of ML give us a general understanding of the computational processes underlying 'large language models' (LLMs). Recently, LLMs such as OpenAI's ChatGPT (Generative Pre-trained transformer), Google's Gemini and Anthropic's Claude received much attention. They are used for many purposes, including generation and classification and were built using both supervised and unsupervised learning. An LLM is pretrained using self-supervised learning (*unsupervised* learning that supplements itself with

[2] The following characterizations of supervised and unsupervised learning are adapted from my unpublished manuscript on the Watson-Sterkenburg debate on clustering algorithms and natural kinds (see Mollema, 2024b).

[3] A neural network autoencoder is a paradigmatic example. The architecture is comprised of two components: an encoder model and a decoder model. Underlying this architecture is the idea of compressing an input into a vector by the encoder model. A vector is an array of (real) numbers that comes to represent the inputted instance via quantitative encoding. The embedding into vectors is the 'latent informational space' of the data that represent the patterns in them. These vectors are subsequently fed into a *decoder* model. The decoder translates the elements of the latent informational space into an output that has been affected by the data's general properties as captured in the compressed data (Roy, 2020).

[4] A generative adversarial network (GAN) is a prominent example of a generative algorithm. This is a two-part neural network: a 'generator' that produces possible outputs and a 'discriminator' that judges that output. The generator takes a random statistical distribution and produce a 'fake'; in turn, the discriminator pairs this with training dataset examples and has to choose which is 'real'. Learning proceeds via the interaction between generator and discriminator with as goal that the generator becomes good enough to fool the discriminator (Brownlee, 2019).

contextual information)[5] and is finetuned using reinforcement learning through human feedback (*supervised* learning with data labeled by humans). Purportedly, the usage and development of ML and LLMs are what gives rise to AI colonialism.

1.2. Methodological distinctions

Some methodological clarifications are in order regarding how to analyze AI colonialism and colonialism and how to determine a concept or metaphor's explanatory value.

1.2.1. The literal, metaphorical and conceptual engineering interpretations

First, we should distinguish between the possible interpretations of AI colonialism, in addition to that it can concern AI systems, D&D or both. There are three candidate interpretations:

i. The *literal* interpretation: AI literally instantiates the injustice(s) characteristic of colonialism.
ii. The *metaphorical* interpretation: AI instantiates injustice(s) that can be metaphorically likened to colonialism.[6]
iii. The *conceptual engineering* interpretations: (a) *lexical expansion:* a new concept for speaking of AI injustices is introduced that closely resembles the concept of colonialism (Cappelen, 2020) or (b) usage of the term 'AI colonialism' is a sign that the concept of colonialism stands in need of *amelioration* (Haslanger, 2020).

It is important to distinguish between the three interpretations, because if (i) holds, this will have different consequences for the conceptualization of appropriate forms of redress than if only (ii) holds. It should also be noted that (i) and (ii) are mutually exclusive, and so are conceptual engineering interpretations (iii.a) and (iii.b). (i) and (ii) are both compatible with (iii.a) and (iii.b), but (i) and (ii) can hold regardless of them. The objective of the thesis is to choose the right set of interpretations.

I develop the argument that the literal interpretation is too strong, because all AI injustices dubbed 'colonial' don't instantiate colonialism's linguistically normative core. For this argument to succeed, the AI systems as well the AI D&D delineations are discussed. Subsequently, I defend the metaphorical

[5] The 'self-supervision' all of these models make use of is realized by an architecture called a 'transformer'. The transformer was popularized by Google in 2017 (Vaswani et al., 2017). The transformer architecture is an encoder-decoder that makes use of 'attention blocks' to improve the translational process of encoding input data, processing it through the model's informational space, and then decoding it for outputs. The idea behind encoder-decoder models is that you get an input of length *n* and you want an input of length *m* as an output. As such, the input is fed into an encoder, which maps and compresses to a context embedding of a certain number of dimensions. The decoder on the other hand, takes the context and translates it into a sequence of length *m*. To encode is then to take a word and transform it into a context vector and to decode becomes to take a context vector and to translate it into a word (Alammar, 2022). The attention mechanism is the introduction of a *context vector* (representing the input) to the encoder-decoder. All hidden states of the encoder are squeezed into the context vector for a supplementation of the input information to each step in the decoding process. This architecture, if scaled, has proved extremely successful for processing all forms of input data that can be successfully tokenized.

[6] The distinction between a literal and a metaphorical interpretation of a sentence or phrase stems from the theorizing of metaphor from Aristotle's *Poetics* and onwards (Johnson, 1981, 7). As Lakoff (1993, 203) writes, on the traditional theory of metaphor, "by definition, […] What is literal is not metaphorical." More generally, what is literal is thus not intended as a trope (metaphor, metonymy, hyperbole, etc.).

interpretation because AI colonialism is a conceptual metaphor that structurally connects the domains of AI and colonialism. Subsequently, I show that (iii.b) must be dismissed, because AI colonialism isn't used to replace colonialism or to ameliorate the shortcomings of the usage of the term colonialism. Rather, I show that (iii.a) is theoretically enabled by the metaphorical interpretation's vindication. The conceptual metaphor calling attention to the resemblance between AI systems and their usage to colonialism introduces a "new meaning" which makes it a case of *lexical expansion* (Cappelen, 2020, 142). However, its metaphorical nature doesn't detract from AI colonialism's value for normative theories: qua it being a *conceptual* metaphor – a "mapping across conceptual domains" (Lakoff 1993, 203) that preserves an analogy between colonialism and AI injustices – it fulfills epistemic functions. A theory gains in explanatory value by incorporating AI colonialism's expressive power.

1.2.2. Expressive power, explanatory value and metaphor

The foregoing introduced multiple technical terms that require unpacking. The *expressive power* of a literal term should be understood as the disclosure of some aspect of reality that cannot be explicitly grasped without it. Metaphors have expressive power too. Conceptual metaphor theory (CM-theory) regards a conceptual metaphor as the systematic communication of some form of structural similarity (Kövecses, 2016, 6; Lakoff, 1993). Building upon CM-theory, Nikola Kompa (2021) argues metaphors are epistemically valuable in their disclosure of aspects of a target domain through the projection of properties from a source domain. If it is a conceptual metaphor, AI colonialism's expressive power is beyond merely aesthetic language-use because it articulates *conceptual relations*.

Here expressive power's interrelation with *explanatory value* becomes relevant. Explanatory power pertains to theories. In theories, concepts do normative or descriptive work and together give rise to a theory's power to explain its target phenomena. Theories' precision, cognitive salience, non-sensitivity to counterfactuals, ease of integration with existing knowledge and factual accuracy are common determinants of explanatory power (Ylikowski & Kuorikoski, 2010). Theory-constituent concepts have a certain indispensability with respect to this explanatory power, i.e., without their *expressive* power, the theory would lack in or have less explanatory power. Borrowing connotations of 'expressive power' from logic and computer science, where it means the *breadth of representations* a formal language can communicate, we can say that a concept's expressive power constrains a theory's explanatory power.

For metaphors in particular, the explanatory power is analogical: "only certain relations and inferences allow themselves to be mapped onto each other" (Kompa, 2021, 42). The distinction between metaphor and analogy is that analogy is a reasoning process and metaphor is a figure of speech, but "the meaning-giving aspects" of analogical reasoning underly metaphors (Nercessian, 2015, 134). This role of metaphor and analogy extends to political theory as well. As Mary Hesse (1988, 7) argues, "novel theoretical concepts come to supplement ordinary descriptive language by the use of scientific models and metaphors". So a theory's explanatory power is determined by metaphorical or literal terms' indispensability for understanding its explananda (target phenomena). Therefore the expressive/explanatory power framing can be applied to the previously identified interpretations of AI colonialism. As conceptual metaphor, it would show language-users similarities between two separate domains: a "metaphorical naming" that "makes political things visible" (Rosenthal, 1982, 294). For some, such as Hesse (1988), Gill (1979) and the CM-theorists (Lakoff & Johnson, 1981; Kövecses, 2016), language-use is fundamentally metaphorical because of this 'making visible'. On these views, metaphors underpin and make possible the capturing of reality by language and

analogical thought. In any case, AI colonialism's expressiveness is its articulation of conceptual relations that one couldn't do without it *or* its correct identification of the injustice of colonialism.

Part of my hypothesis is that AI colonialism discloses relations between AI injustices that escape our attention without it: it hands us a political change in perspective. For example, as will be seen in §2, in colonialism racialization, epistemic violence and cultural imposition are interrelated, and, as will be shown in §§3-5, the mirroring of this interrelation in the domain of AI is *made visible* by naming it as 'AI colonialism'. This works analogously for the relations between other injustices related to colonialism, such as extractivism and exploitation, and such successful projections of aspects from colonialism onto AI translate into explanatory value once embedded into a theory of (what is wrong with) the social reality of AI.

The epistemic value of the literal interpretation is clear: if it obtains, *there is colonialism in AI*. The epistemic value of the metaphorical interpretation is more slippery. Analogies and metaphors can be epistemically valuable as well as harmful beyond natural and political theories' basic dependency on analogical and metaphorical postulates (e.g., atoms, genes, computers, networks, organisms), as "analogical inference" lies "at the center of problem solving where little is known about the target phenomena and conceptual innovation and change are on-going processes" (Nercessian, 2015, 144). So not all metaphors are epistemically successful, but good metaphors, based on good analogies, are theoretically indispensable. Metaphors (analogy carriers) constitute the inferential motor of theory development and *good* analogies preserve relational structures, make isomorphic mappings and are systematic with respect to the connected target and source domains. Kompa (2021, 43) provides similar criteria for the epistemic success of metaphors: (i) the existence of postulated entities; (ii) domain-to-domain correspondence in patterns of inference; (iii) disanalogies for the metaphor; (iv) the theoretical fruitfulness of what the metaphor *shows*; (v) the obscuring of other important aspects of the target domain; and (vi) the metaphor's fertility for new metaphors.

The challenge for the metaphorical interpretation is to display AI colonialism's epistemic success of AI colonialism. Through discussing the previous criteria, in §5 I will show AI colonialism is epistemically successful because it is *inference preserving* and *theoretically fruitful*: it can be a target and postulate for political theory with pragmatic, unifying and generative benefits. The clustering it prescribes isn't accidental or arbitrary, because a redescription of AI in terms of other normative concepts loses explanatory value; relations between AI injustices and colonialism that AI colonialism makes visible are lost. Therefore AI colonialism is best regarded as conceptual metaphor with epistemic value that articulates meaningful relations between AI injustices and invites political perspectival change.

1.2.3. Analyzing colonialism

Finally, for analyzing colonialism, I introduce the following axis for partitioning it.

i. Colonialism's *constitutive* wrongs. A constitutive wrong is the injustice that determines what counts as colonialism. It is noncontingent for colonialism, i.e., invariably part of it. If colonialism is the case, then, necessarily, such a wrong is committed (more on this in §2). These wrongs are part of colonialism's 'conceptual core'.
ii. Colonialism's wrongful *effects*. Wrongful effects are the contingent wrongs made possible by the colonial frame constituted by (i). These effects are historically strongly associated with colonialism, but aren't necessary conditions for colonialism; they belong to colonialism's 'conceptual periphery' (Van Wietmarschen, 2018, 171).

In applying this distinction, finding colonialism's 'essence' isn't pursued. Instead, it has pragmatic use for the current project. Regarding wrongful effects as *symptoms* of colonialism (proxies, minor similarities) and constitutive wrongs as *criteria* of colonialism (distinctive similarities) will yield not strictly speaking a universal definition of colonialism, but rather a justified approximation of a definition: a demarcation of its contours (Pitkin, 1972, 127). This distinction plays two roles in the subsequent chapters.

First, it yields a conceptual scheme for understanding and defining colonialism. Operationally defining colonialism rests upon finding an adequate set of constitutive wrongs. A wrong is distinctive for an injustice if, when it is identified, the object of the identification cannot be redescribed in terms of another wrong. We can therefore approach the definition of colonialism by classifying its main features as being either a constitutive wrong or a wrongful effect. Applied to (AI) colonialism, if an event *E* can be described using the term (AI) colonialism, but it can perfectly be redescribed in terms of other concepts, then (AI) colonialism is explanatorily superfluous, because it has no distinct explanatory value.

Secondly, the distinction helps order the *grounds of comparison* between colonialism and AI injustices. Applying the distinction there shows what components of colonialism are actually instantiated in the AI injustice in question. On the one hand, if only some wrongful effects are instantiated, but no constitutive wrongs, then the literal interpretation will be false, because only contingent properties of colonialism are instantiated, and none of its conceptual core. On the other, the metaphorical interpretation gains traction as soon as an AI injustice shares informative, albeit contingent, properties with colonialism. But when the constitutive wrongs are identified, then the literal interpretation of AI colonialism is warranted. Furthermore, Tashmia Sabera (2021) aptly distinguishes between colonialism's distinctive wrong and its definition. While the former is essentially normative, the latter can contain descriptive as well as normative elements. Any attempt to show that talk of AI colonialism should be taken metaphorically rather than literally additionally has to deal with the possibility that AI may not resemble colonialism's normative definitional elements, but *does* resemble its descriptive definitional elements (or vice versa). If there is only descriptive convergence, that is essentially the weaker result, because in lack of normative overlap, 'AI colonialism' cannot be used as a form of moral/political critique equivalent to the condemnation of historical colonialism.

That being said, the distinction between the description and rule for description (norm) is never rigid (Pitkin, 1972, 228). Within different language-games, "the same proposition may get treated at one time as something to test by experience, at another as a rule of testing" (Wittgenstein, 1975, §98). This is to say that words and their connotations can vary: they can play descriptive as well as regulatory conceptual roles in different usages. Call this the distinction between *linguistic normativity* and *linguistic descriptivity*. Applied to the present project, in the final analysis it shouldn't be the distinction between linguistically normative or descriptive usages of colonialism that forbids the use of the word *like that* altogether; at the same time, the transition of the word colonialism from description to linguistic norm should be regarded with suspicion. Rather, one should be open to the possibility that the term colonialism is used *now as a description* ('describing *X* using colonialism') and at other times *as a rule for judging* ('explaining why *Y* is wrong by identifying it as colonialism'). On the side of linguistic normativity, the concept functions as norm in a way that the semantic content of the concept gives us a yardstick for measuring the cases, while the descriptive usage of the concept is comparative and deals with empirical patterns and where a certain use fits in those patterns.

Furthermore, colonialism's wrongs can simultaneously be *semiotic* wrongs. Brennan & Jaworski (2015, 1055) describe semiotic objections as "a form of symbolic expression that communicates the wrong motive, or the wrong attitude toward *X*, or expresses an attitude that is incompatible with the intrinsic dignity of *X*, or would show disrespect or irreverence for some practice, custom, belief, or relationship with which *X* is

associated." A semiotic wrong is thus a minimalistic type of wrong. For colonialism, one of its necessary or contingent features could be wrong for what it *communicates*. As §2 will find, the type of domination constitutive for colonialism is semiotic: it communicates a deep disrespect towards the colonized. The metaphorical interpretation of AI colonialism could also be vindicated if AI injustices deeply disrespected those subjected to it akin to how historical colonialism did so. This semiotic wrong would then function as one of the structural components of the conceptual metaphor.

1.3. Structure

In §2 an operational definition of colonialism is constructed based upon the relevant politico-philosophical literature. Normative debates on the concepts domination, exploitation and colonialism are presented to provide a vocabulary tailored to the objectives of the thesis. §3 reconstructs six types of 'AI colonialism': exploitation, extractivism, objectification, cultural imposition, epistemic violence and racialization. The chapter contributes to understanding how and why AI systems and AI D&D are charged with colonialism. §4 (based upon §2's normative vocabulary and §3's investigation of AI's colonial aspects) argues that AI colonialism, applied to AI systems and AI D&D, shouldn't be interpreted literally. Finally, in §5, I argue that AI colonialism is a conceptual metaphor that expresses contingent conceptual connections between AI injustices and historical colonialism. It is also shown how it warrants conceptually engineered lexical expansion, but forecloses conceptual amelioration of the traditional conception of colonialism. Finally, it is argued that AI colonialism is an epistemically successful politico-conceptual metaphor. Chapter 6 concludes.

2

An operational definition of colonialism

> **colonialism** *n.* Policy of maintaining colonies (esp., derog.) alleged policy of exploitation of backward or weak peoples, […].
>
> *The Concise Oxford Dictionary*

> I hope I may be forgiven for asking that those who take it on themselves to describe colonialism remember one thing: that it is utopian to try to ascertain in what ways one kind of inhuman behavior differs from another kind of inhuman behavior.
>
> Frantz Fanon, *Black Skin, White Masks*

Western potentates, the traditional colonizers, colonized various peoples in South-America, Asia, and Africa and this involved the appropriation of land and natural resources, the subjugation and enslavement of native peoples and violent forms of alien rule (Getachew, 2019). Colonialism originated from the 16th century's migration dynamics and led to the subjugation, enslavement and eradication of, among others, South American civilizations by the Spanish and Portuguese, African civilizations by the French, British and Dutch, India by the British, Indonesia by the Dutch (Mignolo, 2020). The tales of these atrocities are worth recounting and remembering, and many of today's theorists agree that Euro-American prosperity – 'modernity itself' – was built on the suffering of the colonized (Mbembe, 2017; Mignolo, 2020). As Charles Mills (1997, 35) writes: "colonialism 'lies at the heart' of the rise of Europe". The *Oxford Dictionary*'s commonsensical definition is informed by this legacy. While dictionary definitions are philosophically unfashionable, they provide functional introductions to *standard* uses of a term. For colonialism the definition communicates two uses, typified as 'policies': *settler* colonialism and the *exploitation* of one community by another. This chapter's finale will make clear the shortcomings of this definition.

The political conceptualization of colonialism is contested. Contemporary political philosophers agree that *many* wrongs were involved with colonialism, such as murder, cultural imposition, racialization, (sexual) violence, exploitation, and the taking of land, and that it involved domination. However they disagree on this domination's *form* and what wrong (or configuration of wrongs) is *distinctive* for colonialism. In the following, I venture to demarcate colonialism's linguistically normative core in order to come to understand if the literal interpretation of AI colonialism makes sense: what is *unjust* about colonialism should be mirrored in AI systems or D&D for this interpretation to be justified.

Frantz Fanon's epigraph is at odds with that endeavor. Fanon's (1963; 2008) work provided invaluable insights into colonialism's psychological and racial dynamics, but is explaining colonialism's wrongs really a utopian project? On the contrary, the assumption I depart from is that colonialism not only serves historical descriptions of injustices (of which Fanon is speaking), but also articulates a noncontingency applicable to contemporary or future events. If this assumption is justified, then colonialism must also be distinguishable from other concepts – *but on what grounds?* Fanon would regard this analysis an indignity towards the colonized, but I will show precisely the opposite is the case, because instances of colonialism are distinguished by a procedural and simultaneously semiotic wrong and this conception of colonialism can help us understand contemporary injustices.

Therefore, colonialism's conceptual orrery requires elucidation. What concepts are part of it?

The main contenders are exploitation and domination. Other candidates are cultural imposition, violence and land-taking. Firstly, they are to be classified as (i) wrongs constitutive for colonialism or as (ii) wrongful effects. Here (ii) can be causally related to (i) in that they historically 'travelled together', but (ii) remains indistinctive of colonialism. For example, the (sexual) violence in colonized societies, which is an important wrong, also pertains to noncolonial societies. Secondly, the distinctions between (i) and (ii) and that between semiotic/non-semiotic are combinable. Part of the working hypothesis is that there is a *semiotic* dimension to colonialism's constitutive wrong(s): the mere act of colonizing *symbolically communicates* a disregard for the colonized's dignity.

My goal is to demarcate a *mixed structure* of the concept of colonialism as operational definition: a conceptual delineation containing one necessary condition and a multitude of non-essential family resemblance attributes (Barrenechea & Castillo, 2019). A sketch of the conclusion is as follows.

(1) The necessary condition for colonialism (its linguistically normative core) is: *the institutional subversion of the self-determination of the colonized by the colonizer*. This domination-type has a semiotic dimension: *dehumanization*. The procedural and semiotic are co-dependent, because the dehumanization is communicated by the particular objectifying curbing of political communities' autonomy constituted by colonization.

(2) Colonialism's contingent family resemblance attributes (extensions of its descriptive use) are found to be: exploitation, violence, land-taking, cultural imposition and racialization.

Colonialism is procedurally and morally wrong, regardless of the wrongful effects completing its definition.

Moving towards this conclusion, domination and exploitation are discussed as candidate 'building blocks' for colonialism. These concepts help provisionally classifying AI injustices (§3), before we analyze whether they constitute colonialism literally understood. Correspondingly, §2.1. presents the structuralist and republican accounts of domination. I explain how the structuralist notions of oppression and domination can be subsumed under republican domination and the resulting conception of domination I subscribe to is *an agent's arbitrary interference with the choice set(s) of other agent(s)*. This operational definition explicates that the structural obstruction of self-determination constrains the dominated party's freedom and is helpful for pinpointing colonialism's linguistically normative core. §2.2. defines exploitation by contrasting a domination-based account with an analytical Marxist account. I argue exploitation *modifies* domination, which (1) enables that cases of exploitation are seen as interferences with someone's choice set; and (2) makes domination primitive to exploitation, which explains its role in colonialism. §2.3. summarizes the debate about the necessary conditions for colonialism so that in §2.4. I can propose colonialism's mixed structure.

2.1. What does domination mean?

Understanding domination precedes understanding colonialism. Domination is also useful for identifying the wrongness of AI injustices: as instances of domination, they can be noncolonial, or forms of (AI) colonialism. On the structuralist account (§2.1.1.), domination is entangled with what *injustice* is. Conversely, republican domination (§2.1.2.) is tied to *freedom*. Based on both, I pragmatically choose to conceive of domination as the arbitrary interference of one party with the choice sets of another party (§2.1.3.). This definition retains both views' explanatory powers: the structuralist emphasis on the relational self-determination of social groups and the republican pairing of non-domination with freedom. Subsequent chapters benefit from this generally accepted and broadly applicable conception of domination.

2.1.1. The structuralist account

Structuralists distinguish domination from oppression. Oppression is concerned with the *inhibition of the self-development* of social groups, while domination is concerned with *obstructing their self-determination*.[7] Young (1990, 41-42) explains oppression involves "systematic institutional processes" disabling people to learn or exhibit skills or "institutionalized social processes which inhibit people's ability" to participate in the public sphere. The 'inhibition of self-development' concerns the restraining of individuals' forms of expressing and doing and giving shape to their social lives. Racism, sexism and ageism are forms of oppression on Young's account. Conversely, domination "consists in institutional conditions which inhibit or prevent people from participating in determining their actions or the conditions of their actions". The 'obstruction of self-determination' therefore means that individuals' *conditions of action* are determined by another party without reciprocation. What is structural about the concepts of oppression and domination is the *systematicity* with which they affect individuals' experiences and self-determination. The structuralist distinguishes oppression from domination because, although "oppression usually includes or entails domination", constraints on self-development that are oppressive don't always arise out of dominating relations (Young, 1990, 38). According to Young, oppression and domination overlap, but there are cases where people are dominated but not oppressed, such as bureaucratic domination and domination in the workplace, and cases where people are oppressed but not dominated, like powerlessness and marginalization. The social ontology underlying this regards the social group as ontologically prior to individuals in the sense that oppression and domination are conceived of as structural relations *between* differing *groups* (Young, 1990, 45).

For the structuralist, a sufficient definition of injustice is the *presence* of relations of oppression and/or domination between social groups (Young, 1990, 37). Young conceives of 'justice' as the absence of domination and oppression. This definition's merit is the inclusion of fairness with respect to distributions, while not being restricted to it, because it also includes fairness in terms of *recognition* (tied to social groups' expressing and doing). To recapitulate, oppression is the inhibition of "developing and exercising one's capacities and expressing one's experience", while domination is the obstruction of "participating in determining one's action and the conditions of one's action" (Young, 1990, 38).[8]

2.1.2. The republican account

Having comprehended structuralist oppression and domination, republican domination can be discussed.

For the republican account, what counts as a dominating relationship is determined by freedom's function as political ideal. For Philip Pettit (1999, 109), freedom is defined in terms of *non-domination*.[9] According to Annelien de Dijn (2020, 347), the republican notion of freedom belongs to the conceptual tradition of "freedom [that] required democratic self-government", rooted in ancient Greece and Rome. This makes sense, as for the republican the protection of individuals against the dominating exercise of power is a

[7] I treat Iris Marion Young's views in *Justice and the Politics of Difference* as paradigmatic for the structuralist account, because of their influence.

[8] Here I paraphrase the values of justice that lie at the core of Young's view of what oppression and domination are. At hear of oppression lies the following value of justice: "(1) developing and exercising one's capacities and expressing one's experience […]". The value of justice that opposes domination in the structuralist account is "(2) participating in determining one's action and the conditions of one's action […]" (Young, 1990, 38).

[9] The work of Pettit lies at the roots of many republican accounts of domination, and hence I treat his work as paradigmatic for them.

guiding ideal institutional and political legislation. Pettit (1999, 271; 274) presents republican freedom as a 'third-way' alternative for the dilemma of giving primacy to either positive or negative forms of freedom.[10] Being free is the absence of domination, which is "the capacity to interfere on an arbitrary basis in [other's] choices" (Pettit, 1999, 272). In other words, one party dominates another if the former is able to determine the latter's options for action at will, without reciprocation. Domination, the impairment of freedom, varies in intensity and extent, scaling with the power of one agent's capacity to impose its will on another.

Building upon the republican conception, scholars have sought to characterize *digital* forms of domination. The following discusses two of them because of their relevance to AI injustices.

Bernd Hoeksema (2023) reviews domination through surveillance, data collection and algorithmic decision-making. For surveillance and data collection, Hoeksema argues that they institute "the capacity" to interfere in the form of exercising power over agent's choices regarding the taking, hiding or sharing of (personal) information: "our control over who has access to our information and what it is used for […] is limited", e.g., resulting in breaches of privacy. Algorithmic decision-making satisfies the condition of arbitrary interference via its determination of the preselection or disqualification of choices that are digitally presented to individuals, shaping their choice-situations (Hoeksema, 2023, 8).

Uğur Aytaç (2024) analyses domination on social media platforms, whose role in daily life, facilitating personal contacts and fora for public spheres, has become irreplaceable. On social media platforms, domination is 'digital' because it takes place in or through digital facilities for intra-citizen communication, i.e., in the processes constituting digital extensions of the public sphere. For Aytaç, "[s]ocial media companies dominate individuals qua citizens because of their arbitrary power to interfere with citizens' political participation in the public sphere". Specifically, social media companies narrow the choice-set individuals have as citizens (Aytaç, 2024, 5-6). Therefore, their participation in the public sphere is arbitrarily interfered with: social media companies have the capacity to do so.

2.1.3. Towards structural domination

With two minor argumentative moves I will interrelate structuralist and republican domination. First, (1) I argue that they aren't mutually exclusive if (i) structuralist obstruction of self-determination is regarded as the *systemic* arbitrary interference in choices and (ii) the structuralist distinction between domination and oppression is softened. Second, (2) using the work of Dorothea Gädeke, I argue the systemic interference with choice sets conception of domination is preferable for this thesis's objectives.

(1) Structuralist domination's virtue is that it takes into account background conditions: the (global) reach and history social groups' interrelations (Young, 2005, 709). Historical colonialism was a global phenomenon spanning multiple centuries after all. However, republican domination can also be applied in trans-state contexts. Lillian Cicerchia (2022, 10) convincingly argues the republican discourse's neglect for phenomena such as "slavery, colonialism, capitalist displacement" lies in that "[c]apacities to dominate may be inherent in one's *social position*, not just granted by the state". The way out of this apparent controversy is an "account

[10] The distinction between positive and negative liberty was popularized by Isaiah Berlin, as De Dijn (2020, 336) writes: "Berlin defended the key idea that freedom as noninterference—what he called "negative" liberty—was the only kind of freedom worth having, and that any other, more "positive" definition of the concept was an obfuscation or a lie".

of structural domination" tailored to "how a capacity to dominate results from the social relationships that shape the capacities that individuals have for domination" (Circerchia, 2022, 12).[11]

The first step towards this desideratum is made when structuralist obstruction of self-determination is reframed as arbitrary interference by drawing on Gädeke's (2020, 200; 210) *systemic* conceptualization of domination that "highlights the systematic disempowerment the dominated suffer over and beyond their relation to a particular dominator" in absence of a particular dominator with an actual capacity to interfere". Distinct from systemic domination is *interpersonal* domination: an asymmetric relationship of interference between identifiable persons (Gädeke, 2020, 206).[12]

Gädeke's account of structural systemic domination can subsume the structuralist account for two reasons. First: for Young and Gädeke, the dominator can be personal but needn't be. Structural domination can occur by virtue of a society's *system of power* itself rather than due to *specific* actors which are identifiable (interpersonal domination): society, the state at large or 'capitalism' are indeterminate actors that are *im*personal. These systemic 'actors' cannot be identified as individuals, but the control over the "institutional conditions" giving shape to peoples' capacity for self-determination lies with these systems, whether or not they actively enforce those institutional conditions. Young thinks the same is the case for oppression.[13] Second, Gädeke and Young (1990, 218) think of domination as the imposition of barriers on peoples' options for actions and choices: "freedom, in the sense of self-determination [is] the ability of people to decide what they will do and why". Therefore systemic domination encompasses the structuralist virtue that domination of social groups can proceed via the structures of power upheld by social groups/impersonal actors.

Turning now to Young's oppression/domination distinction, softening this distinction completes the compatibility with republican domination.

Young says oppression without domination is possible and vice versa. Reducing the inhibition of expressing and doing to obstruction of self-determination is therefore not self-evident. Young's paradigmatic examples of oppression are exploitation, cultural imperialism, powerlessness, violence, and marginalization. In §2.2. I show exploitation necessarily presupposes domination. The other named injustices are modifiers of individuals being inhibited in their self-development, because individuals lack the power to self-develop (powerlessness), another community's culture is imposed on them (cultural imperialism), or because they are excluded from the system of labor (marginalization). But through a republican lens, all these examples can be redescribed as the obstruction of social groups' capacities for self-determination and decision-making. In marginalization some actor bars the marginalized from *choosing to partake* in the labor system; in

[11] I will not further engage with Cicerchia's specific answers in the context of the labor market, however, the satisfaction of this desideratum is also what is needed in the context of colonialism.

[12] Gädeke also thinks that the structural character of power holds for both systemic and interpersonal domination, but her argument for that is irrelevant for our present purposes. Gädeke writes: "Domination, I propose, is best conceived of as generally characterized by the structurally constituted capacity to interfere (arbitrarily)." (Gädeke, 2020, 205). In turn, Young has written that "Domination consists in institutional conditions which inhibit or prevent people from participating in determining their actions or the conditions of their actions. Persons live within structures of domination if other persons or groups can determine without reciprocation the conditions of their action, either directly or by virtue of the structural consequences of their actions" (Young, 1990, 38). This excerpt makes clear how Young also thinks of domination in terms of the imposition of barriers on peoples' options for actions and choices.

[13] Some textual evidence corroborates this picture of Young's ideas: "While oppression in other parts of the world is also structural and systemic , in many societies it is easier to identify oppressor groups [than in the United States]" (Young, 1990, 258).

cultural imperialism ways of giving shape to indigenous culture *are being interfered with*; and in powerlessness institutional conditions are in place that *constrain freedom*.[14] It namely follows that if one agent has the capacity to determine how other agents express themselves and what they do, whether this is in accordance with those agents' wills or not, this capacity is one of arbitrary interference in Gädeke's sense. The result is that both structuralist oppression and domination are forms of republican domination, *systemically* construed.[15]

(2) The benefit for our current objectives is related to that the historical colonizers were collective agents (complexes of natural persons), such as large corporations (e.g., East India Company), states (e.g., the Dutch Republic) and other market participants) that were not diachronically stable. Different configurations of actors upheld the subjugation of the colonized communities for centuries on end. With respect to current indictments of 'AI colonialism' on the other hand, again, large multinational companies, states and market forces seem to together constitute and uphold forms of domination. The literal interpretation of AI colonialism *prima facie* lays claim to colonialism literally obtaining in AI injustices and these injustices are sometimes global and structural.

Therefore the retainment of the structuralist account's emphasis on domination pertaining to large-scale societal relations, in parallel with the republican emphasis of the determination of one party's options by another, is, firstly, beneficial for analyzing century- and continent-spanning colonization and its definitional elements. Secondly, it helps determining the role of the concept of colonialism in AI injustices: is there domination without colonialism, colonialism *tout court*, or no domination at all? Thirdly, using domination as *explanans* for injustices with a digital dimension (like the AI injustices that interest us here) is promising, as Hoeksema and Aytaç's respective successes show.

Finally, domination can be demarcated. *Structural domination* is conceived of as:

> the arbitrary interference with the choice sets of a social group (its self-determination) through the exercise of power or the capacity to do so that pertains to the structure of society.

The motivations for this definition are that (a) it is a prominent explainer of wrongs with a digital dimension and (b) it applies well to global injustices related to collective actors. This definition will be applied to exploitation and colonialism in the subsequent sections and to AI injustices in the next chapter.

2.2. Situating domination within exploitation

This section investigates exploitation and domination's conceptual relation, for two reasons. First, some regard exploitation a good explainer for colonialism's wrongness (Butt, 2013a). Understanding whether domination and exploitation are interrelated sheds light on their roles in colonialism, which contributes to determining whether domination and exploitation belong to colonialism's linguistically normative

[14] Violence is a special case that might not be domination per se, but is often a corollary of it: violence always exists in an *act*, as modifier of it. Likewise, there is room for the argument that violence per se is also not best conceptualized as form of oppression. It is however beyond the scope of this chapter to make this niche argument.

[15] This becomes clear from Dorothea Gädeke's summary of Pettit's account: "the mere capacity to interfere is sufficient to constitute the status-denying dependence on the will of someone else characteristic of domination" (Gädeke, 2020, 219). However, terminological trouble for my distinctions is spelled by how Gädeke calls 'domination's power relation that stretches over time' the 'structural dimension' of the concept. This dimension distinguishes it from 'mere interference' (Gädeke, 2020, 202). However, on analysis, this terminological convergence points to the convergence of Gädeke's conception of domination with Young's structural emphasis.

conceptual core. Second, exploitation is a candidate for explaining the wrongness of some AI injustices, such as worker exploitation, the datafication of human behavior for profit and the extraction of resources from countries in the global South (see §§3.1.-3.2.). I won't settle the debate on exploitation, but a robust operational definition is needed because subsequent chapters' analyses require a generally acceptable definition.

2.2.1. Exploitation with and without domination

With the notion of systemic domination in mind, I turn to the controversy over the primacy of domination over exploitation. John E. Roemer (1985, 47) contended that domination sometimes entails that there are exploiters and exploited, but that *exploitation isn't logically or pragmatically reducible to domination*. Nicholas Vrousalis (2013, 156) argued, *contra* Roemer, that "exploitation implies domination of workers by capitalists"; i.e., that domination is a necessary condition for exploitation. Why contrast these two accounts? On §2.1.'s conclusion, Vrousalis' account would make exploitation a *modifier* of domination. Roemer's view challenges this conclusion and could point to shortcomings in this domination-based account. I thus want to establish either (a corrected version of) the structural domination-based account or Roemer's non-dominative account of exploitation. In the end, I will motivate a choice for adhering to the domination-based account, because of shortcomings in Roemer's account.[16]

The following reconstructs Roemer's theory of exploitation in tandem with Vrousalis' domination-based theory. Afterwards the contrast is contextualized with respect to the aims of the thesis.

Roemer and Vrousalis jointly depart from rejecting the traditional Marxist conception of exploitation: the "exchange of labor for goods" that is unequal because the labor value of the goods the worker can purchase with his income is less that the "amount of labor" the worker is exposited for his income (Roemer, 1985, 30). Vrousalis (2018, 147) summarizes this as an agential extraction of "forced, unpaid, surplus labor". But qua endpoints, Roemer thinks an interest in domination (like we have), doesn't warrant theorizing about exploitation, while Vrousalis situates domination sturdily within exploitation as *necessary condition*.

So how does Roemer (1985, 31) define exploitation? Roemer doesn't regard it as a relation between two parties: it pertains to the wealth transfer between an individual and society at large: a *society* exploits a *coalition* (a grouping of members of the society). The two prerequisites for exploitation then are that (i) a specific social group A would be better off (would enjoy more of the fruits of its own labor) in alternatives where it

[16] There is a third account: Young's structuralist account. It conceptualizes exploitation as a structural form of oppression along the axis of wealth transfer. For the structuralist, exploitation is the systematic usage of the labor of one group to the benefit of another group. It is the systematic transfer of wealth and power through the labor of workers from one social group (working group) to another social group (capitalist group) that impairs the self-realization of the working group. On this view, it qualifies as form of *oppression* (Young, 1990, 49-50; 53). The transfer of powers characteristic of exploitation (impairing freedom, self-realization, etc.) makes that feminists have emphasized different modes of exploitation, gendered, racial, and sexed modes, with class being the paradigmatic mode (Young, 1990, 51). Exploitation, for the structuralist, is as such a concept that can be combined with other forms of oppression to give rise to a concept that explains a niche, hybrid injustice. Take the case of race for example. As Mills have shown, slavery was a form of racial *super*exploitation, which shows exploitation can differ in scale (Mills, 2017). Exploitation is most frequently understood as a problem of distributions. This is not wholly mistaken according to Young, but she is wary of the conclusion that exploitation is extinguished by distributive justice alone. Young contends that reform of institutions (of decision-making) are required, because not only the distributions are unjust (Young, 1990, 53). Following §2.1., structuralist exploitation can be seen as a modifier of systemic republican domination.

is treated differently than in the current society and (ii) the current society would be worse off if *A* is treated differently (would enjoy less of the fruits of *A*'s labor) (Roemer, 1982, 194). In layman's terms, the first clause holds that if you're part of *A*, there are scenarios available where you would receive your labor's worth, whereas you currently don't. The second clause: if you're a part of the society enjoying parts of the fruits of *A*'s labor, then in alternative scenarios where *A* is better off, your net welfare due to *A* decreases. For example, consider the case of a worker in a smartphone factory and a smartphone manufacturer, the owner of the factory. Here the worker *is exploited* and the manufacturer *is an exploiter* by virtue of the society they work in. The worker toils to make smartphones and receives wages. The manufacturer sells smartphones and its net welfare increases, because in the current society smartphones sell for more than the worker is paid. This is a Roemerian case of exploitation because in an alternative society, the worker would receive his exact share in the smartphone sale's profits, in which case the worker is better off and the manufacturer is worse off.

How does Vrousalis' definition differ then? Vrousalis (2013, 132). *does* conceive of exploitation *as a relation* between parties:

> *A* exploits *B* if and only if *A* and *B* are embedded in a systematic relationship in which (a) *A* instrumentalizes (b) *B*'s vulnerability (c) to extract a net benefit from *B*.

Here, 'instrumentalization' denotes one actor using another as means. How instrumentalization proceeds depends on the exploited's vulnerability. Vulnerability is either *absolute* (the agent faces an unavertable risk) or *relational* (the agent is at risk in relation to another particular agent). Relational vulnerability is a helpful concept, because of its use in comparatively determining two or more actors' vulnerabilities with respect to each other. Lastly, one can read "extract a net benefit" as an increase in one agent's well-being (Vrousalis, 2018, 132-135). Vrousalis (2018, 136) restricts his account of exploitation to *economic* vulnerability, where the vulnerability condition takes into account differences in position with regard to relations of production, which are "systematic relations of effective ownership, and therefore of power, over human labor power and means of production in society". Applying this to the smartphone example, the worker isn't exploited by a *society*, but by the *manufacturer*.

Economic vulnerability's precedence by power relations means one agent can have economic power over another. Therefore domination is at the heart of exploitation (main point of contention vis-à-vis Roemer). Vrousalis (2013, 139) collapses the power-based account of economic exploitation into his account of domination that is *the systematic relation between agents where one agent disrespectfully takes advantage of its power over the other agent.*[17] The collapse is enacted by first taking 'instrumentalizing another's vulnerability' to imply "taking advantage of one's power over another". This is a solid move, because treating someone as means thanks to some vulnerability implies that one can do so and that one has the power this requires. Second, he appeals to intuition to show that if one actor instrumentalizes another's economic vulnerability, then that is *disrespectful* to the instrumentalized party (Vrousalis, 2018, 141). This is a weak point in his argument (as appeals to intuition often are), but since I share the intuition, I will not make a problem out of it here.[18]

[17] Does this correspond to systemic republican domination? Yes, because a systemic/structural relationship is identified in which one party can exercise its power over another, which can be translated into one party being at liberty to affect another's choice set.

[18] So Vrousalis thinks domination is important while Roemer does not – at least, not anymore? Peculiarly, Vrousalis harks back to Roemer's earlier work where a dominance condition *was* present. A third condition was taken up, the 'dominance' condition: the society "is in a relationship of dominance to" social group *A*. Roemer added the following

Roemer (1985, 39) however distinguishes between two forms of domination: (domination[1]) "domination in the maintenance and enforcement of private property in the means of production" and (domination[2]) "domination at the point of production (the hierarchical and autocratic structure of work)". While both forms of domination can engender situations that lead to exploitation, the converse doesn't hold: a case of exploitation isn't necessarily a case of domination.

Domination[1] can lead to exploitation because of (a) an unequal distribution in the ownership of the means of production that is state-enforced or (b) when the value of a property isn't determined by the market. In (b), 'intrinsic value' or asset worth enables owners to exhibit coercive power over another by controlling the value in the market. Roemer's example of the latter is that of an irrigation canal, but one can think of others, such as scarcity in housing, passing the Suez-canal, etc. But on Roemer's view of this scenario, exploitation is an unimportant corollary of the domination present in fixing the market value. Exploitation can entail domination if there is distributive injustice in the means of production, but then exploitation is the source of domination (a contingent corollary). So domination[1] and exploitation aren't fundamentally related and when they are, this is always *contingent*.

Domination[2] concerns the *employment* of some by others, not necessarily some being exploited (Roemer, 1985, 35). Its unequal standings in interpersonal hierarchies put the conditions for exploitation in place, but it doesn't follow that exploitation *implies* domination[2], as Roemer (1985, 42-43) claims that *everyone* 'selling labor/taking credit' is exploited and that everyone who 'hires laborers/gives credit' exploits in accordance with his definition. Some types of capitalism imply domination[2], but others do not, while being fundamentally exploitative. Hence domination[2] cannot be a necessary condition for exploitation.[19]

---

footnote regarding this 'dominance condition': "Although the dominance condition is not defined, it should at least entail the following: that the coalition of *S'* [society] prevents the alternative from being realized, which gives rise to its exploitation of *S* [social group]. Because I do not pursue the dominance issue here and various pathological examples, the account of exploitation here cannot be considered exhaustive" (Roemer, 1982, 195). This dominance condition reappears only once more on page 237 of *A General Theory of Exploitation and Class*: "Finally, it must be pointed out why the dominance condition (3) of exploitation is necessary." Here Roemer comes up with the example of a disabled person that extracts value from the rest of society, without returning the labor value due for the exchange to be equal. He does not want to admit that the disabled person *dominates* society; instead, society dominates the disabled person by foreclosing some possible choices. On the same page, in another footnote, Roemer tells us someone else pointed out the dominance condition to him and that in a later publication he used the first two game-theoretic conditions only. Vrousalis uses Roemer's example of the disabled person himself (though without evocation of dominance on page 145 of (Vrousalis, 2013). While Vrousalis is sympathetic to the early Roemer, he contests the views of the later Roemer.

[19] Roemer claims economies without unjust hierarchies of power are possible (so without domination[2]) and the example he gives is that of a credit market (lending and taking loans only) where exploitation is replicated without hierarchical relations being present in this economy. See (Roemer, 1982) for an in-depth and technical development of this hypothetical credit market economy. For Roemer's class division of society, "those who only hire others" and "those who hire others and work on their own account" can be the *only* dominators in an economic sense (Roemer, 1985, 45). He argues a third class "those who only work on their own account" can neither be dominated or dominating in correspondence to them being exploited or them doing the exploiting. Vrousalis concludes that in technocracies like today's societies, "Differences in talents and knowledge can generate domination, thus paving the way for "postindustrial" forms of economic exploitation" and that nontechnocrats are exploited by them (Vrousalis, 2013, 157).

With Roemer's view in place, I return to Vrousalis, who argues that Roemer's domination[1&2] are actually not the only types domination, a fact he uses for his subsumption of exploitation under systemic domination.

What Vrousalis takes issue with is that when prior inequalities (pertaining to the society as a whole) are modified by individual preferences (creating voluntary disadvantages), they don't become the *root* of exploitation. Roemer thinks a social group is exploited by society as a *whole* and that exploiters are only indirectly related to the exploited,[20] which conflicts with Vrousalis' account's fundamental relationality. That being said, Vrousalis agrees with Roemer's argumentation regarding domination[1&2]. But how can he do so without relinquishing his own argument? His tricky response is showing there are *more* types of domination than Roemer's selection. Even in perfectly competitive markets, owners of the means of production instrumentalize nonowners' vulnerability and "get" them to do things they "would not otherwise have done […] in order to make money" (Vrousalis, 2018, 155). Vrousalis concludes Roemer has not disentangled exploitation from domination.

2.2.2. Characterizing exploitation

The *relationality* of Vrousalis's notion of exploitation is indispensable for defining colonialism, but Roemer's game theoretical possible economies aren't. Possible parlance on colonialism like 'The Dutch Empire colonized the Indies and exploited the people of Java', requires a clearly identifiable colonizer/exploiter standing in a direct relationship to a colonized/exploited. So if the colonized are exploited in colonialism, it is *the colonizer* that does so, not 'global society at large'. Consider the awkwardness of giving it a Roemerian twist: 'The Dutch Empire colonized the Indies and the people of Java were exploited by global society at large and the Dutch Empire was an exploiter in that society'. The direct ethical significance of one party dominating another at the heart of instrumentalization is lost here.

Concludingly, I think Vrousalis rightly harked back to the early Roemer and his dominance condition.[21] Vrousalis stressed the *disrespect* every dominating relationship expresses. Roemer's exploitation may suit abstract game theory, but not the concrete project undertaken here. Just like the game-theoretical concept of Pareto-dominance is seemingly harmless until it is applied to human beings, so the capability of exerting power over another cannot be left out of a conception of exploitation. As a result, I will regard *exploitation* as a special case of *domination* (in line with §2.1.):

> The relationship between two actors such that one actor dominates the other in order to instrumentalize the latter's vulnerability for the purpose of extracting a net benefit.

From §3 and onwards, this definition figures in the analysis of AI colonialism and in the following section, it is employed in order to show it isn't a necessary feature of colonialism.

[20] The argument made by Vrousalis is not directly relevant for the main text. He does this by showing how exploitation need not depend on asset injustice: if *A* can rescue *B* without costs, but asks money for it, *A* exploits *B*'s vulnerability. It seems that where Roemer wanted to distinguish between ethical and technical exploitation, Vrousalis reunites the two in the most elementary cases: "The reason why exploitation survives in the absence of distributive injustice, […] is that exploitation constitutes procedural injury to status, and status-injury is not reducible to distributive injury" (Vrousalis, 2013, 149-151).

[21] See footnote 18.

2.3. Defining colonialism

Defining colonialism can finally commence. §§2.1.-2.2.'s results enables discerning the conceptual relationships between domination, exploitation and colonialism. How were the colonized wronged and what is common to colonialism? These questions evade straightforward answers. Following Sabera (2021) and Butt (2013a), close attention needs to be paid to the difference between colonialism's distinctive wrong and its full definition. The following slowly approximates colonialism's defining features by surveying accounts of colonialism's wrong(s). Most accounts agree colonialism entails some form of domination. Equally uncontroversial is that colonialism was morally wrong and that its historical instances involved domination, cultural imposition and exploitation.

To ease the exposition, I introduce two combinable lines along which I classify the different approaches: (1) *colonialism's procedural (in)distinctiveness*; and (2) *single-wrong* or *multi-wrong accounts of colonialism*. These categories enable intelligibly representing the debate. Colonialism's procedural (in)distinctiveness comes in two flavors:

> *Proceduralism*: Colonialism is constituted by a form of domination.
>
> *Anti-proceduralism*: Colonialism is wrong for non-procedural reasons.

Three stances are available regarding colonialism's *constitutive wrongs*:

> *Single-wrong*: A single noncontingent wrong is distinctive of colonialism.
>
> *Multi-wrong*: Colonialism is a conglomerate of multiple noncontingent wrongs.
>
> *Indistinctive-wrong*: Colonialism is constituted by a (conglomerate of) contingent wrong(s).

Because these positions are combinable, there can be 'single-wrong proceduralists', 'multi-wrong anti-proceduralists', etc. Important to note is that none of these positions wants to disqualify colonialism's conceptual usefulness.

2.3.1. Introducing multi-wrong proceduralism

I will start with 'multi-wrong proceduralist' views that hold colonialism is defined by a *combination of wrongs*, one of which is the dominating relationship between the colonizer and the colonized. On this view, not all wrongs that constitute colonialism have to be procedural.

Consider first Butt's (2013a, 2) simple threefold definition: colonialism is a conglomerate of "*domination*, *cultural imposition* and *exploitation*". Butt's colonialism clusters wrongs and the three features make the event in question 'colonial'. Here domination is understood as the subjugation of a *people*, regardless of the institutional means by which subjugation is realized; cultural imposition signifies the forced adaptation of customs, religions and knowledge by the colonized; and exploitation is the appropriation of human and natural resources from the colonized to be capitalized upon by the colonizer. The colonizer dominates and instrumentalizes the colonized in order to extract wealth from this dominion.[22]

Other multi-wrong proceduralists provide a phenomenological deepening of Butt's criterium of cultural imposition. Colonialism not only involved the forced adaptation of customs and culture, but also the disruption of identity through the racialization and epistemic disqualification of the colonized. Uchenna Okeja's (2022, 205) characterized colonization as a "conceptual adjustment program" that replaced the rule

[22] Note how this loosely fits the previous definitions of domination (as a community's self-determination is obstructed (their capacity for decision-making is systematically interfered with)) and exploitation (as domination for self-enrichment).

of customs by written law, polygamy by monogamy, and polytheism by monotheism through the exertion of extreme forces of domination and conversion, thereby committing an epistemic vice. Likewise, historian Sabelo Ndlovu-Gatsheni (2019, 208) stresses that "the commercial non-territorial empire and the cognitive empire are, inextricably intertwined". In the work of Fanon we find these cultural, epistemic and racial aspects of colonialism theorized together. Fanon's epigraph to this chapter already challenged the endeavor of providing an account that satisfies *all* of the colonial wrongs experienced by the colonized. I will return to Fanon's remark later. Now I discuss the *racialization* and *violence* he thought characteristic of colonialism next to the dominating relation in colonization.

Fanon (1963; Mbembe, 1992) sought to bring into perspective colonialism's psychological and racial dimensions which hang together with its institutional dominion. Integral to Fanon's characterization of the experiences of the colonized was the conceptualization of their relation with the colonizer as a 'double consciousness'. Double consciousness can be understood as an additional reflexive awareness of oneself. The standard reflexive awareness is that of *self*-consciousness. The second self-consciousness comes about by the internalization of the Other's (the colonizer) *racializing* gaze (Fanon, 2008, 169-170). This is existentialist jargon that says the colonized perceive themselves according to the colonizer's dehumanizing evaluation of them. The colonizer views them as subhuman, objects to be used.

According to Achille Mbembe, colonial domination makes possible that the colonized become "exploitable object[s]". He explains that "Blackness" – colonialism's mark of the subhuman – is the product of "a social link of subjection and a body of extraction, that is, a body entirely exposed to the will of the master, a body from which great effort is made to extract maximum profit" (Mbembe, 2017, 18). The colonized's double consciousness comes about due to the colonizer's gaze: the colonized are degraded to the subhuman, which they internalize. This internalization leads the colonized to associate subhumanity with their bodies rather than with the colonizer's valuation. The nonwhiteness of the body of the colonized becomes entangled with subhumanity and is ontologically opposed to the whiteness of the colonizer's body. Fanon (2008, 172) uses the neologism 'epidermalization' to denote how this domination of colonial whiteness pins the Black man's degradation and dehumanization onto his own skin.

Epidermalization is part of the broader phenomenon Mbembe calls 'racialization'. Racialization is integral to the colonized societies' relation to the colonizer: "Racialisation must be understood as the capture and conscious deployment of a set of techniques of power […] that aim at producing a reality, namely race, that there is then a concerted attempt to naturalize" (Mbembe, 2022, 15). On top of the psycho-existential consequences Fanon writes about, racialization is also felt in political obstructions of citizenship: the invention of race enabled the colonizer to impose identities and disable civil participation (Mbembe, 2019, 61). The concept of race, "a mechanism for objectification and degradation" (Mbembe, 2017, 162) was invented during colonialism and employed to *dehumanize* and *rationalize the subjugation of the colonized and their exploitation.* Mills (1997) argued that the concept of race permeates and underlies modern conceptions of the state as a social contract with citizens. Furthermore, Mills (1998) adds, 'whiteness' (denoting an ontological category based on an idealization of skin color) defines who counts as human in Western societies.

Since in colonial domination two communities are fundamentally opposed, writer Édouard Glissant (1997, 17) wrote that "the Western nation is first of all an 'opposite,' for colonized peoples identity will be primarily 'opposed to' —that is, a limitation from the beginning". Equally, in Fanon and Mbembe's accounts of colonialism, the domination in question is always institutionally realized: the dominion of one community ("Western nation") over another ("colonized peoples") constitutes the frame that prefigures racialization.

The postcolonial tradition following Fanon also regards *violence* of one culture against another as constitutive aspect within the frame of colonial domination. Mbembe (2017, 105-106), citing Fanon, explains three forms of violence are involved in colonialism's "*struggle to the death*", "on the level of the spirit as in "muscles, and blood"": "violence in the "daily behavior" of the colonizer toward the colonized; "violence in regard to the past" of the colonized […]; and violence and insult in relation to the future". These forms of violence terrorize the colonized, who are continuously prey to physical and mental assaults and lack the protections of citizenship. However, violence isn't monolithic: it's an *intentional harming* and so general that it's implicated in colonial and noncolonial injustices alike. Therefore it can't be deemed distinctive of colonialism. As will become clear later, the harm of dehumanization in racialization explains what is *semiotically* wrong with colonialism.

So far, Butt named domination, cultural imposition and exploitation and Fanon and Mbembe added racialization and violence. As horrifying a testament of colonial existence as Fanonian violence and racialization are, they further complicate the predicament Butt left us with: determining a wrong that's *distinctive* of colonialism. Distinguishing *constitutive wrongs* and *wrongful effects* shows that many harms pertaining to colonial regimes are *facilitated* by colonial rule, rather than constitutive of this rule. Considered as wrongs in and of themselves, they can be concretely explained by reference to injustices other than colonialism (e.g., murder or rape). Surely, violence was a systemic aspect of colonial rule (as Fanon (1963) showed for Algeria and Mbembe (1992) for the 'postcolony') and integral to the understanding of a particular regime's coloniality. This systematicity can be acknowledged while simultaneously being classified as 'wrongful effect' that is *colonial by association*. Violence and racialization take place *within* the colonial frame, i.e., within the vile and degrading relationships made possible by schema of colonizer-colonized, but don't *create* or *constitute* the colonial frame. While we have elucidated important wrongful effects, the search for the wrong of the *colonial frame itself* continues.

2.3.2. Single-wrong proceduralism *vs.* indistinctive-wrong anti-proceduralism

The previous multi-wrong accounts were analyzed to be incomplete because of the 'wrongful effects' category. This section will turn to the debate over the centrality of domination in colonialism: is a single form of domination distinctive for and constitutive of colonialism?

'Single-wrong proceduralist' Lea Ypi rejects that colonialism's central wrong is the infringement upon territorial rights or nation-states' boundaries. Why? Ypi (2013, 161) thinks settler colonialism is only a contingent *manifestation* of colonialism, just like population extermination and economic extraction of natural resources.[23] *Many* things are wrong about colonialism, but most of them are of linguistically *descriptive* rather

[23] The infringement of territorial rights account is insufficient for colonialism's injustice, because Ypi shows how the principle to 'leave as much as is good to newcomers' can underpin forms of foreign settling by introducing the obligation of original inhabitants to share with migrants in need, and neither is occupation in itself sufficient, because the justifications for occupation being a wrong per se are lacking, as it is conceivable that it would benefit the newly occupied in terms of rights and jurisdiction and, in some cases, historical forms of colonialism have been able to flourish without having to occupy any state. The example Ypi gives here is that of the cooperation between the British East India Company and the home authorities of the colonized state (Ypi, 2013, 166-173). Equally, she considers the 'status quo objection' to investigate how protected native inhabitants are from the duty to share land with migrants in need, she concludes that "Even if claims to appropriate land and resources can sometimes be permitted, no legitimate institutions can be constructed on their basis" and that this holds for both settler and commercial variants of

than of *normative* importance for defining colonialism (Ypi, 2013, 163). Ypi's (2013, 158; 185-186) counter-proposal is that colonialism's wrong is the "creation and upholding of a political association that denies its members equal and reciprocal terms of cooperation". The denied "ideal of equal consideration of each other's claims and of reciprocity in communication" should underlie every political association and this vindicates the rejection of the territorial rights approach because the duty to associate equally and reciprocally "takes precedence over the recognition of self-referential claims to territory" (Ypi, 2013, 174-176).

Liberally understood, Ypi's cosmopolitan ideal of equality and reciprocity in political cooperation fine-tunes Butt's criterium of domination: the failure to associate with another group equally and reciprocally constitutes an arbitrary interference in the social group that is *denied* full political association.[24]

Laura Valentini's 'indistinctive anti-proceduralism' however opposes Ypi's single-wrong proceduralism. Valentini acknowledges many factors made colonialism morally wrong, but claims nothing is *distinctive* for colonialism. Valentini (2015, 330) generally agrees with Ypi, but deems the "procedural wrong of unilateral takeovers"[25] too strong qua being "contingent on the target collectives' meeting standards of reasonable justice". What does this mean? I read it as that Valentini thinks Ypi's conclusion is doubly mistaken: (1) the violation of equality and reciprocity in decision-making isn't *noncontingently* procedurally wrong; and (2) it isn't "distinctive" of colonialism.

Correspondingly, Valentini distinguishes the *aggregate* interpretation ('a political collective is a conglomerate of the wills of individual members') and the *corporate* interpretation ('the collective as a whole has a will of its own') of Ypi's view. Supposedly neither stands on analysis. The aggregate interpretation leads to an 'overascription of wrongdoing': neither the violation of equal subjection to laws, its negative impacts on subsequent generations, nor negligence of conflicts with all colonized individuals' wills are distinctive for colonialism, because protections to these things aren't even realized in contemporary democracies; surely those aren't colonial. Valentini's corporate interpretation challenges the will-tracking of the *collective*. The collective's will, she claims, can only be the *aggregate of all individuals' wills*, which makes the interpretation susceptible again to the 'impossibility of will-tracking argument' against the aggregate interpretation (Valentini, 2015, 316-323). So long for (2). For (1), Valentini's (2015, 321-326) counterargument is that it isn't necessarily wrong to take over a reasonably unjust political collective, if this is in the interest of its members. The political take-over inherent to colonialism could be *just* by virtue of it taking over a *more unjust* political collective. The colonial rule improves on its predecessor in terms of justice.

Where does Valentini leave our prospect of identifying colonialism's distinctive wrongs? Valentini (2015, 328) concludes colonialism's distinctive wrong isn't procedural, but that there are still wrongs familiar to colonialism: she thinks colonialism is wrong for many reasons, but just not *specifically* for Ypi's one.

However, Sabera (2021) defends Ypi's 'single-wrong proceduralism' against Valentini. Firstly, Sabera (2021, 63-64) rejects that Valenini's 'just unilateral takeover' rebuttal of Ypi's view works. Valentini's example of 'Sweden *justly* taking over Canada' in the interest of the Swedes because it replaces a harmful government with a better one, is namely inconceivable "without any sort of coercion or manipulation" taking place that *does* violate the self-determination of the subjugated. I agree and consider this a refutation of Valentini's rebuttal, because the colonial replacement of an unjust rule with a more just one doesn't change that Canada imposes itself as ruler on the citizens of Sweden. This shows that Valentini's argument (1) is mistaken

colonialism, so "both residents' unilateral enjoyment of certain benefits and outsiders' unilateral appropriation of them are wrong"(Ypi, 2013, 183-184).

[24] This is congruent with §2.1.'s results and with the Fanonian wrongful effect of violence.

[25] This is Valentini's paraphrase of Ypi's 'equality and reciprocity in political cooperation'.

because it still denies its new members equal and reciprocal terms of cooperation: they become colonized rather than liberated, even if this is, under some moral calculus, 'for their own good'.

Secondly, Sabera thinks Valentini's aggregate/corporate distinction misconstrues the difference between individual and collective as solid. She proceeds to strengthen Ypi's view with a duty-based account that corroborates colonialism's inherent wrongness. But this loses the identification of wrongs *distinctive* of colonialism out of sight, because duties to strive for just institutions apply to *any* imposed rule: it doesn't say *why colonialism is wrong in contrast to other injustices*. But Sabera's criticism of Valentini stands because of Han van Wietmarschen's contrasting of Valentini's aggregate and corporate interpretations with a third one: the *plural* interpretation of the colonized. Applied to his running example of the colonization of Java by the Dutch government, the plural interpretation holds that "the Dutch rule of Java instantiated the wrong of colonialism if and only if the Dutch subjugated the relevant plurality of Javanese people" (Van Wietmarschen, 2018, 173). Although criteria for the plurality's sufficient relative size are lacking, this evades Valentini's charge of overascribing wrongs because now it suffices that a *plurality* of the collective's members is harmed for colonization to ensue. 'Will-tracking' of *every* member of the collective thus doesn't play a role anymore (Van Wietmarschen, 2018, 173-174), which shows that Valentini's argument (2) is also misguided.

Van Wietmarschen and Sabera's arguments therefore show that Valentini's criticism of single-wrong proceduralism wasn't justified.

2.3.3. Multi-wrong *vs.* single-wrong proceduralism

With the discrediting of anti-proceduralism, proceduralism seems vindicated: the obstruction of the colonized's self-determination likely is *the* constituent of the colonial frame. However, it hasn't been settled if it's the *only wrong* constitutive for colonialism. We need to know whether colonialism's conceptual core contains one wrong or multiple wrongs.

Margaret Moore's multi-wrong proceduralist position compares Butt's, Ypi's and Valentini's[26] accounts. What complicates any single-wrong account, is Moore's (2016, 448) difficult question of distinguishing colonialism from imperialism. Any single-wrong account should be able to show what is *unjust* about colonialism in particular. Regarding Butt's criterium of exploitation, Moore thinks that on Vrousalis' definition colonialism was rampant with exploitation. But she counters that: (1) not all exploitation is colonial and exploitation doesn't cover *all* instances of colonialism; and (2) colonialism wasn't perceived as unjust at the time. If true, (1) prohibits exploitation from being a necessary condition for colonialism. What sways me to accept (1) is Moore's (2016, 450) argument that deeming exploitation "an unavoidable feature of colonialism" is "problematic because there is in fact considerable controversy about the extent to which the exploitation narrative applies to all the empires and across all the colonized areas"; a non-exploitative colonizer is conceivable and exploitative rule can be continuous across the transition from colonization to decolonization.[27] But (2) is unconvincing. Consider the following schema. 400 years ago, *X* intentionally and cruelly

[26] Against Valentini's conclusion that colonialism was only contingently unjust, Moore brings to bear that it can no longer explain the "deep intuition that colonial rule invariably wrongs those subjected to it" (Moore, 2016, 449). While Moore doesn't discuss them, we can see in this thought a ground for the support of arguments like that of Sabera and Van Wietmarschen.

[27] Moore's example here is the case of a decolonized state that still dominates and exploits its citizens – would that also count as colonialism? I concur with Moore that this isn't so.

killed *Y,* but didn't regard killing as unjust. Is our contemporary *Z* then *not* justified in regarding *X*'s killing of *Y* morally wrong? Surely not.

Regarding cultural imposition, Moore obstructs the inference that mere cultural change due to intercultural contact is cultural imposition. Cultural imposition is an epistemic injustice with psychological and social structural components, e.g., what Fanon accounted for in relation to colonialism. But the conditions for what makes something *imposed*, defer to colonialism's domination criterium and aren't found in cultural imposition itself (Moore, 2016, 450-451). Therefore I support Moore's conclusion that cultural imposition isn't constitutive for colonialism.

For Moore (2016, 453), domination is the "*destruction* of previously self-governing communities". Building on Anna Stilz's (2015) work, Moore asserts colonial rule's wrongness is that the colonized cannot subjectively affirm the political institutions they are dominated by: they cannot *choose* to being ruled like this. Political domination qua alien rule is Moore's *central wrong* of colonialism: *disrespect* is communicated by the colonizer's obstruction of a people's forming of collective communities. And Moore (2016, 459) adds the *taking of land*, such that both criteria oppose "colonized people's collective aspirations to be self-governing […] within their own territory". Land-taking is an injustice because, even without expelling inhabitants, it disrupts meaningful aspects of life. But given Ypi's displacement of territorial rights as criterium, does Moore's reintroduction work? No, because it only defines *settler* colonialism, not colonialism in general (Ypi, 2013, 161). While disruptive, land-taking is a *wrongful effect* of colonialism, and only necessary for settler colonialism: subjugation in the colonial sense can proceed without territorial displacement.

Taking stock, domination's centrality for colonialism is clear. Notable is domination's *semiotic* aspect: the disrespect expressed by the mere act of colonization. Domination semiotically communicates a disrespect towards the colonized's members, which explains what's wrong with colonialism. Also, by Vrousalis' (2013, 139) definition, one should see exploitation as *modifier* of domination: exercising one's power over another *for profit.* The disrespect Vrousalis situated in exploitation because of instrumentalization matches the usage of the colonized for profit. This facilitates colonialism's rampancy with exploitation, and explains why Moore is right that colonialism isn't necessarily exploitative. But the quest for colonialism's constitutive wrong isn't finished: the disrespectful subjugation of the colonized requires further analysis.

2.3.4. Completing single-wrong proceduralism

Previous accounts pointed towards domination's residence in colonialism's conceptual core: colonialism is a subclass of domination *or* significantly overlaps with domination. The subclass view, as defended by Ypi, Van Wietmarschen and Sabera, emphasizes the institutional enforcement of political association without equal and reciprocal cooperation. Contrarily, Butt's overlap view added little semantic content to the 'subjugation of the colonized'. Moore sits somewhere in between by regarding colonial domination as disrespectful alien rule. To unite these views, a final discussion of proceduralism is needed.

Massimo Renzo (2019, 362) combines Ypi's account with Moore's alien rule accent and argues that the *subversion of the capacity for self-determination of the colonized*[28] is the distinctive colonial wrong. This addresses the still open question of the difference between imperialism and colonialism head-on and is the *specific way* colonialism undermines self-determination, distinct from undermining self-determination in general (i.e.,

[28] Renzo leaves it at that and does not seek a complete description of colonialism in terms of necessary and sufficient conditions.

other variants of structuralist domination, such as patriarchy). The subversions' two components, for political communities *V* and *A* are, according to Renzo (2019, 362):

1. *V's* collective deliberation is controlled by *A* insofar as *V*'s members' inputs have (little to) no weight in *A*'s deliberative process; and
2. *V*'s members lack legal rights against being treated this way.

The colonized (*V*'s members) can no longer author laws, policies, etc., because their decision-making power is usurped by *A* (the colonizer) (Renzo, 2019, 366). These conditions distinguish colonialism from annexation, because in annexation, a political community becomes part of a *different* political community, which absorbs them, and, according to Renzo (2019, 363), the absorbed community's institutions cease to exist. It can still be dominated, but ceases to be a self-authoring political community. To contrast, colonialism denies the colonized *any political membership* and so the political community (the periphery) persists but is denied self-determination without being absorbed into the colonizer (the metropole).[29]

This distinguishes colonialism from patriarchy. Patriarchy is the oppression of a society's members externally identified as women. One could argue women's self-determination is disrespectfully subverted here, but the devil is in the details: the obstruction of women's self-determination occurs *within* a society that women share with their oppressors/dominators (men). Women aren't a separate political community originally entitled to self-determination of their own as separate from men. Patriarchy concerns the 'intrasocietal' subversion of self-determination. Colonialism however concerns *inter*societal subversion of self-determination. One politico-cultural community (colonizer/metropole) subverts the self-determination of another (colonized/periphery), without the latter *absorbing* the former. As subclasses of domination, patriarchy and colonialism are sibling concepts, but their scopes of application differ.

Now I can end with a metadiscussion of all the previous accounts' overlaps. The discussion over multi-/single-wrong (anti-)proceduralism can be more tractable rephrased. Synthesizing the surveyed authors' accounts, the type of domination particular to colonialism circumscribes the following features:

i. *institutionality*, one political community non-accidentally takes control over another;
ii. *intentionality*, the colonizer is responsible for this exertion of control;
iii. *obstruction of self-determination*, the colonized are dominated because their self-determination as community is obstructed; and, lastly,
iv. *disrespectfulness*: with Fanon we can say that the process of this obstruction doesn't respect the colonized's dignity in their treatment as less than human.

The takeaway is the importance of the dominating relationship between colonizer and colonized for defining colonization. Mere obstruction of self-determination isn't enough: one community should intentionally and institutionally subjugate another, such that the colonized are denied political membership. If this isn't the case, 'colonizers' and 'colonized' which – just like Roemerian exploitation – don't stand in direct relationship to one another, but are only vaguely connected through 'the mist of society', would be possible. The institutional subversion of self-determination is disrespectful in a dehumanizing way, because intercommunal reciprocation is denied, which treats the colonized community as subhuman: goods to be bartered in, objects

[29] Renzo writes: "When *V* is colonized by *A*, *V* persists as a distinct political community (the periphery), […] but the will of *V* is now subjected to the will of *A* (the metropole)" (Renzo, 2019, 364). In short, the colonized are prevented from shaping "the future of their community in light of choices autonomously made" (Renzo, 2019, 372-373).

to be controlled. Racialization is a particular way in which this has historically played out. Exploitation isn't essential for colonialism, but, as modifier of domination, is a *motive* for colonial dominion.[30]

To conclude, the dominative procedure of erecting the colonial frame creates the colonized as subhuman subjects or what Fanon (1963) calls the "wretched of the Earth". The institutional denial of self-determination establishes the disrespect that Moore and Renzo stress and the dehumanization of Fanon and Mbembe's racialization: they are intertwined into a single feature. While dehumanization can be construed as a separate harm, its combination with domination constitutes a *double communicative wrong*. The *mere act* of colonizing curbs the autonomy of the colonized community's members. This facet could be said to be communicated by domination in general. But Fanon and Mbembe's dehumanization goes further: the colonial subversion of self-determination renders the subjugated individuals *subhuman*, denying them parts of their humanity through the withholding of membership to the political community. So alongside the subversion of self-determination, colonialism's semiotic facets are (1) the disrespect communicated by domination and (2) the denial of humanity.

2.4. Conclusion: operationalizing colonialism

The previously surveyed authors didn't inquire into the *roles* colonialism's injustices play in its conceptual orrery.[31] Any attempt to define a concept is a *practical approximation* (see §1) – an operationalization for a specific purpose – which is why I employed the distinction between constitutive wrong (criterium) and wrongful effect (symptom). Not every instance needs *all* wrongful effects. As I showed, cultural imposition, exploitation, land-taking and violence are wrongful effects of the colonial frame, not the constitution of that frame. The constitutive wrong I discerned is the institutional and dehumanizing subversion of self-determination of one political community by another, which is a specific form of domination. This pattern runs through instances of colonialism in accordance with colonialism employed as linguistic *norm*. The wrongful effects 'travel together' with this linguistically normative core and shape colonialism's role as *description*. Colonialism can thus be part of two language-games, one pertaining to linguistic normativity and another to linguistic descriptivity.

Without the periphery colonialism's historical diversity would be lost on us. Using Wittgenstein's concept of family resemblance (Pitkin, 1972, 65) it becomes clear that the instances of colonialism form "a complicated net of similarities that overlap and intersect; sometimes fundamental similarities, sometimes similarities in details" (Wittgenstein, 2009, §66), like for any concept that partitions human behavior. As Aytaç (2021, 424) states, language-games, with their interrelated patterns of resemblance, "are not reducible to conceptual schemes" which points to the tension between operationalizing colonialism and that meaning will always be a function of use by a "communicative network of individuals". Fanon's epigraph expressed awareness of this intertwinement of language and embedding in civilization: having a *certain* language implies possessing "the world expressed and implied by that language" (Fanon, 2008, 9); having *this* but not *that*

[30] A politically dominating colonialism that deliberately does not exploit people or their natural resources in the strict sense of profiting from them is strictly conceivable (think of Roemer's first two necessary conditions for exploitation). Exploitation thus isn't a necessary condition.

[31] As a sidenote, I state that the final conceptualization should transcend the problem of anachronism any charge of colonialism faces. Colonialism shouldn't be restricted to clustering historical events: it should also apply to present-day phenomena that instantiate its conceptual core. A new need has arisen, namely that of subsuming injustices involved with AI systems/D&D under socio-culturally inherited categories of injustice. Therefore one can take a stand on which similarities of the instances of colonialism are most important and "justifiable to *us* now" (Queloz & Cueni, 2021, 16).

conception of colonialism implies possessing a way of seeing social reality that takes *this* to be colonialism, but *not that*. The need to lay down conceptual boundaries prohibiting the spilling over of particularities is thus fundamentally frustrated if the pragmatics of the concept's function aren't taken into account. Only in this sense of drawing up definitive boundaries was Fanon right to call the endeavor of defining colonialism utopian.

Political concepts are "contested under specific circumstances" (Haugaard, 2010, 423) and are enveloped in applying "agreed-upon and cherished notions in new contexts to render unseen injustices visible or identifying additional examples of virtue, with varying success" (Norberg, 2015, 657). What counts as colonialism can equally be contested, *but not indefinitely*, lest its function for intertemporally chaining patterns of injustice be lost. I will therefore operationalize the central pattern identified in the previous section. For the linguistically normative definition to do work for us, the central pattern needs to be maintained.[32] Heeding Fanon's warning, but inevitably going beyond him, the patterns of similarity running through the instances of colonialism are *c*lustered around this central similarity: the dynamic of colonizer/colonized in the institutional subversion of self-determination of one political community by another.[33] This core sets colonialism apart from other injustices.

However, what justifies operationalizing a central pattern of resemblance? Sally Haslanger's distinction between *operative* and *manifest* concepts provides a clue. The former is the determination of the application of the term, while the latter is what individuals think they mean with the term (Aytaç, 2021, 425). Regarding the construction of an operational definition (drawing a boundary around a concept), Wittgenstein (2009, §69) reminds us:

> We don't know the boundaries because none have been drawn. To repeat, we can draw a boundary – for a special purpose. Does it take this to make the concept usable? Not at all! Except perhaps for that special purpose.

The passage's liberating insight is that a "special purpose" is at hand. For determining how to interpret AI colonialism, a clearer partitioning of colonialism's conceptual orrery is needed than its everyday referencing provides. For the theorization to be viable and insightful, a boundary *must* be drawn. As Black (1962, 94) noted, if "rules are not yet firmly established in the language", such rules cannot produce "necessary statements in any strict sense of 'necessary'" and only "pragmatic considerations of convenience and usefulness" can be offered. Every analysis presupposes some things to stand fast, just as 'impossibility' is prefigured by rules that allow for certain possibilities and not others (Black, 1962, 140-141). The boundary I seek to draw doesn't make manifest usages of colonialism obsolete. It harnesses the central similarity for colonialism qua *operative* concept. Its intended use is showing if comparisons of AI with colonialism are (un)warranted "to the extent that their beliefs and utterances are [(not)] coherent with the operative" concept (to paraphrase Aytaç (2021, 425)). The need for analysis and the awareness of these limitations justifies operationalizing colonialism by laying down a conceptual scheme.

[32] The dictionary definition of colonialism this chapter started with also makes that clear. Standard parlance of colonialism wants to recognize that it is concerned with *policies*, so that consequently, in the absence of such policies, one cannot speak of colonialism. Examining the various attempts at defining the use of the concept of colonialism led to the uncomfortable conclusion that colonialism is fundamentally possible without such a policy being in place.

[33] Part of the family resemblance are the injustices that historically 'travelled together' with the (interstate) political domination and exploitation: foreign settling or taking of land, cultural imposition and violence. Especially the first two are strong proxies for the presence of colonialism, but not as central of a similarity as the institutional and intentional subversion of self-determination.

I propose a *mixed structure* is the right scheme: a combination of a necessary condition with "non-essential attributes like those found in a family resemblance structure" (Barrenechea & Castillo, 2019, 115-116). A mixed structure is a form of *bounded polysemy*: a family resemblance in which multiple combinations of features are sufficient, but in which "there is a limited number of combinations that are sufficient for concept membership" (Barrenechea & Castillo, 2019, 108). This acknowledges the family resemblance nature of colonialism while overcoming the shortcomings of the family resemblance of political concepts. The mixed structure is depicted in **Figure 1**.[34] Colonialism's wrongful effects are represented as points in the periphery; likewise for colonialism's constitutive wrongs and its linguistically normative conceptual core. A descriptive 'conceptual orbit' corresponds to an instance of colonialism, revolving around the core and one or more of colonialism's wrongful effects. The orrery's center is occupied by the following linguistically normative core:

> The intentional and institutional dehumanizing subversion of one political community's capacity for self-determination by another, such that the subverted community loses all its self-determining power and becomes a distinct periphery (colonized) dominated by a metropole (colonizer).[35]

Nota bene (1) the *roles* of colonized and colonizer in the intentional degradation of the colonized community by the colonizer; and (2) that the incorporation of institutionalization into the linguistically normative core gives colonialism its explanatory power: its specific *mode of institutionalization* within world capitalism is what ties its paradigmatic injustices together. What this means is that colonialism is a particular mode of institutionalization, *an organized activity*, and that its domination-type depends on intentional political agency.[36] I take this broad conceptualization of institution from De Pina Cabral (2011, 492), who writes that "Instituting is a process of shared intentionality carried out by persons who, being mutually constituted, are in the process of becoming singular persons". The erection of the colonial frame is therefore never accidental: it creates an organizational structure for making possible myriad forms of control and harm and the intentionality underpinning colonial domination is embedded into its institutionalization.

This framing obtains, because were the domination involved devoid of institutionality and intentionality, colonialism would lack in explanatory power, as this would construe colonialism's normativity as adventitious, while none of historical colonialism was accidental. Conversely, hesitation is justified regarding calling contemporary dominating injustices with a sufficient degree of accidentality to it 'colonial'. As §4 will show, the emphasis on intentional and institutional domination is important for pre-empting the misunderstanding of cases of domination or restrictions on autonomy as colonial. Counting as colonialism requires more, as the latter two can namely lack institutionalization, while still counting as an *invasion* on another's choice set, or lack institutionalization and/or intentionality, but nevertheless causing the *vitiation* of another's choice set (Pettit, 2013, 38-40). Therefore, we shouldn't call those latter cases 'colonial' in the sense of the operational definition's mixed structure.

The myriad of related injustices makes family resemblance an indispensable insight for understanding colonialism. But there are interconceptual differences as well: not all concepts are composed of

[34] **Figure 1** shows the examples of two conceptual orbits that have as similarity the dehumanizing subversion of self-determination, but are distinguished by, respectively the wrongful effects of exploitation and the taking of land, and cultural imposition and violence.

[35] For this semantic criterium, I shall sometimes use 'institutional subversion of self-determination' as a shorthand for this core.

[36] I am indebted to Uğur Aytaç for this insight.

noncontingent features only, at least, they are of no use to us if construed like this, especially political concepts. Following Haugaard's (2010, 427) description of power as a family resemblance concept, "usefulness" rather than "essence" is "our criterion of evaluation": there is no noncontingent essence of colonialism, but it's specifiable which aspects of colonialism are necessary for its linguistically normative and descriptive roles.

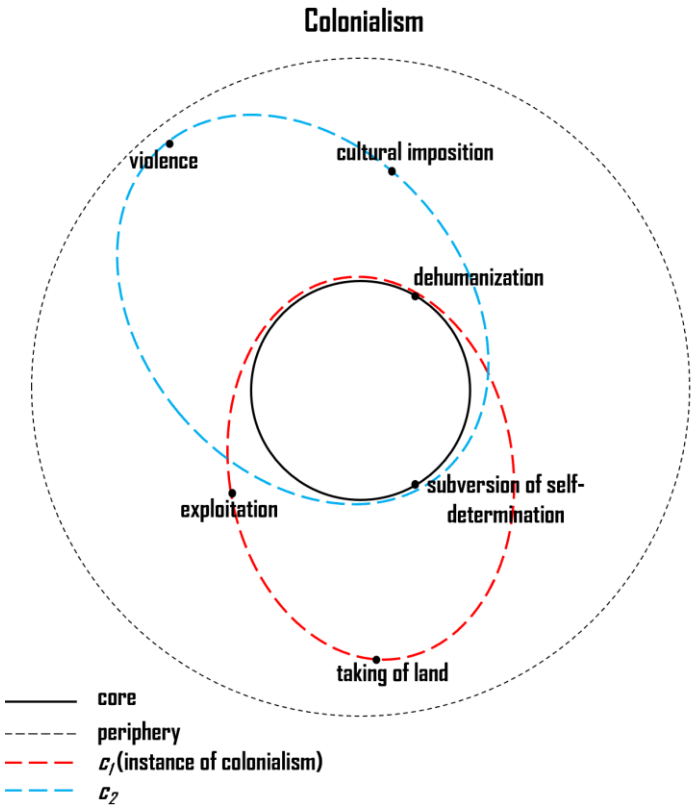

Figure 1 Schematic conceptual orrery of colonialism.

One counterpoint needs to be pre-empted to defend the mixed structure. Suppose contingent features of colonialism yield such a peculiar patterned, repeated phenomenon, that it seems functional to call it colonialism, although we don't observe any of the linguistically normative conditions or nothing that completely satisfies them. Family resemblance seems to allow for this. But the mixed structure addresses this problem: it stresses the importance of the wrongful effects, but is able to strip the descriptive semblance observed in this objection from any *normative semblance* belonging to the linguistically normative core. The mixed structure overcomes family resemblance's main limitation, namely that drawing no line at all renders some concepts theoretically useless: theories are "special purposes" that depend on conceptual boundaries. It recognizes that while many attributes are relevant, *not all attributes are equal* for colonialism's language-games. Some attributes are more central than others to the point that the concept can hardly fulfill one of its roles without it. The mixed structure employs this insight in order to bring the benefits of the meaning of a word as being constituted by its use to bear on definitional analyses. This requires that institutional dehumanizing subversion of self-determination is instantiated in tandem with at least one of colonialism's wrongful effects, which explains why historical appearances of colonialism concern many different wrongs.

To conclude this chapter, exploitation was situated as an *extension* of colonialism's linguistically normative core and colonialism qua norm was construed as *subclass* of domination. The mixed structure of colonialism was constructed and justified and multiple wrongful effects were identified. The 'institutional dehumanizing subversion of self-determination' is the shorthand for the semantic criterium tailored to the analysis of the similarities between AI injustices and colonialism. In §4, the target AI injustices can be projected as conceptual orbits onto colonialism's conceptual orrery, which elucidates whether there is (i) a valid instantiation of the constitutive wrong or (ii) only of some wrongful effect(s). If (i) is the case, then the target satisfies the success criteria for the literal interpretation. But if only (ii) is the case, then the literal interpretation fails: the resemblance to colonialism is *apparent* or at best *structural* and the metaphorical interpretation possibly applies.

3

The contours of AI colonialism

> Serfdom is no more, but the capitalist system is growing. And in the very heyday of emancipating ideas, […] the majority feeds, clothes, and defends the minority while remaining hungry, inadequately clad, and defenceless. Such an order of things can be made to fit in finely with any tendencies and currents of thought you like, because the art of enslaving is also gradually being cultivated. We no longer flog our servants in the stable, but we give to slavery refined forms, at least, we succeed in finding a justification for it in each particular case.
>
> A. P. Chekhov, *My Life*

Parlance of 'AI colonialism' was kickstarted by Abeba Birhane (2019, 391) who argued the 'algorithmic colonization of Africa' is taking place: "Algorithmic colonialism, driven by profit maximization at any cost, assumes that the human soul, behaviour, and action is raw material free for the taking." Africans are regarded as "merely data producing 'human natural resources'". Likewise, regarding AI and Africa, AI-expert Seydina Moussa Ndiaye says that "The biggest threat […] is colonization. We may end up with large multinationals in AI that will impose their solutions throughout the continent, leaving no room for creating local solutions. […] We could have Africa being used as a Guinea pig to test new solutions, and this could be a great, great threat for the continent" (UN News, 2024; *cf.* Kwet, 2019). These findings reverberate in the concept 'data colonialism': "the appropriation of human life" in the form of 'data relations,' digital interactions between humans and computer systems, that can be extracted via digital means in order to make profit from them (Couldry & Mejias, 2019, xiii). In an influential paper, Shakir Mohamed, Marie-Therese Png & William Isaac (2020) brought AI and colonialism together. They use the term "algorithmic coloniality" to extend data colonialism to AI systems.[37] Algorithms impact "the allocation of resources, human socio-cultural and political behaviour" and "coloniality features in algorithmic decision-making systems as they generate new labour markets, impact geopolitical power dynamics, and influence ethics discourse". The usage of algorithms to extract (and use extracted data) is compared with the problems of colonialism.

Since then, the indictments of AI colonialism have come from various sources, such as academics, consultancy bureaus and corporate and governmental researchers. The independent policy institute Chatham House calls AI "a new vector of colonial harm" that requires decolonization (Zimeta, 2023). "AI [is] created for "universal users", and "will neglect the languages and needs of individuals who do not fall into the category of 'white male,'" says Bobina Zulfa from Ugandan data research collective Pollicy (Harrisberg, 2023). Songyee Yoon (2023) compared AI D&D to historical colonialism's cultural imposition. AI systems such as ChatGPT are said to impose "a particular set of principles" and have the "capability to erase the cultural identity of non-western views through its ease of accessibility at a massive global scale". In

[37] Others that try to conceptualize the AI economy in terms of coloniality (the continuation of colonialism in today's world order) have supported criticisms of the data colonialism paradigm. The field isn't unequivocally convinced of the data colonialism thesis, as Muldoon & Wu (2023, 7) illustrate. The goals of both approaches and the aspect of the analogy with colonialism they find most important differs. While data colonialism wants to stress the comparison with colonialism in terms of 'extractive rationality', the coloniality of power framework stresses the colonial aspects the world order that AI systems and their deployment emerges from. In terms of how the historical process is reproduced today, their normative arguments thus differ.

considering these various claims, it is helpful to see how African jobs platform CA Global Finance (2023) defines AI colonialism:

> The term 'AI colonialism', or 'digital colonialism', doesn't refer to enslavement, land capture, or physical violence, but it does relate to the history of colonialism. Colonialism at its core was about using inhabitants of colonies for cheap labour, unfairly extracting their resources for their own financial gain, and maintaining economic and social inequality through various means of control.[38]

AI colonialism is framed as the *instantiation* of what colonialism was *about*. But further inquiry is needed as to *why* the previous comparisons are made: *explicit grounds* for the comparison are needed.

Others have warned AI engenders a 'technological' colonialism "of a small number of entities" (Dobrin, 2024). The case of Cyprus is illustrative. As media theorist Alexia Achilleos (2023) writes:

> Cyprus is under AI colonialism's occupation. A few Western AI companies have centralised all power in their hands and control the global AI infrastructure, pushing their own values and worldviews as universal, without taking into consideration the rich diversity of experiences of communities around the world. They also exploit the global population and the natural environment for their own profit, from mining raw materials in ex-colonies, used to build hardware that run their technologies, to extracting internet users' data to train AI models.

For Achilleos, AI colonialism thus involves the interplay of Euro-Western imposition, extractivism and the obstruction of indigenous people's self-legislation. From an economic perspective, think-tank director Ussal Sahbaz (2019) sees the risk of AI colonialism in the position of American and Chinese Big Tech companies in the markets of countries in the global South. Their developing markets are *dominated from the outside*. Former IBM chief AI officer Seth Dobrin (2024) echoes Sahbaz in the context of generative AI models. Their deployment by Western corporations leads to "the imposition of these entities' cultural values, biases, and societal norms on a global scale, often resulting in the marginalization of diverse cultural identities, exacerbating socioeconomic disparities, and potential homogenization of global cultures". While the targets are mostly previously colonized societies in the global South (or majority world), the actors aren't only ex-colonizers (e.g., European states) but also tech companies and formerly colonized nations that have started expansions of their own, such as China (Muldoon & Wu, 2023, 17).[39] Accessible AI, another civil society organization, surmised six aspects that capture how AI reproduces the dynamics of 'coloniality': algorithmic discrimination, the hidden energy and labor costs of AI, global power asymmetries reenforced by AI, the invisible labor behind AI datasets, beta-testing on marginalized groups and the usage of AI as a tool of oppression (Gausen, 2023).

Finally, the *AI Decolonial Manyfesto* (Krishnan, et al., 2023), signed by a manifold of academics, corroborates the idea that current AI injustices are intertwined with historical colonial injustices: "removing the echoes of coloniality in AI will require reparations for present and past material and epistemic injustice and

[38] CA Global Finance writes about African startups' resistance to AI colonialism and argues that AI is colonial because of (1) data's sources and (2) the types of data gathered.

[39] For concrete examples of China's interference in Africa's digital world, see Gwagwa & Mollema, 2023. Muldoon & Wu, in line with Birhane's work, also stress the ongoing "scramble for Africa" in which tech companies offer purportedly "free" and altruistic services in order to harness the power of users' data for profit" (Muldoon & Wu, 2023, 6; *cf.* Timcke, 2024).

dispossession". The mode of associating colonialism and AI in the *AI Decolonial Manyfesto* concerns 'the coloniality of power'. 'Decolonialism' is their point of departure.[40] The decolonial 'turn' in data and technology scholarship recognizes continuities between historical colonialism and modern Western societies, placing these societies' state-of-the-art technologies at the spearhead of a continuation of colonialism (Couldry & Mejias, 2023). Corresponding claims about how AI systems and AI D&D perpetuate, reproduce or cause colonial injustices have arisen. The goal of their analyses is to situate AI technologies and their deployment in the 'coloniality of power'.[41] This is supposed to explain why these technologies and the supporting economy constitute a contemporary form of colonialism.

The indictments thus deal with the catalysis and perpetuation of *existing* colonial injustices or the engendering of *contemporary* colonial injustices. The epigraph due to Chekhov concerned the substitution of exploitation through serfdom for exploitation through capitalism that occurred in nineteenth-century Russia. The authors that compare AI systems or D&D with colonialism follow an analogous schema: an injustice is thought to be over, but it again manifests itself in a different guise:

i. Injustice *X* is said to be over.
ii. But phenomenon *Y* engenders the same harms as *X*.
iii. Therefore *Y* constitutes a new form of *X*.

Filling in AI for *Y* and colonialism for *X*, this yields approximately:

[40] A short note on what is meant by 'decolonialism' might be helpful. Decolonialism concerns the political processes of undoing alien rule in the colonies ('decolonization') (Getachew, 2019, ch. 2-3) as well as the abolishment of colonialism's cognitive and epistemological components (Fanon, 2008; Mbembe, 2015). Decolonialism strives for the undoing of colonialism's harmful and/or unjust effects via changes to the colonized's relation to themselves, whether in terms of political self-determination and self-government (the undoing of what Stilz (2015, 4) calls the morally problematic "continuing political association between colony and metropole") or in cognitive or epistemic reconstitutions of identity (Mbembe, 2019). The decolonial picture of AI tries to incorporate the ecological, epistemic and politico-economic aspects of colonialism. Mbembe's (2022) critique of computational technology is an example of this, while the scholars working in the 'coloniality of power' tradition of Aníbal Quijano are another example (Ricuarte, 2019; 2022; Muldoon & Wu, 2023).

[41] For some, clarifying what the 'coloniality of power' might be needed. Decolonial scholars use the concept to challenge the "pretensions of universalism" inherent in the modern Euro-American world (Ndlovu-Gatsheni, 2019, 3). The term is due to Peruvian decolonial sociologist Quijano and describes "the current world-systems entanglement with historical colonialism" (Muldoon & Wu, 2023, 4), i.e., how distributions of power emerged and have abided by virtue of colonialism's subjugation of peoples. It conceives of power as a global system of control that originated in Europe's colonial expansions: "in the control of labor and its resources and products, it is the capitalist enterprise; in the control of sex and its resources and products, the bourgeois family; in the control of authority and its resources and products, the nation-state; in the control of intersubjectivity, Eurocentrism" (Quijano, 2000, 545). Ndlovu-Gatsheni (2019, 16) characterizes the reality of colonialism/coloniality as the reason for decoloniality's existence: the coloniality of power is a conceptual tool that "that decolonial theorists use to analyze the modern global cartography of power and how the modern world works". Its complements are the colonialities of *knowledge* and *being,* which are concerned with the colonial schemas inherent in "who generates which knowledge and for what purpose" and "the making of modern subjectivities and issues of human ontology" (Ndlovu-Gatsheni, 2019, 17). Coloniality's central idea is that practice of colonialism was fundamental to the construction of the modern world, not just in terms of explaining why Western countries were able to build empires and accumulate capital, but also in terms of how Western forms of life have come to dominate other cultural forms of life, in terms of the imposition of Western scientific knowledge and the creation of ideas of race.

i. Colonialism is said to be over.
ii. But the usage of AI (ML in particular), continues the injustices $i_1 \ldots i_n$ of colonialism.
iii. Therefore the usage of AI constitutes a new form of colonialism.[42]

This chapter is concerned with sketching the injustices $i_1 \ldots i_n$, i.e., the justifications of the comparisons between AI and colonialism. The corresponding research question is: *How are the harms involved with AI related to the harms involved with colonialism?*

The following provides descriptions of the various AI injustices and the reasons why they are colonial. My reconstruction yields that the comparisons of AI with colonialism come in six different forms – *extractivism*, *objectification, exploitation*, *cultural imposition*, *racialization* and *epistemic violence* – that ground why 'AI colonialism' mediates apprehensions of social reality in a way that *enriches* and *explains* rather than *obfuscates* and *confuses*. Serving these aims, the structure of this chapter is as follows. First, §3.1. discusses the extraction of natural and human resources for and by AI and the way AI systems objectify human behavior. Second, in §3.2. the worker and sales market exploitation involved in the 'techno-colonial market' of AI technologies are discussed. Third, in §3.3. the racialization, imposition of cultural values and norms and epistemic violence related to AI systems and D&D are examined. Finally, in §3.4. the contours of 'AI colonialism' are identified.

3.1. AI's extractivism and objectification

The first comparisons concern the indictments of AI systems and D&D with extractivism (§3.1.1.) and objectification (§3.1.2.). The AI economy is dependent on natural resource extraction in the global South, and historical colonialism's geopolitical schema is similar to AI D&D's geopolitical impact, while AI systems are said to objectify human behavior through datafication, the extraction of value from human behavior through degrading objectification.[43]

3.1.1. The AI economy's extractivism

The AI economy depends on the extraction of natural resources from the global South, mirroring the appropriation of resources and land by colonial potentates in 1500-1950. AI systems and the supporting infrastructure are embedded into a cyclical supply chain with ecological deficits, especially in the global South, such as energy consumption, mineral extraction and toxic waste pollution (Mbembe, 2022).[44]

[42] The dominant view within Western society on colonialism is that colonialism is supposed to be a thing of the past, ending with the *formal* decolonization of African, South American and Asian states. European states have in recent years for example been involved in 'making excuses' for their colonial history. In these apologies, colonialism is treated as something to be remembered, the consequences of which reverberate into the present, etc., but its *continuation* until today is acknowledged let alone defended. This is a controversial claim that scholars hailing mostly from the majority world have fiercely attacked (Mbembe, 2019; Quijano, 2000; Mignolo, 2020; Muldoon & Wu, 2023).

[43] In some identifications of AI colonialism, such as Birhane's (2019) discussion of the algorithmic colonization of Africa or Achille Mbembe's (2022) discussion of the techno-colonial market, these two grounds travel together: their work wants to show how Africa is occupied by foreign powers for the appropriation of natural resources and Africans are objectified and exploited qua digital resources.

[44] Jung (2023, 8) provides details on the waste production of the global tech industry: "Every year, the world discharges almost 50 million tonnes of e-waste. The overwhelming majority from the Global North ends up being

'Critical cartographies' of the ecosystems AI is premised on (Crawford & Joler, 2018) have put forward the claim that 'colonial extractivism' figures in it (Joler, 2020) and that this is central to AI's rationality (Pasquinelli & Joler, 2021). Vladan Joler and Kate Crawford (2018) analyzed the supply chain and deployment infrastructure of Amazon's Echo, a home interface technology depending on ML. They charted the Echo's ecology and documented the mineral extraction, milieu pollution, low-wage labor and data gathering practices it depends on.

Crawford (2021) took the cartographical project one step further in her *Atlas of AI,* where she explored the mineral extraction and water consumption needed for the building and maintenance of the data centers mainstream ML-training and deployment depend on and traced their impact across the globe. Due to LLMs like GPT, Gemini and Claude, ML-models' total energy consumption has skyrocketed. Already in 2018 "a single NLP [natural language processing] model produced more than 660,000 pounds of carbon dioxide emissions". Current LLMs are magnitudes larger than the models for which that estimation was determined; the resources required for these models has increased conforming to OpenAI's estimation "that since 2012, the amount of compute used to train a single AI model has increased by a factor of ten every year" (Crawford, 2021, 42).[45] In the past three years, AI's global impact on climate and natural reserves has only increased (Crawford, 2024). ML-models' training often depends on the 'compute' of hundreds of thousands of GPU's working a week long. They are used to compress training data size of the entire Internet to provide a new baseline for an LLM (Karpathy, 2023).

But what is *colonial* about AI D&D's consumption footprint and resource extraction? Crawford & Joler (2018) situate this extractivism in colonialism's tradition of natural resource appropriation. Citing postcolonial scholar Vandana Shiva, they write: "The 'enclosure' of biodiversity and knowledge is the final step in a series of enclosures that began with the rise of colonialism". Colonialism concerned the motive of enclosing and extracting as much resources as possible. AI D&D is its latest manifestation and a continuation of the same enclosing practice: "Traditional colonial practices of control over critical assets, trade routes, natural resources and exploitation of human labor are still deeply embedded in the contemporary supply chains" (Joler, 2020). The resulting argument is that AI's extractivism implies the *asymmetric colonial schema of Southern countries* (many of which are former colonies) *having to bear ecological disruptions while Northern countries* (many of which are former colonizers) *reap the economic benefits of AI*, with minimal ecological backlash. This variant compares AI with colonialism on the basis of a resemblance of historical colonialism's geopolitical schema to AI D&D's geopolitical impact: AI D&D's prerequisites *reinvigorate* colonialism.

But this isn't a singular conclusion. Mbembe (2022, 43) stressed the vast scale that sea bottoms are scraped for minerals and how African countries are mined to satisfy Big Tech's need for cobalt, copper, tin, manganese and nickel.[46] Conflict-free certifications for mining operations are "under question" (Crawford,

exported to the 'Majority World', from North America and Europe to Nigeria or Ghana, from Japan to China, from Singapore to India. Most of the waste ends up in landfills, where heavy metals such as lead, mercury, cadmium, and other toxins leak into the ground and contaminate the groundwater and the food chain." See also Muldoon & Wu (2023, 10), who report that "E-waste increased to 6.8 kg per capita in 2021, with long-term estimates predicting over 120 million metric tons of e-waste per year by 2050."

[45] Likewise, "At an organisational level, Google has released figures which show that machine learning accounts for 15% of the company's total energy consumption" (Muldoon & Wu, 2023, 10).

[46] Crawford also devotes specific attention to the lithium mining, for example in Bolivia, a country which, because of its lithium reserves, has suddenly become the stage of a power play between AI-powered companies. Other

2021, 43) now that is has become apparent that tech giants have little insight in the origin of their minerals because "suppliers are themselves linked downstream to tens of thousands of component manufacturers […] which are supplied by unknown numbers of traders that deal directly with both legal and illegal mining operation" (Crawford & Joler, 2018). Likewise, the full gravity of the environmental impact of these mining operations is also opaque to them.[47] But it is well-known that mining these minerals is toxic to workers and proceeds under lethal circumstances, such as is well-documented for tin mining in Indonesia (Crawford & Joler, 2019). James Muldoon and Boxi Wu (2023, 2; *cf.* Hao, 2022) conceptualize this as a 'colonial supply chain' that is embedded into the coloniality of power. Therefore AI D&D *extends* historical colonialism. They view AI D&D not as analogous to colonialism, but as the next episode in a "longstanding material and epistemological subordination of countries in the majority world from which resources are extracted and labour is exploited". Environmental costs are disproportionately shouldered by the countries that are mined for the AI economy's resources and examples of mine nationalization show how "the geopolitics of mineral extraction" is formed by "historical and ongoing colonial patterns of exploitation and dependencies" (Muldoon & Wu, 2023, 9-10).[48] "North-South systemic asymmetries" (Ricaurte, 2022, 727) are reinforced by AI.

Likewise, Maximillian Jung states 'digital capitalism is a *mine*, not a cloud' and that the extractivism of natural resources is colonial because it depends on "disposable lands and disposable people". Jung (2023, 3; 7-9) names three aspects of the careless 'colonial logic' of digital capitalism (*cf.* Mohamed, Png & Isaac, 2020, 10). First is digital capitalism's historical emergence out of colonial dependencies: the technological innovations of industrialization that gave rise to today's market economy were paid for with the capital taken from the colonies and are currently sustained by extractivism. Second is the treatment of these countries as *peripheral colonies*: "Nowadays, former colonised countries and peoples are treated first and foremost as resources that can be tapped into rather than connecting them in their own right". Third is that those who toil to extract resources receive subhuman treatment and are frequently subject to human rights abuses: the violence isn't as *overt* as it was in historical colonialism, but people are seriously harmed as part of the machinery of AI production (e.g., miners have to work unprotected under toxic conditions, digital workers are exposed to harmful content) (Jung, 2023, 5-8).

---

prominent lithium mining sites in the global South are "Congo, Mongolia, Indonesia" and in the global North Nevada and Western Australia (Crawford, 2021, 33). As several African scholars write: "Big Tech is not the saviour Africa needs to look up to and their presence in Africa is driven primarily by profits, monopoly and a rush to grab power more than anything else" (Eke, Wakunuma & Akintoye, 2023, 41). What is also notable is the amount of violence involved in suppressing resistance to the mining operations. More than 1733 environmental defenders where assassinated in the past ten years, defending their lands from being exploited (Jung, 2023, 7).

[47] Crawford & Joler cite natural resource strategist David Abraham: "Only 0.2 percent of the mined clay contains the valuable rare earth elements. This means that 99.8 percent of earth removed in rare earth mining is discarded as waste called "tailings" that are dumped back into the hills and streams," creating new pollutants like ammonium. In order to refine one tone of rare earth elements, "the Chinese Society of Rare Earths estimates that the process produces 75,000 liters of acidic water and one ton of radioactive residue.[…] Furthermore, mining and refining activities consume vast amount of water and generate large quantities of CO2 emissions. In 2009, China produced 95% of the world's supply of these elements." (Crawford & Joler, 2018).

[48] This claim is echoed by others. They "stress the importance of AI development being "for Africans by Africans" to ensure that colonial cycles of extraction by Western entities and historical dependence on foreign aid don't impede what could be a viable pathway towards economic freedom" (Eke, Wakunuma & Akintoye, 2023, 48).

To recapitulate, the first comparison of AI with colonialism is (i_1) *extractivism*. Harking back to the previous chapter's normative vocabulary, extractivism can be related to colonialism's land-taking (albeit without state usurpation) and exploitation (the instrumentalization of other communities for profit). Additionally, Jung's third point relates to dehumanization (in the way people from formerly colonized societies are treated as disposable workers), which §3.2. elaborates on. AI D&D's extractivism is *prima facie* a form of colonialism for extending historical colonial practices into the present and implying historical colonialism's geopolitical schema of extracting profit from foreign natural reserves.[49]

3.1.2. AI and the coloniality of objectification

Up next is objectification through datafication, where humans become 'digital subjects': resources to be appropriated in AI D&D. Human life is increasingly commodified in the form of data for AI systems, accumulating profit for the systems' developers and deployers. These data are the product of an information market about human beings that depends on their digital interactions, the human labeling of data instances, and the capturing of their behavioral traces in the digital realm. Instead of minerals, *online human behavior* is mined. This happens by means of algorithms underlying search engines, web page tracking, etc., and with all human beings leaving digital traces qualifying as resource, is global in scale. In historical colonialism, power, land, labor and resources where what colonial exploiters were interested in. But the extractivism of the digital economy has extended to new objects. The appropriation of land, labor and natural resources now includes *digital spaces* (as new form land), *data production* (as new form of labor) and *data relations* (as new resource to be mined) (Couldry & Mejias, 2019; Mbembe, 2022).

AI D&D is said to *objectify* human life like colonialism did. (i) Humans are now treated akin to how colonialism treated natural resources and (ii) the colonial project of profiting from the colonized's labor power is extended towards profiting from the colonized's digital relations among themselves by means of AI systems and digital platforms. Lambrechts, Sinha & Mosoetsa (2022, 11060) state that there is a ""mining" of people for data": "There are similarities between classic and digital colonization where, most notably, developed entities aim to enforce power and control over developing minorities". The objectification through datafication by algorithms is reminiscent of the colonial appropriation of human beings. This aligns with the thesis that AI D&D depends on a form of 'data colonialism' (the practice of the datafication of human life and behavior by means of algorithmic extraction (Couldry & Mejias, 2019)). On an economic level, these data are profited off by 'the market in digital futures' – a market driven by the datafication resulting from human interactions with "*organo-computational machines*", such as smartphones, laptops and anything with internet access and user identification (Mbembe, 2020; 2022). The market is realized by AI systems that predict consumers' next steps and wants in order to be able to play into them.

Micheal Kwet names the data extraction from humans in the global South a defining element of 'digital colonialism'. "By its very design," Kwet (2019, 13) holds, "Big Data violates privacy. […] When applied to

[49] During a talk called "AI & the Climate Crisis" organized by Dan McQuillan, speaker Sebastián Lehuedé responded to the author's question 'What is the importance of decolonizing AI for dealing with the environmental impact of AI?' that "there is skepticism about tying up the colonial with AI" but that "with the world being completely co-opted by companies" forms of resistance like decolonial AI are imperative for presenting "alternatives to the current global system" (Lehuedé, 2024). Lehuedé's talk named four points that makes AI colonial: the extraction of data and natural resources, the expansion of monopoly infrastructure without consent, the outsourcing of environmental damage and (his own addition) the imposition of techno-ecological visions (Lehuedé, 2024).

people, it collects personal and historical information to predict the future" and requires constant mass surveillance in order to prosper. Integral to AI D&D becomes the gathering of data, the infringement on people's privacy (whether legalized or not) and the perspective that human existence in the form of data is a resource whose value can be detached from the persons whose behaviors are harvested. It depends on the objectification of humans as resource to be exploited. Joler (2020) goes further and conceives of data production through interaction with interfaces of the digital sphere as inherently disrespectful: 'the forced labor of the twenty-first century'. Not only raw data are targets: existing datasets are as well. In 2015 DeepMind "got access to the health records of 1.6 million identifiable patients of Royal Free hospital" and this privatized public data through AI: "A dataset may still be publicly owned, but the meta-value of the data – the model created by it – is privately owned" (Crawford, 2021, 120).

To summarize: *humans are treated as means to an end in datafication of their digital traces*. Through the lens of colonialism's conceptual orrery, some important similarities become visible. 'Data colonialism' and 'digital colonialism' qua objectifying datafication seem to be premised on extending historical colonialism's extractivism to human behavior. This echoes colonialism's wrongful effects (symptoms), namely land-taking, exploitation and forms of violence done to the datafied party. Another is that between the degrading nature of objectification (humans being treated as means for making profit without their explicit consent) and colonialism's dehumanization. These similarities are further unpacked in §4's analysis of the literal interpretation of colonialism. The argument supporting the coloniality of (i_2) *objectification* is that AI systems extract value from datafying humans: their digital traces are algorithmically harvested through forms of surveillance, especially in the global South, where data protections are less secure. AI systems require these data in order to be finetuned to the populace and AI D&D depends on this process of objectification for the maximization of its utility both in terms of perfecting the systems' use cases and making profit.

3.2. AI's forms of exploitation

Jung brought the dehumanization of digital capitalism's laborers to the fore. His case concerned exploitation as part of the supply chain underlying AI. Purportedly, it is in the AI supply chain that workers' vulnerabilities are instrumentalized. This section will bring to bear the characterization of exploitation from §2.2. on examples of exploitation involved in AI D&D. This exploitation comes in two variants: exploitation (1) with *workers* as exploited and (2) with *markets* as exploited. In both, a colonizer is said to extract benefits from a colonized party. The argument supervening both is that the AI economy is 'colonizing' the working class through the exploitation of social groups' vulnerabilities and their objectification as means for profit (discussed in the previous section).[50] In what follows, specific attention is given to how AI D&D engenders exploitation. It has to be acknowledged beforehand that apart from the consequences of AI D&D, a continent such as Africa remains a source of "resources to the rest of the world, which often involves exploitation and violence" (Eke, Wakunuma & Akintoye, 2023), which should be regarded as in continuity with colonial times and not inherently dependent on the use of AI systems *per se*. However, accepting this does not negate that the forms of exploitation related to AI D&D have specific colonial dynamics.

[50] To be legitimately colonial, they should involve the institutional dehumanizing subversion of a community's self-determination from which the colonizer extracts benefits. At first glance this is plausible. §4 examines whether this holds on analysis.

3.2.1. AI and worker exploitation

Mohamed, Png & Isaac (2020, 9) introduced the term 'algorithmic exploitation' to conceptualize forms of exploitation that are specific to AI D&D: "the ways in which institutional actors and industries that surround algorithmic tools take advantage of (often already marginalised) people by unfair or unethical means".

A first example of algorithmic exploitation is the worker exploitation enabled by AI-powered platform applications in the global North. Delivery services, mostly focused on speed, use migrants as source of cheap labor and interact with them through smartphone apps with algorithmically determined performance targets. Algorithmic targets create harsh working conditions for deliverers that (have long) receive(d) no labor protections (Cant, 2023). They are economically and socially vulnerable in that they often have little alternative jobs available to them. This is instrumentalized by their employer so that they accept low wage and unfavorable, algorithmically determined and impersonal, working conditions. However, it is a tough question whether this exploitation is *colonial* in any significant sense, because migrants can be construed as a separate politically *peripheral* community, but this doesn't happen at large, to the entire 'migrant community' (but then this characterization would only work in the case of people who do not originate from the global North).

A second, more important example is the worker exploitation present in the outsourcing of data labeling practices to low-income countries in the global South (Williams, Miceli & Gebru, 2022; Crawford, 2021) and the usage of so-called 'ghost workers' for processing mentally harmful data (Mohamed, Png & Isaac, 2020; Perrigo, 2023). 'Ghost workers' from the global South are invisible for AI systems' users, but necessary for the production of data suitable for the practice of the supervised ML. Sometimes the labor is exercised by citizens instead of employed workers (such as in the case of Google's reCAPTCHA). Crawford & Joler (20218) describe this as "invisible, hidden labor, outsourced or crowdsourced, hidden behind interfaces and camouflaged within algorithmic processes".

Data labeling work has often been outsourced via platform services such as Amazon's Mechanical Turk, for example for the ImageNet dataset, the most important training dataset for object recognition (Crawford, 2021, 108). Outsourcing to low-wage countries (ghost work) yielded the recent and rightly notorious example of the labeling of sexually violent or graphic material by Kenyan workers for OpenAI's development of ChatGPT (Perrigo, 2023), which testifies how widespread this practice is. For current LLMs, the tactic of reinforcement learning through human feedback (RLHF) is deployed, where one has to kickstart the learning process with large datasets labeled by humans. Subsequent learning stages can then incorporate human forms of feedback, e.g., user inputs or 'red teaming' efforts by the developers themselves to disable undesirable responses. Other examples are readily available. Crawford (2021, 65) discusses the situation of workers at x.ai, which were "sometimes putting in fourteen-hour shifts of annotating emails in order to sustain the illusion that the service was automated and functioning 24/7. They couldn't leave at the end of the night until the queues of emails were finished". Muldoon & Wu (2023, 13) discuss the case of Sama, an "ethical AI" company, "claiming to have lifted 50,000 people out of poverty in East Africa through their business model." In actuality, their wages were "between US$1.32 per hour after tax and US$1.44 per hour", which is less than the US$1.52 a receptionist in Nairobi earns. But one must be aware that these aren't singular cases, but are representative of a new industry which is growing enormously fast: "data preparation tasks represent over 80% of the time consumed in most AI and machine learning projects, and that the market for third-party data labelling solutions is $150M in 2018, growing to over $1B by 2023" ((Tubaro, et al., 2020), cited in Muldoon & Wu, 2023, 12).

The forms of maltreatment, underpayment and abuse of workers constitute the AI economy's worker exploitation. The economic uncertainties and lack of alternatives lowers the threshold for accepting harmful and unsafe working conditions. This constitutes the *economic vulnerability* that is instrumentalized by Big Tech corporations in order to reduce the costs of their data labeling. The sector's riches illustrate better paid or safer alternatives could have been provided. Corporations' power over low-wage workers is thus exercised to extract a net benefit from them. Workers shouldn't be exposed to the abusive working conditions constituted by the exposure to harmful data; this may lead to trauma or emotional desensibilization, such as in, e.g., OpenAI's Kenya case or in that of Appen, a data production contractor with one million employees in 170 countries, with whom Alphabet recently ended a contract over "precarious labor conditions" (Taylor, 2024). In the briefing of the UK House of Lords Communications and Digital Select Committee, McQuillan (2023b) also explicitly named the data labeling economy supporting large language models as a "colonialised and exploitative outsourcing to the Global South".[51]

A final variant of algorithmic exploitation is the 'beta-testing' of algorithms in countries with fewer data protections. Beta-testing turns citizens into guineapigs for novel technologies. The paradigmatic example is that of Cambridge Analytica testing algorithmic tools during the Kenyan and Nigerian elections (Mohamed, Png & Isaac, 2020, 11). Eleonore Pauwels (2020, 12) writes:

> Relying on the aggressive, incendiary campaigns generated by PR companies like Cambridge Analytica, domestic political parties can exploit citizens' personal profiles and information networks for spreading rumours, targeted propaganda, hate speech, mis- and disinformation. The rationale behind such sophisticated disinformation architecture is to immerse citizens in an alternative, virtual reality where they themselves become producers of digital manipulation. In Africa, the capacity to manipulate populations and information is increasingly imposed through the 'Internet of Bodies and Minds'.

This caused 'information disorders' that aggravated existing tensions between ethnic communities over the pending elections, leading to sprees of intergroup violence. Preliminarily, beta-testing fits the operational definition of exploitation, since the citizens are informationally vulnerable (they had little means of discerning information from disinformation), which was taken advantage of for profit by Cambridge Analytica.

In short, the first type of exploitation associated with AI is the worker exploitation involved with data labeling, ghost workers and beta-testing: the economic vulnerability of workers and citizens is instrumentalized by corporations (acting as exploiters) for extracting profit, through chartering cheap labor and the further deployment of their AI systems.

3.2.2. AI and market exploitation

The other form of exploitation concerns the global South's *markets*. The exploitation of these markets by AI D&D is dubbed 'technological colonialism' (Dobrin, 2024; Sahbaz, 2019). The *sales markets* of the global South are placed in the role of the exploited. While this sounds like a function of *market power*, there is a case to be made for colonial market exploitation through AI.

The sales markets of the global South are economically vulnerable in that for example Chinese and American companies have capabilities to produce software and hardware for AI that these countries lack.

[51] The danger of letting the deployment of large language modes continue in this fashion over the next three years, according to McQuillan (2023b), is that the narrative of progress in AI resulting in prosperity masquerades for increasing job losses, precaritization, algorithmic violence and environmental impacts.

The asymmetry in technological capabilities is exploited by Chinese and American companies, who come to function as 'gatekeepers' for technological development, enabling them to exhibit control over public life (Gwagwa & Townsend, 2023). The reason the asymmetry between these markets exists, has to do with the global South's colonial history. The current international division of labor stems directly from past exploitation of former colonies and the extermination and domination of non-European societies in the centuries giving rise to modernity. The global North used them as deposit markets for their own produce. As Muldoon & Wu (2023, 11) argue, AI is placed in direct continuation of this international division of labor. AI D&D replicates the colonial division of labor because the West accommodates highly-paid AI labor, while "low paid and often precarious clerical, data entry, assembly and mining work" is outsourced. Muldoon & Wu (2023, 12) critique the invisibility of the labor that AI systems require and corroborate the picture of outsourcing degrading microwork on data to low-income countries in the majority world (currently mostly to India, Pakistan, Bangladesh, Indonesia and the Philippines). The market is economically vulnerable: it delivers resources and labor for less than the intrinsic value, but depends on the technology, infrastructure and software for AI to be returned. An example of this clear asymmetry is "Nigeria, one of the more technically developed countries in Africa, [that] imports 90% of all software used in the country. The local production of software is reduced to add-ons or extensions creation for mainstream packaged software" (Birhane, 2019, 396). As long as this vulnerability is exploited, the market is 'held hostage' in a detrimental situation. Therefore a diverse range of parties (Chatham House, CA Global Finance and multiple academics) have called for the decolonization of AI and the production of *AI from the global South*, to secure economic benefits and shift asymmetries.

Furthermore, the automation of production and social processes through AI systems aggravates the already unjust reallocation of the benefits AI yields. Now that LLMs require so much energy, data and infrastructure resources so that only the world's largest technology companies (and possibly the wealthiest countries aided by technological expertise) can produce them, the global South is placed in a systematically disadvantaged position. The "development of surplus value in economic, behavioural and epistemic dimensions", Paola Ricaurte (2022, 734; 737) writes, will not be allocated with those where resources and data were gathered. Automation thus serves the accumulation of wealth among capitalism's current beneficiaries and consolidates their dominion over the "global underclass" and their decision-making power over "the destinies of the poor".

According to Muldoon & Wu (2023, 13), on the level of the market as a whole, the degrading character of the work and the corresponding mental health risks show the embedding of the AI market in the *colonial matrix of power*, because the unwanted tasks necessary for AI welfare are cast upon those deemed disposable. This multifaceted disrespect mirrors historical colonialism's systemic disregard for the humanity of the colonized, who were treated as disposable laborers, slaves, cattle. The model of the coloniality of power applied to the data economy of the global South explains how data colonialism is embedded into a transnational process that "perpetuates exploitation" and "the extinction of alterity" with the dominant agents of the Western hemisphere as controllers of "processes that determine the collection, storage, access, analysis, and use of data", so Ricaurte (2019, 355) contends. With AI as new face of this data economy, I take her argument to extend to AI D&D as well.[52]

[52] Consider Birhane's testimony that: "Like traditional colonialism, wealthy individuals and corporations of the Global North continue to profit from some of the poorest communities except now it takes place under the guise of 'revolutionary' and 'state-of-the-art' technology" (Birhane, 2019, 402).

Taking a step back, it is worthwhile to see how the techno-colonial AI market at face value constitutes exploitation (§2.2.). The exploited and the exploiter are both *market actors* (global North *vs.* global South). This makes the 'relation' central to exploitation less direct than in personal cases. The plausibility of the applicability to these actors stands in need of further analysis; but then again, the colonial exploitation that would satisfy our definition of colonialism isn't about 'Mary exploiting John'. Rather, it's about schema's concerning collective actors like 'The Portuguese Empire exploiting 'the people of Senegal"; so let's grant the applicability for the time being. The market of the global South stands in a power asymmetry with that of the global North thanks to historical colonialism that has been transformed into an asymmetry in terms of AI capabilities. The economic vulnerability that is instrumentalized thus is the lack of AI capabilities, and the presence of people in need of work without alternatives. This benefits the global North as a whole through economic growth in general and direct profit for the companies involved. The case for market exploitation looks solid at first analysis.

To summarize: economically dependent sales markets are exploited thanks to global power asymmetries rooted in AI capabilities. The final analysis on whether this satisfies the literal interpretation of AI colonialism is undertaken in the next chapter.

3.3. AI's cultural imposition, epistemic violence and racialization

In relation to colonialism, cultural imposition, (epistemic) violence and racialization were categorized as wrongful effects, rather than defining criteria.[53] Because they remain important proxies for colonialism, it is worthwhile to investigate their intertwinement with AI systems and D&D as grounds for comparing AI with colonialism.

Cultural imposition is concerned with the expulsion of local cultural norms by the Western norms and values embedded into AI systems' training data and outputs (§3.3.2.). Epistemic violence is similar, but pertains to the displacement of non-Western standards of justification by those embedded in AI systems (§3.3.1.). Both aren't new to the global South. Racialization concerns the production and spreading of racial stereotypes rooted in colonialism: some scholars argue AI systems racialize by creating new forms of discrimination and reinforce societal racism through a propagation of 'whiteness' (§3.3.3.). Juxtaposing these grounds yields a picture of AI systems/D&D as tools for a global reinforcement of Western values and standards of justification, akin to how colonialism dethroned colonized peoples' ways of thinking and relating to their cultural heritage and identities (Okeja, 2022; Fanon, 2008; Ndlovu-Gatsheni, 2013).

3.3.1. Epistemic violence related to AI

The concept of epistemic injustice is due to Miranda Fricker (2007) and denotes a group of injustices that are specifically concerned with wrongs pertaining to the epistemological dimension of human interaction. Fricker distinguishes two kinds of epistemic injustice: *testimonial* injustice, where someone "is wronged specifically in her capacity as a knower" and *hermeneutical* injustice: "the injustice of having some significant area of one's social experience obscured from collective understanding owing to hermeneutical marginalization" (Fricker, 2007, 20; 158). Both concepts stand in need of further clarification before AI's 'colonial' epistemic violence can be discussed. Testimonial injustice is the case when one of the speakers in a discourse faces a

[53] Cultural imposition, racism, and forms of violence were explicitly named by some of the multi-wrong proceduralists surveyed in the previous chapter (Butt, Fanon, Mbembe). In particular, the discussions of Fanon's and Mbembe's work showed the consequences of colonialism, were at once psychological, cultural and epistemological.

credibility deficit. The testimonial injustices that are mainly of interest to the political philosopher are those which are systematic: structural patterns of prejudice that lead the credibility of a social group (e.g., women, immigrants, elderly people) to be structurally doubted (Fricker, 2007, 21; 28). Hermeneutical injustice is an equally structural notion. Its 'background condition', hermeneutical marginalization, precedes the 'eruption' of injustice when someone is actually "handicapped" by it. In short, hermeneutical injustice can occur in societies where some social groups' experiences are structurally marginalized, that is: where an inequality exists in how their experiences are made intelligible to society as a whole. Instead of those experiences being recognized and valued as valid perspectives, they are excluded from public discourse, institutions and the like (Fricker, 2007, 156-159).

The form of epistemic justice that concerns AI colonialism is '*intercultural* epistemic violence'. Thus conceived, 'epistemic violence' denotes the displacement of non-Western standpoints and knowledge to Western standards of justification and is closely related to 'cultural imposition'.

McQuillan, building explicitly on the work of Fricker, conceptualizes hermeneutical violence due to AI as the disqualification of persons in their capacity as knowers by AI systems. McQuillan (2022, 61) holds that hermeneutical violence is enacted by AI "because the complexity and opacity of AI-driven interventions are inherent barriers to any independent effort at comparable sense-making" which comes to "[overlay] already existing cultural and institutional systems of superiority". The related problems with AI systems like LLMs are that (a) LLMs' 'memorize' parts of their training data which they then regurgitate without referencing a source (Cyphert, 2024, 52-53);[54] (b) 'stochastic parroting' and fact hallucination (Bender, et al., 2021); and (c) hate speech production (McQuillan, 2023a). These shortcomings require *post hoc* elimination and their gravity is hidden from the average user[55] and contribute to the disproportionate harming of marginalized groups because (a) and (b) together constitute the emission of phrasings that are regurgitated from dominantly white and Western data as factual or could be hallucinated based upon that, with no direct way for users to check this. Marginalized and non-Western groups interacting with these technologies can therefore be presented with claims to facticity conflicting with their own culturally determined epistemic particulars. For (c), the case is that when hate speech seeps through the established filters, this is in most cases directed against those falling outside of the set of cisgender Western white males. For these reasons, McQuillan (2023a) denounces LLMs like ChatGPT: far from 'trustworthy AI', they end up destabilizing trust in human capacities of knowing in general, and that of already marginalized communities in particular. As McQuillan utilizes Fricker: LLMs amass a *surplus* of credibility and thereby create a credibility *deficit* amongst average users and culturally non-Western users in particular. When it's you against a personification of the Internet's knowledge, you have to be quite sure of your facts.

Yarden Katz (2020) genealogically accounts for the imposition of Western standards of justification. He argues AI has been imbued with forms of white universalism, militarism and racial bias ever since the 1950s. AI systems create false forms of objectivity; the "forgery" of "universal machine intelligence": 'views from nowhere' that universalize Western knowledge without justification. This capacity is increasingly recognized in mainstream computer science, with a prominent AI scholar like Micheal Woolridge (2023) naming AI's diffusion of existing forms of bias and toxicity explicitly in his Turing Lecture.

[54] The dependency on plagiarized training data that was outed in the *New York Times* lawsuit against OpenAI for the near-verbatim repeating of paywalled *New York Times* articles by ChatGPT attests to this (Grynbaum & Mac, 2023).

[55] An example is the Google AI that told users to eat rocks and put glue on pizza to make the cheese stick better, which near verbatim repeated an 11 year old Reddit comment (*cf. De Volkskrant,* 2024; Kelly, 2024).

An example of Katz's universalization of Western knowledge through AI is that AI systems are trained on predominantly Western datasets (the Internet). Based upon McQuillan and Birhane, Muldoon & Wu (2023, 14-15) corroborate this for more recent AI systems. They criticize LLMs' dependency on datasets that propagate Western epistemologies and spread a worldview that is "young, white, male and American". In terms of image recognition, the image databases many algorithms are trained on contain objects that figure in Western forms of life that are classified using categories supplied by the English language. The outputs of these AI systems reflect *those* ways of partitioning visions of the world and not others (Katz, 2020, 110-114). To illustrate, ImageNet, the most used image database, is comprised of 60% Western sources images (US: 45%) and only 2.2% from China and India (Muldoon & Wu, 2023, 14-15). Recently, Stanford University's Artificial Intelligence Index Report highlighted that these biases scale with size: the models' 'elicited toxicity' rises along with increases in the number of parameters (Maslej, et al., 2023).[56]

One could respond that the pre-LLM AIs that Katz is writing about just weren't that good. But that would be to mistake Katz's point. The general point isn't that AI could be corrected for racist or colonial imagery. It is that Western classification of images and opinionated patterns of text are regarded as giving *unmediated access* to the world that one can unproblematically extrapolate from. This imposition of Western standards of justification impairs interactions of AI systems with non-Western communities that diverge from them.[57] What began as a general injustice in Fricker's account, in McQuillan's and Katz's accounts becomes a form of violence in which cultural backgrounds are entangled with hermeneutical backgrounds.

Lastly, AI's epistemic violence is related to colonialism as *hegemonic knowledge production*: AI's embedding in the coloniality of knowledge, a facet of the coloniality of power where the force of colonialism im-/explicitly constrains knowledge that is 'invalid' from the colonizer's perspective. Simultaneously, it is a *re*production of the colonizer's existing knowledge through AI: alternative epistemologies are denied and ways of being and thinking are imposed (Ricaurte, 2019, 351). Muldoon & Wu (2023, 15-16) think the aspiration to universal validity is colonial because they regard it as the algorithmic reinforcement of the West's hegemony of values and knowledge: a perpetuation of the colonial matrix of power. Muldoon & Wu thereby complete the argument that AI systems are related to colonialism in their channelling of an imposition of a white/Western worldview.

In short, AI systems' colonial (i_4) *epistemic violence* is that their social powers 'reorder' the world in the colonizer's image. The reordering reinforces the hermeneutical marginalization of social groups from the

[56] What is specifically glaring in Katz's examples is the political dimension of the experience of marginalized social groups that AI systems cannot recognize in images they are presented with "Palestinian arriving at a checkpoint operated by Israeli soldiers" being classified as "a group of people standing on top of a snow covered slope" or "a 1960 photograph of Ruby Bridges, a six-year-old African American girl being accompanied to a desegregated school by U.S. marshals" being classified as, "a group of men standing next to each other" (Katz, 2020, 113).

[57] In *Responsible AI in Africa: Challenges and Opportunities* this epistemic effect and its colonial connotations are well illustrated: "AI applicability is synonymous with big tech companies from the Global North, therefore raising concerns around dependency in terms of technology know-how, capability, capacity as well as the inculcation of value-systems from the Global North to the Global South. This raises further questions around the potential and possibility of digital/neo-colonialism which can leave Africa grappling with the technology and not being able to understand fully or find solutions for challenges that result from AI as it is applied in different domains. Simply put, the needs of the Global North are different from those of the Global South, as such, it goes without saying that the application of AI on the African continent may be different in terms of the problems it intends to solve and subsequent benefits the technology will have" (Eke, Wakunuma & Akintoye, 2023, 185).

global South by systematically excluding their experiences, who are only marginally part of the training data, or not at all. The hermeneutical injustice "erupts" once the AI system's 'objective' output comes into conflict with a 'subjective' experience that is then *disqualified* instead of made socially intelligible. The analogy with colonialism is deepened by postcolonial theory on the displacement of non-Western 'ecologies of knowledge'. Boaventura de Sousa Santos (2014) introduced the term '*epistemicide*' to capture the eradication of indigenous forms of knowledge and sensemaking through Western colonialism and coloniality.[58] AI systems' 'view from nowhere' epistemically disqualifies knowers and inferiorizes non-Western epistemologies.

3.3.2. Cultural imposition through AI

Values embedded in AI systems can expulse (and conflict with) local norms and values if deployed in the context of different "value systems" than what they're developed for (Lambrechts, Sinha & Mosoetsa, 2022).[59] As Khan (2022) wrote:

> AI colonialism has long been bleeding into our everyday lives. If there is no reflection of the global south in AI technology, then what we consume is a wholly western point of view. The only differences between the colonial powers of the past and tech companies are that there is no physical invasion of our lands, just the hijacking of our minds.

This dystopic picture of "AI invading our minds" isn't without grounds, as Kwasi Wiredu (2002) has analyzed the consequences of something as basic as the imposition of a foreign language on the colonized. The imposition eclipsed forms of concept use prevalent in a local language (say, Akan) by categories denoting abstract concepts that are abundant in the imposed language (English). Language, as a carrier of thought, doesn't completely predetermine the expression of norms, values and concepts, but its use more readily fits the expression of the 'value system' that it has historically been used to convey (Okeja, 2022). As such it is understandable that contemporary African scholars of technology are worried:

> Gendered chatbots in communities in Africa […] could introduce and, in some cases, further complicate gender relations in a way and manner akin to the impact of colonialism on indigenous women's rights which were eroded in colonial and postcolonial societies. […] Already, several studies have highlighted the many ways technology today is a tool of neocolonialism reminiscent of colonial extractivist activities. AI-powered chatbots could therefore introduce and impose new forms of gendered expectations upon women (Eke, Wakunuma & Akintoye, 2023, 132).

Ugar (2023) uses the term 'value colonialism' for these instances of AI systems' influencing conceptions of cultural norms and values. What is clear and already a non-controversial position in technology scholarship is that values can be embedded in AI systems, but qua their relative autonomy and adaptivity "ensuring that the right values remain embodied in them would […] require[s] continuous human oversight and redesign"

[58] One such form of epistemicide is identified by Matteo Pasquinelli (2019), who writes that treating AI as a merely Western invention is itself a colonial judgment: "…claiming that abstract techniques of knowledge and artificial metalanguages belong uniquely to the modern industrial West is not only historically inaccurate but also an act and one of implicit *epistemic colonialism* towards cultures of other places and other times."

[59] Lambrechts, Sinha & Mosoetsa's framework of 'value sensitive algorithm design' wants to incorporate the idea that a culture's value system – "the set of social, moral, political, religious, aesthetic, and economic values that govern how a group of people live, work, and worship" – has to be adequately incorporated into the design of the algorithm before it can be deployed without the harms of cultural imposition applying (Lambrechts, Sinha & Mosoetsa, 2022, 11058).

(Van de Poel, 2020, 405). As a corollary, it is feared that culturally *unwanted* values can enter, while the *desired* values can slip out of these AI, with potentially harmful consequences only emerging in engagements with human interactants.

What makes this worry of cultural imposition reminiscent of colonialism justified? Kidd & Birhane (2023) argue that LLMs' underlying biases and their machinery's obscurity lead people to mistakenly depend on their outputs *as if they were true*. Drawing upon psychology of persuasion and child development, they argue that "People form stronger, longer-lasting beliefs when they receive information from agents that they judge to be confident and knowledgeable, starting in early childhood" and that AI agents come across as confident and knowledgeable. What's more, there is additional evidence that LLMs are capable of feigning the objectivity and impartiality required. As Anthropic CEO Dario Amodei disclosed, in-company research on their latest LLM Claude Opus revealed that it is almost as capable at single-shot persuasion of humans on political topics as average humans are (Klein, 2024). *Here cultural imposition overlays an epistemic injustice*. It is an epistemic problem with cultural ramifications: the threat constituted by the repetition with which LLMs can expose humans to informational biases. This gives substance to speculative claims like that of Yuval Noah Harari (2023), who fears generative AI systems' competency at fooling people. Harari's envisioned dangers are the ensuing vulnerability when people develop affective relationships with AI systems and their influence on affective stances towards marginalized groups. People form their beliefs on the limited pool of available information and LLMs' *souffler* of facts plays into this informational weakness. This might be a general shortcoming of human-AI interaction, but it underpins exogenous norms and values' disruptive power.[60]

The incorporation of non-Western cultural values into AI D&D is therefore seen as indispensable for these cultures' inclusion into the use and legislation of AI systems, which is a necessary countermeasure for combatting AI's cultural imposition and as a ground for the development of non-Western AI.[61]

To conclude, (i_5) *cultural imposition* is linked to colonialism's wrongful effect of imposing foreign norms, values and languages on the colonized. This interferes with native ways of sensemaking and identity formation. Following in Wiredu's footsteps, Okeja (2022, 205) has called colonialism a "conceptual adjustment program" that initialized normative shifts in colonized society, leading to forms of conceptual *loss*. The grounded fear, as the example of the AI-chatbot deployed in African cultures made clear, is that the current deployment of AI systems ingrained with Western norms and values (the content of the Internet) will

[60] The case of the 'information disorders' created by the algorithmic spreading of misinformation surrounding the Kenyan elections provides a 'proof of concept' (Pauwels, 2020) that LLMs can only but improve on in the domain of cultural conflict.

[61] For a defense of these views, see Mhlambi & Tiribelli (2023) and Gwagwa, Kazim and Hilliard (2022). Mhlambi & Tiribelli argue how AI ethics and governance frameworks are attuned to Western, individualistic social ontologies. This attunement makes them not directly applicable to the relational social ontologies that are characteristic of many non-Western societies, such as Sub-Saharan Africa. An example is the absence qua representation of the important African value of Ubuntu in the global AI regulation discourse. Ubuntu means 'a person is a person through other persons' and it implies the priority of relationality over individuality in identity and community formation. Gwagwa, Kazim and Hilliard (2022, 6) have identified the absence of Ubuntu in AI ethics and governance frameworks as a lacuna to be filled: "African values like Ubuntu, as well as the proposed moral ethics principles like harmony and consensus, have the potential to have a significant influence on AI ethics and policy, but this is only possible if the current domination of discourse by the Global North ceases. Greater inclusion would result in AI being more accessible and having less adverse effects for marginalized populations".

aggravate existing forms of conceptual loss or – in line with how AI creates new forms of racial and cultural boundaries – give rise to entirely new forms of conceptual adjustment and cultural expulsion.

3.3.3. AI as racializing technology

Racialization is the last facet of AI's apparent colonialism. According to McQuillan (2022, 136), AI is colonial because of the colonial history of pseudoscientific classification it inherits and "due to its racialized practices of exteriorization". It is a technology of racialization because it

> …lends itself to becoming a racial project, which is simultaneously an interpretation, representation, or explanation of racial identities and meanings, and an effort to organize and distribute resources (economic, political, cultural) along particular racial lines […].[62] AI will expand the operation of race by algorithmically diffracting it. […] AI will sort the distribution of resources and life chances according to statistical classifications, many of which will constitute new kinds of segregation at the level of groups and populations. It will not only reproduce race, gender and class but cross-multiply them with its own optimizing divisions (McQuillan, 2022, 65).

Not only do AI systems neatly reproduce their inputs reflecting the racist societies that produced them, they also *aggravate* their racism by creating new discriminatory division lines.

But McQuillan's view is prophetic rather than concrete. Annette Zimmermann and Chad Lee-Stronach (2022, 6-7) argue that the way algorithmic systems discriminate and divide involves the violation of the 'Like Cases Maxim' – "individuals with morally equivalent sets of features should receive the same treatment". When the Like Cases Maxim is violated by AI systems, this is procedurally wrong because relevant information with respect to structural injustices (pertaining to race and gender among others) is excluded. AI systems include the background configuration of any structural injustice, because this is reflected in their training data, which mirror society. However, this isn't explicitly reflected in the resulting algorithmic model; it *seems* as if like cases are treated alike, but they are treated differently and 'technofixes' that eliminate variables that code for unwanted differences ("*justice-relevant* features") don't work, because they don't eliminate the social reality of injustice that produced them (Zimmermann & Lee-Stronach, 2022, 13).

A paradigmatic example of the ways in which AI assembles forms of racism is due to Joy Buolamwini (2018), who established facial recognition systems' inability to detect the faces of African Americans and other people with a dark skin tone. Birhane (2022) notes that these results have subsequently informed the development of many AI systems. A more radical perspective is due to Katz, who speaks of AI's "artificial whiteness": the tethering of whiteness as *ideology* to AI. The reproduction of an ideology is a subtle affair in that it depends on two "epistemic forgeries". (1) It belongs to "the arsenal of a white supremacist social order" and (2) it "mirrors the nebulous and shifting form of whiteness as an ideology" (Katz, 2020, 153-155).

[62] McQuillan's technical motivation for this is as follows: "AI not only perpetuates racist discrimination but acts, at a deeper level, as a technology of racialization. The only way machine learning knows how to discriminate into output classes is by calculating distances, by determining some abstract metric of difference as a distance. The basis of machine learning is the construction of homophily – statistically induced connections to those who are allegedly 'similar', a forcing of closeness in data space that can be interpreted as a biologized attribute. AI segregates at a data level in the same way that racism itself segregates at a social level" (McQuillan, 2022, 65).

For the first, Katz explains that AI's whiteness exists in the technology serving the aims of a white race conquering nature and life itself (Katz, 2020, 156-159). Whiteness, according to Charles Mills (1997, 14), is "the 'consent' expected of the white citizens […], whether explicit or tacit, to the racial order, to white supremacy". Whiteness is "not really a color at all, but a set of power relations" (Mills, 1997, 127).[63] Following Mills, Katz explains the second claim by pointing to whiteness, as a racial category, being both real and unreal at the same time. Its effects are massive, but there is no biological grounding altogether. This is corroborated by others like Stephen Cave and Kanta Dihal (2020), who explain how AI systems become 'vehicles of whiteness'. For Cave & Dihal (2020, 12-14), AI's 'whiteness' consists in the coupling of imaginaries of the future to forms of white identity and the coalescence of attributes of whiteness to AI (intelligence, professionality and power).[64]

So far the intertwinement of AI with whiteness was conceptualized as ideological or pertaining to imagery, so the presence of whiteness in AI systems has to be unpacked further. Under Mills' conception of whiteness, so Katz maintains, whiteness is undetermined and shapeshifts "to maintain relations of power. AI reproduces this quality of whiteness" (Katz, 2020, 164). It can be seen that AI fulfills this function if one looks at the history of the idea of AI becoming the epistemological types that fit to the "imperial and capitalist agenda (as with the ideology of whiteness)" (Katz, 2020, 166). The first moment was the shift of symbolic AI into connectionism. After a short 'AI winter', connectionism was replaced by the narrative of Big Data and deep learning, that has now transformed into the paradigm of generative AI and LLMs.

Going one step further, McQuillan (2022) envisions AI systems as being cogs in the machine of a biopolitics of race. Following Mbembe's characterization of the result of racism's "techno-algorithmic turn" as "viral" and driven by computation (Mbembe, 2022, 113), McQuillan sees AI as a form of 'necropolitics': *the politics of who lives and who dies*. In earlier work, McQuillan (2015) already theorized the legal dimension of this algorithmic politics. He calls attention to the way in which algorithms are used to constitute 'algorithmic states of exception': suspensions of law that utilize algorithmic technologies in times of 'crisis'. A state of exception is "a space devoid of law where legal determinations are deactivated, especially that between public and private" (McQuillan, 2015, 6). The threat algorithms pose to citizens is their contribution to the constitution of power in states of exception and that they can have *force-of law* in this space devoid of law.

When the law's force is exercised on a racially divided society via AI systems, AI becomes "a racial project that, by assigning different values to parts of populations, will help to 'determine who lives, for how long, and under what conditions'", so McQuillan claims (2022, 65). In other words, AI systems become the means for enforcing the racial contract. Via practices of 'personal redlining' – the labeling or statistical accumulation of conditions that make one person less or more valuable than others –, assigning risk-scores, or the automatization of the economization of health care systems ('who is and who isn't worthy of treatment?'), AI is used to automatize the '*designation of disposability*'. A prominent example of these algorithmic management techniques that states use, is the usage of the 'System Risk Indication' (SyRI) by the Dutch tax agency. Instituted by law, SyRI was "an algorithm designed to identify potential social welfare fraud" (Rachovitsa & Johann, 2022) and played a role in the allocation of welfare child care and daycare budget

[63] As Charles Mills explains elsewhere (1998, 53), the intuitive picture of whiteness is contrasted with an intuitive picture of blackness: "we naturally think of whiteness as being associated with the experience of racial privilege and of blackness as being associated with the experience of racial oppression."

[64] AI systems not only "reflect the milieus in which they are produced", but, in the form of robots or chatbots for example, literally reflect the aesthetics of the white body. The latter is seen both in industry examples as well as in popular imaginations of AI systems (Cave & Dihal, 2020, 6-10).

surcharges and rent surcharges. However, it was found to infringe upon privacy rights as the algorithm didn't disclose subjects that their personal data were being processed. It systematically selected for people with migration backgrounds, resulting in the practice of institutional racism via the systematic discrimination and financial imperilment of already marginalized social groups (see also McQuillan, 2022, 58).

In short, various ways are emerging in which AI systems are used (through an implicitly or explicitly racial lens) in combination with biometric data or other personal identifiers for the management of various dimensions of human life (loans, health care, trustworthiness, etc.). To summarize, one identification of racialization was the AI system as vehicle of the white racial frame. For McQuillan, AI becomes a technology of racialization via the algorithmic creation and enforcement of the discrimination of racial categories concerning the management of society. The violation of the Like Case Maxim as conceptualized by Zimmermann & Lee-Stronach underpins how AI systems can transform these various facets of racism into concrete procedural wrongs, i.e., what, apart from being discriminatory, is wrong with the (obfuscated) division lines that AI is able to create. It has been noted before how the racial categories of today have been produced during colonialism and have given shape to modern states of the global North (Mills, 1997, 64). (i_6) *Racialization* thus is AI systems' inheritance of colonialism's notion of race, because of how they can function as tools in service of ideologies of whiteness and because of their classificatory computational basis that allows for the diffraction and creation of discriminatory lines.

3.4. Conclusion: the contours of 'AI colonialism'

At the start of this chapter, the comparison of AI systems/D&D with colonialism was schematized. Having surveyed many comparative grounds, the time has come to answer the question how the harms involved with AI are related to colonialism. Put differently: *what are the contours of AI colonialism*? Many claims regarding AI injustices have been expounded, but no author has undertaken a rigid analysis to determine whether the AI injustices *constitute* colonialism (the literal interpretation) or *mirror* colonialism's properties (the metaphorical interpretation). Let's therefore finally sketch AI colonialism's contours and summarize this chapter's findings.

Recall that applying the Chekhovian scheme to AI and colonialism yields: (1) colonialism is said to be over, but (2) the usage of AI continues injustices $i_1 \ldots i_n$ of colonialism and that therefore (3) the usage of AI constitutes a new form of colonialism. What remains to be done is to synthesize the previous sections to bring together the six grounds for comparing AI with colonialism. **Table 1** summarizes the injustices and the reasons for being colonial.

Prima facie, the term AI colonialism denotes the apparent similarities of AI systems and AI D&D with colonialism's extractivism (land-taking), objectification (dehumanization), exploitation, epistemic violence, cultural imposition and racialization.

But that being said, sorting out how far these comparisons go, is the topic of the analysis that takes place in the next chapter. There the case for the literal interpretation is made and assessed. The puzzle is whether the six grounds end at being *symptoms* of colonialism, while failing to be *criteria* for it; or if they *do* amount to be sufficient for colonialism *tout court*. The usage of the term can therefore be construed in two ways: as *unifying for these wide-ranging similarities* or as *selective for some of them in particular to point out a specific resemblance*. Whether literally or metaphorically construed, based on this contour sketch, it can already be concluded that AI colonialism is a rich term with moral, economic, cultural, epistemic and racial dimensions.

	AI injustice	**Colonial because…**
i_1	Extractivism	AI D&D's extractivism is a form of colonialism because it extends historical colonial practices into the present and implies historical colonialism's geopolitical schema of extracting profit from foreign natural reserves.
i_2	Objectification	AI D&D and AI systems are in continuity with colonialism in the way they thrive on the objectification of human beings as data resources. Value is extracted from this objectification via datafication. This communicates a form of disrespect that is analogous to the form of disrespect the colonized experienced.
i_3	Exploitation	(1) AI D&D takes advantage of the economic vulnerability of workers in order to extract profit from them, not only by chartering cheap labor, but also through the further deployment of the AI systems thus created. (2) Economically dependent sales markets are exploited by the global North's AI D&D thanks to global power asymmetries rooted in AI capabilities.
i_4	Epistemic violence	AI systems or AI D&D reinforce hermeneutic marginalization and displace non-Western 'ecologies of knowledge' through hegemonic knowledge production, akin to how colonialism has done.
i_5	Cultural imposition	AI systems or AI D&D interfere with native ways of sensemaking and identity formation, through the imposition of exogenous values, much like colonialism succeeded in doing.
i_6	Racialization	AI systems are colonial and stand in direct inheritance of colonialism's notion of race, because of how they can function as tools in service of ideologies of whiteness and because of their classificatory computational basis that allows for the diffraction and creation of discriminatory lines.

Table 1 Summary of the grounds of comparison of AI with colonialism.

4

Against the literal interpretation

> Every colonization or bloody adventure has involved precisely this—a given land is declared virgin or uncultivated, and its occupants are confined to reserves, the typical example of an enclosure and some gaping elsewhere.
>
> Achille Mbembe, *The Earthly Community*

Mbembe speaks not only of the colonization of America, Africa and Asia, but also of the 'technological' colonialism going on today. Should one take Mbembe's words literally or not? The previous chapter expounded the forms of colonialism/coloniality that the decolonial theorists, media scholars and diverse voices from civil society have been calling attention to over the past five years. As Couldry & Mejias (2024) indicate in a clear expression of the literal interpretation: "Coloniality – colonial thinking about how knowledge is produced and by who – is the clearest explanation for the sheer audacity of today's AI giants who see fit to treat everything humanity has produced to date as fodder for their large language and other models." Are the claims that make up the literal interpretation of AI colonialism justified? Is colonialism constituted by the extractivism, exploitation, objectification, cultural imposition, epistemic violence and racialization associated with AI? Is the majority world *literally* subject to colonization by way of AI systems or their development and deployment (D&D)?

The literal interpretation demands more of AI colonialism than the metaphorical or conceptual engineering interpretations. Under the literal interpretation, an adequate referent for 'AI colonialism' should instantiate colonialism *tout court*. For the metaphorical interpretation, establishing structural *similarities* between injustices associated with AI systems or D&D and colonialism's wrongs is enough.[65] But what the literal interpretation lays claim to is more grave than that. Firstly, that the extractivism, exploitation, objectification, cultural imposition, epistemic violence and racialization constitute *colonize* the global South and the treatment of citizens around the world as colonial resources. Secondly, the literal interpretation expects that one or more of these case actually constitutes an instance of colonialism.

This chapter concerns the validity of the literal interpretation that understands the AI injustices as *literal instantiations of colonialism's constitutive wrongs*. I will argue against this view. While similarities between AI qua global sociotechnical phenomenon and colonialism are evident, the AI injustices are in mismatch with the linguistically normative core of colonialism's operational definition. The conjunction of the two, if taken literally, so I argue, misconstrues the concept of colonialism. To undertake this analysis, §2's operational

[65] The conceptual engineering interpretations aren't discussed until §5, because their obtainment is distinct from the question whether to interpret AI colonialism literally or metaphorically. As a reminder, under the first variant of the conceptual engineering interpretation, use of the concept AI colonialism seeks to ameliorate – that is: to extend, shift or constrain – the use of 'colonialism' from its operative use *C* towards a better use *C**. The second variant of the conceptual engineering interpretation seeks independent lexical expansion: the introduction of a new term 'AI colonialism' with a new meaning *M*. Requisites for the conceptual engineering interpretation are that AI colonialism either (i) captures the meaning of some of the most important aspects of AI systems or D&D partly but shows colonialism is a defective concept (*amelioration*) (Haslanger, 2020); or (ii) diverges from colonialism altogether with the introduction of a meaning that differs from colonialism's meaning (*lexical expansion*) (Capellen, 2020).

definitions of domination, exploitation and colonialism are recycled. Recall that operational definitions "ultimately are useful only if they come close to adequate definition" (Pitkin, 1972, 275), but that colonialism's mixed structure was developed with the particular task of analyzing AI colonialism in mind. The *objects* of the analysis are the six AI injustices described in the previous chapter.

In the linguistic analysis of political concepts, one use of a concept is weighed against other possible uses. Is it mistaken how the authors surveyed in the previous chapter use the terms 'colonialism' or 'colonial' when describing AI? In cases like this, one can give reasons for preferring one conception of 'AI colonialism' over others. In short: what can one do with *this* conception that one cannot do with *that* conception? And: in what way is *this* conception confused where *that* conception isn't (Pitkin, 1972, 186)? The intellectual humility with which the attempts at comparing AI and colonialism should be approached is that if they are symptomatically deviant uses of the word colonialism with respect to AI, it is important to meet those deviations half-way and understand *why* they stray from the pattern as they do, before condemning them for doing so (Pitkin, 1972, 313). In short: the reasons I will present for not following the literal interpretation have to do with the AI injustices falling short of colonialism's linguistically normative core. Recall that §2 concluded with sketching the mixed structure of colonialism: colonialism is the case when its linguistically normative core is instantiated, in conjunction with any one of its wrongful effects. This core, which is the target I hold the AI injustices should satisfy, was:

> The *intentional and institutional dehumanizing subversion of one political community's capacity for self-determination by another*, such that the subverted community loses all its self-determining power and becomes a distinct periphery (colonized) dominated by a metropole (colonizer).

The justification for construing this core lay in the features taken from the literature on defining colonialism: intentionality, institutionality, domination in the sense of subverting political self-determination and disrespectfulness. This linguistically normative conceptual boundary delimits the *proper* operative concept of colonialism, i.e., the operative concept that corresponds to colonialism's function as rule for judging cases. The definition is employed for showing if parlance of AI colonialism is (un)warranted insofar as it is (in)coherent with it. Let this be our yardstick.[66]

The distinction regarding interpretative objects – AI *systems* (the technological entities themselves) or AI *D&D* (the social processes and economy of AI development and deployment) – takes center stage now, because they yield two delineations of the literal interpretation that require separate discussion. The first delineation is discussed in §4.1., while the second delineation is the subject of §4.2.. For both, I argue that the literal interpretation isn't justified. Finally, I consider three objections to this conclusion the *drift of language objection*, the *racial exploitation contract objection*; and the *Cortés disanalogy*.

4.1. The colonialism of AI systems

Not all of the purportedly 'colonial' injustices apply to AI *systems* (see **Table 2**). Would these injustices individually satisfy the definition of colonialism? This question should be answered negatively, because they fall short of colonialism *tout court*. I develop an argument for why AI systems themselves cannot be proper

[66] It is worthwhile to note that, when considering the charge of AI colonialism taken literally, a reflexivity with respect to the 'drift of language' is required. A possibility that should be recognized is that this operational definition reflects a *historical use* of the concept of colonialism, where, in fact, after processes of formal decolonization and the rise of neocolonialism and neoliberalism, the *criteria* for the use of colonialism have changed (Pitkin, 1972, 196): its use has 'drifted' towards another meaning. More on this in O_1: the *drift of language objection* (see §4.3.).

colonizers: the Cortés analogy. Subsequently, I show how this argument translates to the domains of racialization, cultural imposition, epistemic violence, and objectification.

	AI systems' injustice	**Colonial because…**
i_2	Objectification	AI *systems* are in continuity with colonialism in the way they objectify human beings as data resources. Value is extracted from this objectification via datafication. This communicates a form of disrespect that is analogous to the form of disrespect the colonized experienced.
i_4	Epistemic violence	AI *systems* reinforce hermeneutic marginalization and displace non-Western 'ecologies of knowledge' through hegemonic knowledge production, akin to how colonialism has done.
i_5	Cultural imposition	AI *systems* interfere with native ways of sensemaking and identity formation, through the imposition of exogenous values, much like colonialism succeeded in doing.
i_6	Racialization	AI *systems* are colonial and stand in direct inheritance of colonialism's notion of race, because of how they can function as tools in service of ideologies of whiteness and because of their classificatory computational basis that allows for the diffraction and creation of discriminatory lines.

Table 2 The grounds of comparison of AI ***systems*** *with colonialism.*

First, some point-wise stage setting is needed. Firstly, the literal interpretation concerns colonialism qua semantic criterium that depends on a relation between two actors, where the colonizer is responsible for the colonization of the colonized. The colonizer was often a collective actor composed of multiple agents. This collective actor can rightly be deemed responsible for the harm caused by the *organized* behavior of its members (Pettit, 2007, 172; 193-197). Secondly, the question whether *humankind* or *their technologies* are responsible for harmful consequences of using technologies is an old one. Unable to sketch its history here due to spatial constraints on this text, it can still safely be asserted that the question was already old when Karl Marx (2015, 620-621) wrote that:

> Machinery inserts itself to replace labour only where there is an overflow of labour powers. Only in the imagination of economists does it leap to the aid of the individual worker. It can be effective only with masses of workers, whose concentration relative to capital is one of its historic presuppositions, as we have seen. It enters not in order to replace labour power where this is lacking, but rather in order to reduce massively available labour power to its necessary measure. Machinery enters only where labour capacity is on hand in masses.

Following this representative line of thought, AI systems, which are attempts at automizing human intelligence and action, are above all tools of *optimization*, e.g., of speed, accuracy, human likeness, etc. Many of its applications reflect this, like question answering, theorem proving, image recognition and generation, document retrieval, recidivism prediction, self-driving vehicles, chip manufacturing, data analysis, vaccine testing, protein folding, etc. Thirdly, for optimization machinery to be literally colonial, it has to institutionally dehumanize a community and subvert that community's self-determination (by the literal interpretation's success criteria). Integral parts of this shorthand for colonialism's core are the conceptual roles of colonized and colonizer. Some actor *X* (colonizer) institutionally exerts its power over a community *Y* such that *Y*'s self-determination is subverted and *Y* is dehumanized (colonized). As such, *X* colonially dominates *Y*. It is

hard to see how AI systems themselves (*pace* artificial general superintelligence) could be substituted for *X* here, apart from as *part* of *X*'s *automatization of the process of colonizing Y*.

Building upon these three points, I propose that AI systems fall short of the intentionality requirement for *X*. Contrary to toasters, hammers, thermometers and other simple tools, AI systems are no 'mere' artefacts. Sometimes AI systems execute self-sufficient processes without a human 'in the loop'. Scholars have been debating whether this amounts to a 'responsibility gap' for two decades now (Matthias, 2004), but what is clear is that even when the human is out of the supervision loop of an AI system, society has an incentive to hold the party that built and/or activated the system responsible for the harms inextricably tied to its use. This is a fruitful line of reasoning for dealing with autonomous weapon systems (Suchman and Weber, 2016),[67] but it isn't for the ML-cases we're concerned with, because these applications are mostly concerned with reactive tasks part of human-computer interaction systems. On the contrary, autonomous weapon systems are more assertive and their outputs have directly moral consequences. But still: in the case of LLM-powered chatbots, where their creators have no full oversight over their linguistic potential and interactions with users, software companies are the targets for liability when the chatbots malfunction, not the chatbots 'themselves'. Following Daniel Dennett (2017), responsibility assignment requires *comprehension* of the action, not just the *competence* of being able to execute it. AI systems are paradigmatic cases of Dennett's slogan "competence without comprehension". AI systems are part of bringing a harm about, but are *vehicles* of intentionality, rather than *originators* of it. But even under the supposition of AI systems being autonomous, other reasons appear why they cannot be a complete substitute for *X* in the sense of *the type of actor* this substitution demand.

I now develop an analogy that applies the previous paragraphss line of reasoning to historical colonialism. The Spanish conqueror Hernán Cortés "landed on the beaches of what will be Veracruz" (Calvino, 2023, 264) and started the Spanish colonization of America. One of the peoples his army colonized and eradicated were the Aztecs, reigned by emperor Montezuma. Well, *what* colonized the Aztecs? Was it Cortés, his men, their swords, muskets, crossbows, boats or horses? Was it the contagious diseases they spread? Or was it not Cortés as a whole, but only 'the cunning' with which he planted himself in the idea of the battle of the gods in Montezuma's mind (Calvino, 2023, 274)? Affirmation of any of the options clearly is beyond the bounds of sense: under normal circumstances, with 'Cortés colonized Mexico' one means that the Spanish Empire (collective actor), on whose behalf Cortés and his men acted as *conquistadores* (member agents), took control of the region and displaced/eradicated its people (note how description and norm co-occur here). What this makes clear is that AI systems themselves are unfit for the conceptual *role* of colonizer in the subversion of self-determination of a community. Their only role in making colonization happen is as a *tool* employed by the colonizer – as Cortés employed sword and cunning. I shall recycle this idea under the name of the 'Cortés analogy'.

[67] A recent example of this is Israel's 'Lavender' AI system that "scans data regarding every Gazan in the IDF's database and, using this set of characteristics, generates a score from 1 to 100. The higher the number, the more likely that individual is to be a member of Hamas, according to the set of characteristics the AI produced. Lavender outputs these names onto a kill list. Then, after a brief check to confirm that a target is male, commanders turn the name over to additional tracking technologies, ordering the air force to bomb the target once their surveillance technology indicates that he is at home" (Weirich, 2024). Here decision-making over life and death is outsourced to the statistical workings of the AI system, 'relieving' soldiers of the direct responsibility of choosing who to target. But this moral shroud of the autonomy of the system's decision-making does not allow the Israelian military to shirk responsibility for the killings.

The Cortés analogy disqualifies the possibility of AI systems *literally* colonizing. A musket, autonomous drone, or optimization algorithm could be used in a colonial conquest or extractivist endeavor, but that wouldn't make the technology itself colonial.[68] We now understand better what the requirement of counting as 'full colonizers' means for satisfying colonialism's linguistically normative core. Only full colonizers like the Spanish Empire are suitable candidates for *X*, because they can act institutionally, i.e. as an intentionally organized system of behavior, and can form dehumanizing and dominating relationships with some *Y*, the colonized political community.

Having sketched the Cortés analogy, I shall promptly apply it to the AI injustices pertaining to AI systems to show that the literal interpretation cannot succeed.

First consider racialization. In racialization the AI system becomes a vehicle of its training data's racism (McQuillan, 2022; Katz, 2020). The idea of race that underpins both training data and society was fabricated during the colonialism lying at their historical roots (Mills, 1997). By the Cortés analogy, regardless of the importance of race for understanding modern society, AI systems' refraction racialized data, doesn't warrant describing them as literal colonial actors. To illustrate, for identifying the procedural injustice of AI systems' discrimination, Zimmermann & Lee-Stronach's (2022) formulation of the violation of the Like Cases Maxim was used. The violation of the Like Cases Maxim isn't an injustice due to the AI system itself. Rather, the idea of *doxastic negligence* places the responsibility directly with the *operator* of the AI system. Consider Zimmermann & Lee-Stronach's (2022, 19) definition of doxastic negligence:

> *A* is doxastically negligent if *A*, purely on the basis of an algorithmic output concerning *B*, adopts a belief about what kind of treatment of *B* is warranted.

So the agent that directly depends on or benefits from the AI system's functioning is the actor to which doxastic negligence rightly applies, *not* to the AI system. The procedural injustice wouldn't pertain to AI without the algorithm 'hiding' the structural injustice in seemingly objective outputs. But the injustice is only realized once the algorithmic output is *approved of, accepted or acted upon*. An AI system literally instantiating colonialism through racialization is therefore ruled out by the fact that an AI system is embedded into an organized system of behavior for which human agents are responsible. The *AI system* isn't in doxastic negligence, *its operator is*, which is a situated application of the Cortés analogy, because it shows how, in the context of racialization, the AI system doesn't fulfill the requirements for being responsible for a certain outcome, just like one cannot blame Cortés' sword for the consequences of his conquest, pledged to his ruler. AI systems are better described as *mirrors of a society founded upon colonialism* than as instances of a new colonialism themselves.

Secondly, consider cultural imposition and epistemic violence in tandem. In cultural imposition AI systems interfere with native ways of sensemaking and identity formation, and in epistemic violence AI systems

[68] As a further corroboration of the Cortés analogy, consider the case of algorithmic surveillance. In algorithmic surveillance, facial recognition technologies are the enablers of the repression of minorities. Not only does the repression of minorities invoke associations with colonialism, statistical management of subjects that were denied citizenship was developed under colonial rule. Even though this shows statistics is used for colonial purposes, this doesn't make 'statistics itself' an instantiation of the subversion of self-determination. Likewise, AI systems for biometric identification are excellent tools for the colonial management of colonized subjects and one could argue that in the case of Israel's usage of facial recognition systems to survey Palestinians and Israelis alike (Amnesty International, 2023), Palestinians are a colonized community. To be specific, Palestinians are a non-assimilated community whose self-determination is subverted. However, in that case, it is the *government of Israel* that colonizes, not the facial recognition technologies that are (admittedly) organs in the colonial organism.

displace non-Western 'ecologies of knowledge'. Following Pettit (2013), I recognize that both can constrain social groups' autonomy by way of interference with their cultural and epistemic cognitive resources. In line with how this was documented for colonialism (Okeja, 2022; Wiredu, 2002), I grant that this *could* lead to cases of cultural and epistemic conceptual loss. The conceptual loss could have the restriction of choice menus as consequence. But for both cases we can again not follow through with identifying them as instances of colonialism. AI systems may be the media of interference with a community's autonomy, but they aren't to be conceptualized as *interferers* here. Two lines of reasoning support this. First, harking back to the Cortés analogy, we cannot indict the AI system for the consequences of *the interaction with* the AI system. Scapegoating AI systems enables rather than closes the responsibility gap that AI's semi-autonomy gives rise to. Second, while AI system's presentation of a community with exogenous cultural and epistemic perspectives as if they were right and true may excite conflicts with their own values or standards of justification, this doesn't constitute the institutional subversion of the self-determination of a political community. As will be developed in more detail in the next section, these new AI injustices, while foreseeable once awareness of them is raised, are accidental *vitiations of resources for self-determination*, rather than coherent institutional attempts at undermining political communities' self-determination. Of course, Pauwels (2020) showed that Cambridge Analytica succeeded in disrupting election processes in Sub-Saharan Africa, but this incidental case of domination gives us no carte blanche for deducing the presence of the capacity of structural and institutional domination that colonialism's linguistically normative core requires. But more importantly: this case illustrates that *collective actors* subvert political community's self-determination in the sense that concerns colonialism, *not* sociotechnical interaction nodes like AI systems.

Furthermore, the conditions for colonial actors that the Cortés analogy presents us with, are morally relevant. Without distinguishing, e.g., how the Spanish Empire was a colonial actor from how Spanish horses were not (even though the latter were part of colonialism's causal network) responsibility for colonialism cannot be rightly allocated. Applied to AI systems, if we treat AI systems as originators of intention, their deployers and developers could evade responsibility for the racial, epistemic and cultural effects of their systems. Until AI systems come to qualify as legal persons and originators of institutions and intent in their own right, we cannot accept this conclusion.

Turning now to objectification via datafication, here AI systems capture human behavior. Objectification is a facet of the extractivism of human data resources, but it's only by virtue of ML processing and data collection systems that behavior becomes a 'virgin' resource. In datafication, humans are instrumentalized, as the resulting data enables agencies to perfect their advertising and sales algorithms or sell the data on the market for future behavior. Much like colonial actors infringed upon native rights to natural resources, mined them and profited from them, human rights are now being infringed upon and behavior is mined and profited off. Like exploitation, objectification *disrespects* who it instrumentalizes. Therefore what's wrong with datafication can insightfully be construed as *exploitation*. Applying §2.2.'s definition: humans are instrumentalized, treated as means to an end (profit), and this plays into their lack of legal protections or inescapable vulnerabilities, which are due to using digital interfaces. But even if this modelling of datafication as exploitation is granted, this is insufficient for colonialism. For exploitation to imply colonialism, it is required that (i) *AI systems* dominate and disrespect in an intentional and institutionalized manner and (ii) that this is inevitably part of datafication. As the reader might expect, I see problems with both conditions.

Starting with the first, under the Cortés analogy, AI systems cannot be viewed as colonial *actors,* since it is only the organizations employing them (Meta, Alphabet, the government of Zimbabwe, etc.) that could hope to satisfy the conditions for *X* in '*X* colonized *Y*'. AI systems therefore can't exploit when they

automatically register data about users – partly because they couldn't have done so otherwise. While it may be the case that, much like Aytaç (2024) shows for social media platforms, engaging with an AI system may place me in a relation where another 'exercises power-over' me, it is rather the collective actor forcing me into a (tacit) engagement with the AI system that overpowers me. Also, the modelling of objectification as exploitation also suffers from susceptibility to the Cortés analogy. As per §2.2., the exploitation that is of interest for colonialism is the systematic relation between *agents A* and *B* such that *A* dominates and extracts benefit from the instrumentalization of *B*'s vulnerability. Just like the AI system by itself cannot fulfill the role of *X* in '*X* colonized *Y*', the Cortés analogy yields that it also falls short of the role of *A* in '*A* exploited *B*'.

Secondly, datafication is not generally inevitable and often involves users' own choices. Datafication could arise when I become a data subject in interaction with a digital interface. However, any overpowering of myself in this relationship isn't due to AI systems *per se*, but rather due to a multiplicity of related collective actors and institutions: I may not have access to this digital service without using a smartphone, some data tracking is necessary to use the service, or the privacy laws that apply to me don't prohibit this type of data tracking. Thereby my choice set is narrowed because (a) some choices are being predetermined without consulting my preferences, but I'm not generally *exploited*, unless I'm (b) vulnerably dependent on the digital service. For many digital services, such as social media, online games, etc., (b) doesn't apply since I could shape a fulfilled life without them. But other cases, such as the increasing digitalization of banking and buying goods, that, depending on my own contingencies, *do* make me vulnerably dependent on digital services in an inevitable way. However, we can conclude that the requirements of colonization aren't inherently part of datafication and that when there are overpowering relationships involved in datafication, this isn't due to AI systems.

Finally, what happens if one considers these injustices taken together? Would the literal interpretation *then* be warranted? I argue not, because the conjunction of these cases would only satisfy the conditions set by colonialism's mixed structure if they would together be *sufficient* for AI systems counting as full colonizers, institutionally subverting self-determination. But invocation of the Cortés analogy renders this possibility implausible as well. After all, the AI injustices under consideration aren't all effected and orchestrated by a unified 'rogue AI'. As it stands, these disparate effects relate to the global development of a new type of technology, and even if we take all the AI systems together, they don't form a 'distinct political community' – calling them that equals committing a category mistake. But could we acknowledge their disparity while still summing the injustices' harms and attributing them to *AI systems in general*? Does this 'the sum is more than its parts' line of reasoning vindicate the view that AI systems literally colonize? No, because this still misconstrues the concept of colonialism: as the Cortés analogy shows, there is no colonization *without a suitable colonizer*.

So to summarize, using the Cortés analogy I showed that AI systems are insufficient for the conceptual role of colonizer required by colonialism's linguistically normative core. The literal interpretation is therefore unwarranted for AI systems. With the systems delineation of the literal interpretation disqualified, I now turn to the AI D&D delimitation.

4.2. The colonialism of AI's development and deployment

Because AI D&D concerns collective actors, rather than systems occupying a middle ground between tool and agent, the Cortés analogy cannot be appealed to. At first glance, the collective actors involved in AI D&D (states and corporations) align rather well with historical colonialism's collective actors, like

'Portuguese Empire' and 'Dutch East India Company'. But before diving into the analysis, a note on structure is in order. **Table 3** contains AI injustices that apparently warrant the literal interpretation of AI D&D. But contrary to these appearances, my argument will be that for AI D&D the literal interpretation is also unwarranted, because the AI D&D delineation doesn't satisfy the linguistically normative core of colonialism. AI D&D strongly converges with the operational definition of colonialism in *effects*, but not in constitutive *wrongs*. In other words, AI D&D and colonialism share *symptoms*, not *criteria*, because (i) the domination or autonomy restrictions effected under AI D&D are accidental *vitiations* rather than intentional *violations*; and (ii) the AI injustices are too disparate to count as co-institutionalized like colonialism's constitutive wrong and wrongful effects are. After some conceptual preparations, the following applies these two arguments to the five injustices in question.

	AI D&D's injustice	**Colonial because…**
i_1	Extractivism	AI *D&D*'s extractivism is a form of colonialism because it extends historical colonial practices into the present and implies historical colonialism's geopolitical schema of extracting profit from foreign natural reserves.
i_2	Objectification	AI *D&D* is in continuity with colonialism in the way it thrives on the objectification of human beings as data resources. Value is extracted from this objectification via datafication. This communicates a form of disrespect that is analogous to the form of disrespect the colonized experienced.
i_3	Exploitation	(1) AI *D&D* takes advantage of the economic vulnerability of workers in order to extract profit from them, not only by chartering cheap labor, but also through the further deployment of the AI systems thus created. (2) Economically dependent sales markets are exploited by the global North's AI *D&D* thanks to global power asymmetries rooted in AI capabilities.
i_4	Epistemic violence	AI *D&D* reiforces hermeneutic marginalization and displaces non-Western 'ecologies of knowledge' through hegemonic knowledge production, akin to how colonialism has done.
i_5	Cultural imposition	AI *D&D* interferes with native ways of sensemaking and identity formation, much like colonialism succeeded in doing.

Table 3 The grounds of comparison of AI ***development and deployment*** *with colonialism.*

Before introducing the two main arguments, the case of objectification can already be eliminated based upon what the previous subsection established. It was argued that AI systems are conceptually unfit for exploitation and colonization and that when there is an overpowering relationship, this isn't the particular institutional subversion of self-determination colonialism requires. Does any of this change on the D&D delimitation? I answer negatively, because even if sentences like 'AI D&D exploited/colonized *Y*' make more sense than with 'AI systems' in the place of *X*, the second conclusion still obstructs the goal of datafication being necessarily exploitative and colonial. This can be seen if we develop the previously taken example. If I (white Dutch male) interact with a digital interface and my search, click and purchase behaviors are datafied and sold to a party with (somehow) also access to (biometric) identifiers (passport photo, IP-address, official address, etc.), then it follows that this party has the power to constrain my future interactions with digital interfaces in subtle and less subtle ways. My choice set as social participant in the digital realm can be constrained and arbitrarily interfered with without reciprocation. This is an awful power asymmetry, but not a *colonial* one, because from myself being digitally dominated by some data broker, website or

platform operator, it doesn't follow that I'm dominated qua *political community*.[69] This is due to the capitalist beauty of the datafication mechanism: the profiting off of individuals (interpersonal domination) can proceed without the undermining of a community's political self-determination. Instead, the mechanism benefits from political stability across the board, because that enables citizens to be more proficient consumers and interactants with digital interfaces. Therefore the case of objectification can be dismissed.

(i) Now I turn to the accidentality of AI D&D's domination and autonomy restrictions. By itself, the subversion of self-determination encompasses both domination and violations of autonomy, but not all forms of domination and violations of autonomy are *colonial*, because domination and autonomy violence aren't necessarily institutional in the sense that colonialism is. I will show that the AI injustices in question can *narrow down the opportunity set available to people*[70] without satisfying colonialism's linguistically normative core. Let's further unpack this by drawing on what Pettit (2013, 27) writes:

> any factor that can reduce your freedom of choice in that sense constitutes a hindrance that makes one or another option unavailable: it removes some of the objective or cognitive resources required for accessing the option, or it restricts your use of those resources; it impairs or impedes the capacity that those resources establish.

So in for example cultural imposition and epistemic violence, freedom is indeed impaired but not by completely taking options off the choice menu (i.e., making them impossible), but by making certain options *too costly* for individuals to undertake. Once the cost of an option rises, there comes a point where, relative to the social group, it's *as if* it wasn't on the choice menu at all. Domination or restriction of autonomy are undeniably there in the shaping of cost structures of the choice menus of racialized people, besides increasing their relational vulnerability. But in the end, this type of domination remains insufficient for colonialism's conceptual core. To acknowledge that choices are constrained checks off but *one aspect* of colonial domination: (1) the choice menu being narrowed. But what is absent is (2) that this restriction is the result of an intentional actor who acts as a dominator (colonizer). The institutionalized intentionality required for colonial domination, which we see reflected in the mixed structure of the operational definition, is lacking in the deployment of AI systems playing a causal role in the disruption of cultural and epistemic concepts. The lack of (2) means that the autonomy restrictions that might – but not necessarily *will* – ensue are too *accidental* to be sufficient for colonialism's conceptual core because they aren't traceable to any *intent*.

This line of argument tackles cultural imposition and epistemic violence at the same time. For cultural imposition and epistemic violence, the 'usage of concepts' isn't completely analogous to 'options on the choice menu'. If concepts are made slowly unavailable or quickly become overshadowed by exogenous concepts, the concept's novices or forgetters needn't be duped into believing certain concepts cannot be used anymore; something that *is* the case in the institutional influencing of beliefs in order to shape interactions with available cognitive resources.[71] Whether these constraints on cognitive resources lead to

[69] I claim this holds on all possible delimitations of community, i.e., all groupings in its power set (white Dutch males; the Dutch as politically organized community; whites as racial group; males; Dutch males, etc.).

[70] I am indebted to Uğur Aytaç for making me aware of this line of reasoning.

[71] Besides this, any conceptual conflicts coming about through exposures to AI systems only interfere with communities' capacity for cultural-epistemic *identity formation*, which is distinct from (albeit closely related to) their capacity for *political organization*. Identity and polity formation are distinct, much to the dismay of global history, which contains too many examples of peoples that are opposed in terms of identity, but are trapped in a shared and sovereign political organization. One might object this sad fact shows precisely that viable political organization presupposes successful

restrictions on autonomy or to dominative interference, it should be recognized that in practice extremities of exposure regarding the former could lead to a convergence of both in *effects*.

The way to make sense of these effects without conceding that they constitute colonization is to invoke Pettit's (2013) distinction between invasion and vitiation. In §2.4., I showed how they relate to colonialism, as an invasion of another's choice set is intentional but needn't be institutionalized, whereas a vitiation remains the placement of a constraint on a choice set, but can be non-institutionalized *and* unintentional. I propose that the effects of epistemic violence and cultural imposition feared by scholars like Eke, Wakunuma & Akintoye (2023) constitute only vitiations of cultural and epistemic cognitive resources. With this conceptualization I don't want to diminish the importance of preventing these effects. Rather, I try to steer clear of conceptualizing this as colonialism *tout court*, which requires a far more organized subversion of self-determination and presence of the intent thereto. Even if one admits that this involves a deplorable negligence regarding the "expected level of care not to undermine your control" (Pettit, 2013, 39), it doesn't follow from the deploying party's negligence that this party is a colonizer, because we should distinguish *accountability for harms* from *intentional organization of harms' infliction*. Assuming that in general AI corporations strive for profit, not cultural dominion, what one is left with is an accidental imposition of cultural and epistemic particulars via the systematic design of AI in the global North and export of values embedded in these designs to the global South. Sure, the intention of cultural dominion can't be ruled out a priori, and neither can the charge of negligence, for refraining from action is sometimes also culpable. The point is that if we overinflate these injustices to colonial extents, we risk losing sight of what actually happens in AI D&D from the global North in the global South: cultural and epistemic disruption.

The last facet of epistemic violence to be addressed is the *reinforcement of hermeneutic marginalization*: would this amount to colonization? Domination is the relationship between a dominator and the dominated (see §2). Marginalization, as construed by Young, concerned structural oppression and Fricker's work expressed a hermeneutical inflection of this. In §2.1. it was shown that oppression can be redescribed as systemic domination focused on the suppression of individuals' capacity for participating in societal processes. It may seem that this could be redescribed as the subversion of self-determination, but any such linguistic temptation should be halted, because on the consequences of 'marginalization' Fricker (2007, 154) writes that "while a hermeneutically marginalized subject is prevented from generating meanings pertaining to some areas of the social world, she might well maintain a fuller participation as regards others". Unpacking this, it's clear that hermeneutical marginalization is so specific that, by itself, it doesn't follow that when someone is hermeneutically marginalized, her choice set is also invaded, let alone that the political self-determination of her social group is subverted. This isn't to say that the colonized aren't hermeneutically marginalized; again, the evidence presented by Wiredu for Ghana, Mbembe for Cameroon and South-Africa, Fanon for Martinique and Algeria, and Okeja for Nigeria tells us otherwise. It is only that the *conceptual relation* between colonialism and epistemic marginalization isn't that of noncontingent mutual implication, which halts any inference from reinforcement of this marginalization to colonialism.

---

identity formation, and that therefore the political community always depends on the community's shared identity. While *prima facie* plausible, this involves an overgeneralization. It dismisses out of hand the possibilities of (i) a sense of identity *deriving* from a political organization (e.g., the European Union in which formerly incompatible identities politically organize themselves and have derived a shared sense of identity from it since); and (ii) multiple distinct identities politically organizing according to a shared political ideology (e.g., world communism) or matters of concern (Latour, 2004) (e.g., the United Nations Framework Convention on Climate Change).

(ii) To deal with extractivism and exploitation, I argue that AI D&D's institutionalization is in mismatch with colonialism. To start comparatively, in historical colonialism, colonial conquest was enacted by corporations and states, but it was less fragmented than today's AI D&D is. In historical colonialism, next to European colonial empires being established from the 15th century and onwards, one observes a strong interplay between corporation and state as, for example, the Dutch East India Company was a fusion of trade corporations forced by the government of the Dutch Republic to capture military and economic potential, and the British East India Company was partially state owned. In today's AI D&D, the main players aren't co-institutionalized in this sense. Tech giants that have absorbed companies pioneering in AI (Microsoft, Facebook, Google, Amazon, Tencent, etc.) are candidates, and alternatively there are the geopolitical forces of AI legislation and regulation (United States, European Union, China, etc.), which deploy AI in their own right. But these two sides aren't aligned with governmental interests of conquest as in historical colonialism. By itself, this is no problem for the operational definition of colonialism, as it is flexible enough to have any collective actor capable of institutional domination erect a colonial frame. The problem I see in the relation of these corporations and states to AI D&D however, is that, like in the accounts surveyed in §3, corporations and states are amassed into an 'AI market' as an holistic institution. The problem lies with the construal of AI D&D as a market actor with the intention of subverting the self-determination of the global South. It makes sense to speak of AI D&D when one is talking about general patterns running through the economy of AI D&D, but not in a strongly agential sense – *pace* actor network theory and object oriented ontology. It's uncontestable that there are global North/global South asymmetries regarding AI capabilities, where production and resource extraction are located, who reaps and who sows, but the takeaway insight is that while the dynamics we can call AI D&D emerge out of the interplay of actors that could be colonizers, these dynamics themselves cannot literally be one.

Let's test this idea by applying it to exploitation and extractivism. Starting with worker and the sales market exploitation of the global South by the global North, do both cases literally instantiate colonialism via AI D&D? I argue that they don't, for two reasons: (1) agential capacities are mistakenly attributed to AI D&D and (2) they don't differ from capitalistic exploitation in general.

Regarding (1), we can insert argument (ii)'s conclusion and see what happens. In worker exploitation corporations profit from economically vulnerable workers' labor (ghost work, harmful data-labeling and beta testing). The workers' choice sets are constrained by their contractors arbitrarily, because their lack of alternatives (part of their economic vulnerability) is instrumentalized. However, by argument (ii), taking AI D&D to be the exploiter here is misguided because this stretches the relationship between exploiter and exploited to Roemerian abstraction: it isn't the case that AI D&D exploited, e.g., Kenyan workers; no, *OpenAI* did so. If we care about maintaining concrete relations of exploitation and domination, connecting actual legal persons that can be held responsible for their actions, and not abstract entities or market dynamics, saying 'AI D&D exploited/colonized *Y*' is of no use to us.

What about sales market exploitation? The corresponding narrative is that the global South's economic dependence and asymmetry in AI capabilities is exploited by the global North's AI D&D. The global North produces the necessary software and hardware, leaving the global South no choice but to demand in accordance with this supply. The global South's lack of software/hardware production cutting-edge AI needs is the instrumentalized economic vulnerability. Obviously, the global North's corporations profit from adding these countries to their clientele. But can a market *exploit* another market? I suggest answering affirmatively asks for a troublesome redefinition of exploitation, because the 'sales markets of the global North and South', like AI D&D, supervene on the interactions between many disparate collective actors. A claim like

'the AI market of the global North exploits/colonizes the AI market of the global South' thereby equally opts for a Roemerian conception of exploitation. If one accepts such a conception, then one should let go of the capacity to speak sensibly of '*A* exploits *B*', but I think the relational aspect of exploitation is worth having. Regardless of if one agrees with that, on the 'Roemerian stretch' one definitely needs to forestall that this sense of 'exploitation' has anything to do with colonialism's linguistically normative core, because that necessarily depends on the *identifiable relationship between colonized and colonizer*.

Regarding (2), the apparent similarity between AI D&D's worker and sales market exploitation and colonialism can be undermined by showing that they are products of *capitalism at large* rather than of an AI colonialism. If exploitation is concerned with the *relation* between two agents regardless of the good involved in the production of the exploited agent, then it follows that for both worker and sales market exploitation, the instrumentalization involved is standard to capitalism rather than special to colonialism. 'But', it can be countered, 'AI colonialism may still be literally the case if *AI capitalism itself* is colonial!' Granted; and this idea isn't foreign to this project, as the decolonial scholars from the tradition of the coloniality of power hold precisely this and think AI is colonial because of how it stands in relation to the coloniality of capitalism. However, one shouldn't rush to vindicate the literal interpretation because of this. On the coloniality of power thesis, colonialism has been ongoing ever since Latin America and Africa were first colonized by the European Empires. All actors involved have changed diachronically, but the asymmetries with respect to power, knowledge and being have stayed in place (Quijano, 2000; Mignolo, 2020). Even if this strong and completely historical thesis is granted, then there is nothing special about the role of AI in this respect. The minimum that can be deduced from this, is that AI is embedded into the colonial matrix of power (Muldoon & Wu, 2023), just like any asymmetric endeavor of the global North's digital economy in the global South is, such as the smartphone market, the Internet economy, overfishing, etc. If there *literally* is a colonialism of the global South by the global North, the way AI relates to exploitation isn't the reason *why* this is so.

Lastly, I discuss the extractivism that purportedly extends colonial practices and natural resource extraction into the present. On the operational definition of colonialism, AI D&D's extractivism of natural resources is clearly analogous to the wrongful effect of land-taking, combined with exploitation. As Jung (2023, 9) wrote, AI's extractivism is about "disposable lands and disposable people". On Moore's account of colonialism, the taking of land was essential, but Ypi showed it only pertained to *settler* colonialism (see §2.3.3.). An immediate problem that arises for the proposition that 'AI D&D is colonial because of extractivism' is that corporations opening up data centers or branching out to countries in the global South can only weakly be analogized to the total usurpation of land in settler colonialism (a similar case can be made for the intensification of diplomatic relations with Russia, China, the US, etc.). However, the ties of the taking of land to the dehumanization of the colonized are clear and this is strongly reflected in the extractivism of AI D&D. The AI economy demands large quantities of rare minerals, taken from countries in the global South without much concern for local ecological, economic and political impacts. Since similarities between today's extractivism and historical colonialism abound, does it amount to colonialism?

We can more straightforwardly reformulate the effect of extractivism we care about in the context of colonialism's linguistically normative core as *taking communities out of the resource loop and thereby taking away the power to develop their own technology*. Clearly this lays claim to communities' resources for self-determination, but the problem with the current extractivism is that it goes unregulated or is based on licenses and economic deals between political communities, however unjust those may be. Using Pettit's (2013) conceptual resources, we can conceptualize this as an *invasion* of choice sets and analyze the presence of institutionality and intentionality that colonialism's mixed structure requires. Regarding intention, the practice of mining

for resources, employing the local populace for this, etc., isn't arbitrary; on the contrary, this is a rational economic mechanism seeking to minimize manufacturing costs and maximize the profit from the sales income the final product yields. Therefore I think it can be accepted that AI extractivism and colonial exploitation share this feature. The case for institutionalization is however conceptually muddled: first, recall that no single corporation or country has the entirety of its supply chain clearly in sight or under control (Crawford, 2021); second, consider that extractivism equally faces the difficulty of construing the entirety of AI D&D as colonial actor. So the question then is: *how many exploiters does it take to make a colonizer*? While deceptive, this question's framing of the difference between current extractivism and historical colonialism as a gradual one is partially adequate. AI extractivism comes very close to the institutional and intentional subversion of self-determination, but I think a twofold caveat applies: (a) colonialism requires a clearly circumscribed colonizer, while all we have in extractivism is a set containing exploiter-exploited pairs. This means calling it extractivism or exploitation is apt, while calling it colonialism is hyperbolic. And (b) colonialism requires, which is where the mixed structure strongly follows the proceduralists Ypi and Renzo, the removal of the colonized's political self-determination: this only partly happens by interfering with a community's workforce and access to natural resources, thereby maintaining economic power asymmetries. However, this falls short of colonialism's linguistically normative core in any literal sense. Rather, this sounds more like the decolonial scholars' coloniality thesis: modern capitalism is built upon power asymmetries that historically originated from colonial rule. To word it differently, the power asymmetry in extractivism can be called a 'weak' colonialism, as contrasted to 'strong' colonialism that involves the *complete* taking away of self-determination of the colonizer by the colonized. But this isn't insightful more insightful or illuminating than just calling it 'extractivism' or 'exploitation', as it falls short of the demands of the literal interpretation that states it is colonialism *tout court* that is instantiated in the AI injustices concerned with AI D&D.

4.3. The failure of the literal interpretation: some objections

After discussing the two variants of the literal interpretation, both were found unwarranted. But these discussions bore some analytical fruits. The similarities and differences between extractivism, exploitation, cultural imposition, etc., with colonialism, show *why* the usage of 'colonialism' like that strays from the correct pattern of usage (Pitkin, 1972, 313). Remember that the introduction introduced the difference between linguistic normativity (semantic content regulating the identification of instances of the concept) and linguistic descriptivity (the concept's fit into empirical patterns) which I made heavy use of for defining colonialism and for analyzing the contours of AI colonialism. What happens in the comparison between AI and colonialism when taken literally is that empirical patterns of similarity are mistaken for signs of colonialism qua linguistic norm: from an empirical pattern of resemblance one moves to saying '*This* is colonialism'. So when, e.g., extractivism is identified and related to the AI economy, the comparers move from a perceived similarity to identifying extractivism as colonialism. Likewise, when a convergence in effect is observed, such as the conceptual loss feared in cultural imposition and epistemic violence, the comparison moves from descriptive similarity to linguistically normative instantiation. It is, as Pitkin (1972, 313) writes, that "our *verbal* deviations are symptomatic for our patterns of thought": once we use the word 'colonialism' to describe a case, we come to think – for ease of representation perhaps – that the case *is* an instance of colonialism.

So to recapitulate: the descriptive similarities between colonialism and the AI injustices harbor the temptation of moving from the linguistically descriptive to the linguistically normative usage. But, as the argumentation of the previous two sections has shown, this again and again proves too much of a stretch. Apart

from the differences between colonialism and the AI injustices that I identified, we saw that the Cortés analogy showed how AI systems fall short of being full colonizers, and that AI D&D's agential mismatch and the accidentality of its dominating effects prohibited its coupling with 'colonization' in a literal sense. As such, when from descriptive comparison linguistically normative deductions are attempted, unwarranted conclusions follow. *This* is the literal interpretation's philosophical confusion.

That being said, philosophical diligence requires that the argumentation's possible weaknesses are addressed before I rest content with these results. Replies to these objections follow immediately.

O_1: *The drift of language objection*. It might be objected that the analysis of the literal interpretation is mistaken, because of its reliance on colonialism's operational definition. Doesn't the usage of this definition commit an anachronism by constructing the concept of colonialism based upon compatibility with *historical* colonialism? (The focus on the historical use is implicit in the debates surveyed to construct the operational definition.) Based upon the idea that language 'drifts' towards new meanings with humans rarely having a 'say' in it, it can be argued that a new use of colonialism has arisen: colonialism*, that simply doesn't need to satisfy the subversion of a community's self-determination anymore for it to have colonialism's other normative connotations. Attempts at describing what remains 'colonial' about postcolonial situations have been made ever since formal decolonizations began. The paradigmatic example is 'neocolonialism', separately due to Jean Paul Sartre and Kwame Nkrumah. Neocolonialism is concerned with the "remnant" "influences of colonialism and its agents" in postcolonial societies and "the deliberate and continued survival of the colonial system in independent African states" (Taiwo Afisi, 2024). Consider Enrique Dussel's formulation too: Dussel (1985, 13) regards neocolonialism as the "international effort of the transnationals, which structure their neocolonies from within". In other words: the new imperialism of transnational operations of states and corporations that reenforces decolonized states' status as periphery. Édouard Glissant (1989, 49) also introduced another variant of neocolonialism: "successful colonization", where "The desired result has been successfully obtained: the populations concerned cannot, whether they like it or not, do without the American presence for money, goods, culture, education, health. Domination is complete".[72] Departing from neocolonialism, the sales market exploitation rejected earlier for example, seems instead proof of colonialism's continuation. The characterizations of neocolonialism cast the doubt that ruling out cases only because the subjugation of self-determination isn't *active* anymore may be mistaken – the subversion only captures the *advent* of colonialism. Can't 'AI colonialism' apply literally without colonial domination, just like that?

Reply. To counter this objection and the conflation with neocolonialism, I point out three things. Firstly, the operational definition's use of colonialism as norm isn't anachronistic, because that would confuse the decrease of contemporary cases of colonialism for the term itself to have changed in meaning. It would only be anachronistic if one political community subverting another community's self-determination such that they remain distinct as metropole and periphery weren't possible anymore. But this is false, as multiple contemporary instances can be named, such as the colonization of Palestine by Israel, the continuing

[72] Micronesia declared independence in 1979 and entered a compact of free association with the US in 1986, that continues until this day. Later on in his book, Glissant writes: "It is a matter of something on which no one has seriously reflected: the French colonizer, because he is fully aware of the fact that he has managed to put into effect (we shall see how) his particular brand of assimilation; the colonized Martinican because he is upset to see himself look so good in this mirror. It is a case of what I call successful colonization. What is the use, of making ritual and almost magical reference to the forms of decolonization in the world: national army, total revolution, liberation front, if the questioning of this success is not undertaken?" (Glissant, 1989, 89).

institutional subversion of Native-Americans' self-determination, China's colonial domination of the Uyghurs and its 'peaceful quasi-colonialism' of Mongolia and Nepal. In this sense, colonialism's semantic criterium isn't anachronistic and neocolonialism is a supplementary concept that has branched off of colonialism without inheriting colonialism's distinctive wrong, with only the wrongful effects being bequeathed. Neocolonialism describes not the colonial frame but the dependencies left in its wake and is a separate concept, integral to which is the historical legacy of its predecessor.

Secondly, the case of a drift within colonialism's linguistically normative core is itself implausible because it concerns an actual injustice that was avidly practiced less than a century ago. What would it mean if colonialism qua norm today expressed a different normative facet of social reality than it did, say, before 1950? Then it would be analogous to a previously annexed society where post-liberation the meaning of 'annexation' no longer included the *advent* of annexation. Could the use of the norm change through experiential distance from the reality it expresses? Or should this be viewed as bug, or cultural mutation, of the words 'colonialism' and 'annexation'? However, this doesn't make colonialism*'s warranted domain of application equivalent to that of colonialism. To illustrate, if colonialism* no longer requires the subversion of self-determination, then this doesn't make the set of events that colonialism clusters together equivalent to colonialism*'s cluster set. But placing AI injustices within colonialism's normative cluster set is precisely the goal of the literal interpretation. AI colonialism literally intended is unlike neocolonialism in that it is explicitly looking for colonialism *tout court* and not for dependency effects only.

Thirdly, the 'drift of language' from colonialism to neocolonialism inherits colonialism's linguistic descriptivity (conceptual periphery), but not the linguistic normativity (institutional subversion of self-determination). One can see this because neocolonialism is used to *evaluate* postcolonial dependency structures and the structures of economic dependency, cultural imposition, etc., of post-decolonization *are* similar to that pre-decolonization, but colonialism's linguistically normative core is no longer implied by neocolonialism, because what is wrong with them no longer involves the constitution of the colonial frame. The wrongs particular to neocolonialism aren't wrong in the same way that the *advent* of colonial domination is, so when one can rightly identify an instance of colonialism *it cannot be* an instance of neocolonialism and vice versa, otherwise our politico-theoretical hygiene wouldn't be in order, because we wouldn't have two distinctly applicable theoretical concepts for analyzing societies. The wrongs that neocolonialism shares with colonialism therefore are the latter's wrongful effects. These derive their meaning from the legacy of colonialism, but their reasons for being wrong don't appeal to the constitution of the colonial frame. The point of neocolonialism is to articulate *what remains of colonialism* after formal decolonization has deconstructed the colonial subversion of self-determination and isn't a case of the use of colonialism itself drifting towards a different meaning.

Since colonialism qua norm isn't anachronistic, literal AI colonialism is concerned precisely with the operational definition and the point of neocolonialism differs from that of colonialism, I conclude the drift of language objection doesn't reopen the literal interpretation's case.

O_2: *Racial exploitation contract objection*. The operational definition of colonialism doesn't include exploitation in its core. Philosopher of race Charles Mills (1997, 31) has argued that with colonialism, modern states' social contracts became *racial* contracts as well and that the racial contract was intrinsically an *exploitation contract*, adhered to by modern states. For Mills (1997, 32; 37), "the Racial Contract", which is the constitution of a society in which racial categories determine the political or 'who gets what and why', "is calculatedly aimed at economic exploitation" and is "continually being rewritten". Globally construed, exploitation thus always supervenes on colonialism by virtue of colonialism's racial component. If this thought is right, then

the analyses of AI D&D concerning exploitation are mistaken: because of AI D&D's creation of a racialized relationship concerning 'former' colonizer as exploiter and 'former' colonized as exploited, the literal interpretation of AI colonialism would be vindicated thanks to the presence of these two. To be specific, if this objection is right, worker and sales market exploitation amount to colonialism by them being exploitation of non-whites in 'former' colonies. AI colonialism simply writes a new page in the "continually being re-written" racial contract.

Reply. The following retort intends to show that not all exploitation is colonial: the racial exploitation Mills is concerned with doesn't equal colonialism as semantic criterium, because the *racial contract doesn't equal the colonial frame*. The racial contract, with its exploitative character, is broader than colonialism and a separate (albeit related) concept. Historical parts of the racial contract have included for example slavery, white supremacy, apartheid, the Jim Crow laws, segregation, etc., but none of them is coextensive with colonialism. Sure, they are interrelated, but North-American history in particular has shown how the racial contract can be in place without colonialism being the case. From the vantage point of the indictments of AI with colonialism it may seem like colonialism is *that broad*, but qua its linguistically normative core, it is a specific indictment.

However, couldn't all instances of colonialism still be parts of the racial contract? That would make colonialism (i) fundamentally exploitative and (ii) all cases of exploitation colonial. On this phrasing, (ii) clearly doesn't follow, given the operational definition of exploitation (§2.2.). Therefore the cases of exploitation needn't be reconsidered. The case of (i) is trickier, because it would imply that the racial contract's overall property of 'being exploitative' is transitive to its parts. I will only focus on settling this issue for colonialism in particular, which can be done by returning to the exclusion of exploitation from colonialism's conceptual core. Exploitation was deemed to be closely related to but insufficient and unnecessary for colonialism because not all colonial empires were strictly exploitative and exploitative rule has often been continuous across the transition from colonial to decolonial rule. This can be backed up from a mereological angle: if a system as a whole exhibits a certain property, it doesn't follow that that property can be sensibly ascribed to all its parts. Normally the 'mereological fallacy' of failing to heed this *non sequitur* is evoked in the discussion of material systems, such as the brain and its properties (Bennett & Hacker, 2003). But it can be seen that this is applicable to social objects as well: if e.g., the Dutch government as a whole is exploitative, then a priori one of the municipalities that is part of its registrational unity need not be. By analogy, this implies that the transitivity of 'exploitativeness' from the racial contract to colonialism is all but necessary.

It needs however to be acknowledged that the correlation between exploitation and colonialism is significant (Butt, 2013a). But just like Moore (2016) distinguished exploitation from colonialism, we can remember the lesson inherent in Valentini's (2015) image of peaceful colonization. The lesson I read (which isn't what she wanted to teach, namely that unjust procedural takeovers cannot define colonialism) is that colonialism qua norm is about *what comes before* colonial exploitation can take place: the constitution of the colonial frame. While this seems as mere reiteration of my position, Valentini's image takes colonialism not as historical phenomenon only (where exploitation sometimes seems impossible to disentangle from colonialism), but as a more time-unbound concept. That colonialism without exploitation is possible isn't to make the correlation between them unimportant, but rather that placing exploitation at its center distorts its function. In other words, there can be plausible configurations of several contingent features including exploitation (e.g., exploitation + racialization + violence) that one wouldn't want to call *colonial* because the subversive dynamic of colonizer/colonized isn't present. Such cases would be more aptly described by exploitation *tout court* or a direct hybrid concept like *racial exploitation*. Evocation of colonialism can at best

point to empirical patterns of similarity between historical colonialism and a target case. To conclude, it seems that even if a transitivity of 'exploitativeness' is possible, it doesn't apply to colonialism. Therefore the arguments against worker and sales market exploitation amounting to AI don't require reconsideration.

O_3: *Cortés disanalogy*. A problem with the Cortés analogy can be raised. One can object that it's actually a *dis*analogy because it arbitrarily applies a non-essential property of colonialism in the analyses of the literal interpretation. The non-essential property is the conceptual role of colonizer that AI systems were denied. As alternative, the Cortés disanalogy holds that colonialism is conceivable as an emergent social phenomenon without explicit role-players: colonialism is an institutional structure rather than a relation. Colonialism shouldn't be construed as a relational concept like domination or exploitation, but as a conglomerate structure that can be realized by disparate agents, AI systems included. In this sense, the standard use of the concept of colonialism requires amendment. But the upshot of this is that AI systems could be 'full colonizers' and the arguments of §4.1. need to be reevaluated accordingly.

Reply. The Cortés disanalogy is mistaken, because it stretches the concept of colonialism so far that it becomes dysfunctional. It isn't impossible for a concept to be created that articulates this, but it's simply not in accordance with the need that colonialism fulfills and it doesn't correspond to the need AI colonialism apparently fulfills either. This is so because all indictments of AI colonialism regarding AI systems weren't concerned with holding the system responsible or with a detached identification of the 'structure of colonialism' in AI systems. No, authors like Katz, McQuillan, and Birhane want to connect this to Western society, states and corporations in particular.[73] If the Cortés disanalogy is right, AI systems would *also* count as adequate colonizers via their contribution to the colonial structure.

But the responsibility attribution to the newly added class of actors in between object and agent fails. I argue the addition of *semi-autonomous systems* and *reactive models* like AI systems to the linguistically normative core of the concept of colonizer is an undesirable result, because these systems constitute 'responsibility vortices'. With responsibility vortex I mean that, the responsibility ascription *loses its function* – is sucked into a vortex – because this class of systems cannot adequately *repent*, experience punishment in response to societal desires for retribution, nor intentionally contribute to bringing about a change in the deplorable state of affairs. To illustrate, since the workings of LLMs are to some extent opaque in the sense that it cannot be traced why *this* particular text-string was outputted, they constitute a causal nexus in the harm that is the production of racial hate speech for example. But this nexus is also a *vortex*, because it makes no sense to ascribe responsibility to a statistical mechanism, however disjoint from human agency the output was. Just like young children or animals cannot be held fully responsible for their impulsivity and one looks to their parents or owners for retribution, even on the structural emergence interpretation of AI colonialism, AI systems are ill-fitted for the role of colonizer. The need the indictment of AI colonialism literally construed should fulfill is that of being a norm for judging similar cases such that like moral consequences and responses can be deduced from them. An example of such a moral inference is that if the wrongs of colonialism are the case, then decolonization and reparations are required. Because they are responsibility vortices, semi-autonomous systems and other ML-models clearly cannot share in these political forms of redress and the Cortés analogy isn't a disanalogy.

[73] E.g., Katz (2020, 181) thinks that responsibility for AI's whiteness lies with white societies themselves, Birhane (2020, 408) appeals to the role of "policy makers, governments, and firms" to redirect algorithmic colonization, while McQuillan (2023a) indicts the techno-elites. This is unproblematic as societies can rightly be held responsible the unintended consequences of their actions, as can states and corporations, including Big Tech.

***

Having successfully defended this chapter's arguments against a series of objections, I conclude that the literal interpretation of AI colonialism is unwarranted. Mbembe's dictum that in every colonization "a given land is declared virgin or uncultivated, and its occupants are confined to reserves" may be true, but cannot be applied to AI colonialism. As will be seen in the next chapter, the metaphorical interpretation of AI colonialism is more successful.

5

Interpreting AI colonialism as conceptual metaphor

> Why stretch, twist, press and expand, concepts in this way – why try to see *A* as metaphorically *B*, when it literally is not *B*? Well, because we can do so, conceptual boundaries not being rigid, but elastic and permeable; and because we often need to do so, the available literal resources of the language being insufficient to express our sense of the rich correspondences, interrelations, and analogies of domains conventionally separated; and because metaphorical thought and utterance sometimes embody insight expressible in no other fashion.
>
> Max Black, More About Metaphor

Two interpretations of AI colonialism remain: the metaphorical interpretation and the conceptual engineering interpretation. Here I defend the metaphorical interpretation by showing that AI colonialism is a conceptual metaphor. So far we focused on the mixed structure's applicability to AI colonialism. To vindicate the metaphorical interpretation, I will develop a theory of metaphor and metaphor's relation to concepts.

In the opening quote, Black stresses the irreducibility of metaphorical thought. The introduction probed the idea that metaphor permeates "the entire semantic, semiological, and performative field of language", to speak with Paul de Man (1978, 30). In any case, one should recognize "the proliferating and disruptive power of figural language". Applied to AI colonialism, I will conceive of it as a metaphor that disrupts the commonplace understanding of AI by transposing the concept of colonialism onto it, which results in a *perspectival shift*. It contributes to the articulation and disclosure of social reality to the language-user, in line Italo Calvino's (2013, 210) saying that "the only thing that I would like to be able to teach is a way of *looking,* in other words a way of being in the world. In the end literature cannot teach anything else". Similarly, Paul Ricoeur (1979) thought of metaphor as a cogitative counterpart to Wittgenstein's notion of "seeing-as" in which "another concept, related to this one, also becomes important to us". Metaphor supplies a *framing* – the 'stretching, twisting, pressing and expanding' – of one concept by another: "a seeing-as which occurs only while I am actually concerning myself with the picture as the object represented" (Wittgenstein, 2009, II, §199).

Black, de Man, Ricoeur and Calvino's common ground is an awareness of metaphor's potency to disclose such that both simulant and receiver of metaphor become able to relate to aspects of the world differently. As such it amounts to a *social* "thinking tool" (Dennett, 2017): it provides an analogical "submerged model" (Black, 1977, 445) we can share to theorize AI politically.

However, we shouldn't forget the *conceptual engineering* interpretation. There are many different ways to conceive of conceptual engineering, and the philosophical tradition contains many examples (Hopster, et al., 2023). The contemporary discussion about the right way to conceive of conceptual engineering is also lively. According to Jeroen Hopster and Guido Löhr (2023, 8) conceptual engineering responds to a conceptual disruption that is effected by technological change. The introduction, elimination or preservation of concepts address the conceptual gaps, misalignments or overlaps that arise via encounters with artefacts, norms or beliefs. Relatedly, Haslanger (2020) stresses the *ameliorative* dimension of conceptual change. Alternatively, Cappelen (2020) distinguishes lexical improvement (akin to Haslanger's amelioration) from *lexical expansion* as possible routes for conceptual engineering. Either the boundaries of the concept should be changed, or, similarly to Hopster & Löhr's 'introduction', we have something new. Lexical expansion is the

introduction of a new expression with a new meaning. Commonly called a 'neologism', an example of the lexicon's expansion is 'enshittification',[74] a word hitherto unused with a new meaning: the gradual decline in quality of the content of online platforms, such as social media. Furthermore, Hopster & Löhr (2023) introduce *adaptation* as a better conception than amelioration. However, for my purposes, the distinction between amelioration and lexical expansion (preservation *vs.* introduction in Hopster & Löhr's conception) suffices. §§3-4 namely presented no grounds for eliminating the concept of colonialism, while I showed that AI colonialism cannot be taken literally. So it follows that *either* colonialism requires amelioration to fit this new use, *or* that AI colonialism retains a new meaning that heralds the introduction of a new concept.

With this succession of philosophical ideas in place, consider the interpretations' success criteria.

i. For the metaphorical interpretation, the minimal criterium is that AI colonialism expresses *structural relations* between AI injustices and properties of colonialism's conceptual orrery. These structural relations could be epistemically valuable or superfluous. Upon relinquishing the literal interpretation, we let go of the instantiation of colonialism's linguistically normative core. Therefore, the metaphorical interpretation can only articulate a (partial) convergence of AI injustices with colonialism, which however supports rather than detracts from the metaphorical interpretation because, as Kompa (2021, 34) writes, "we prefer to exploit *contingent* properties, not *defining* properties, metaphorically".
ii. For the two variants of the conceptual engineering interpretation on the other hand, different minimal success criteria apply. For (a) conceptual amelioration, reasons should warrant the shift of 'colonialism' from use *C* towards a semantically or epistemically better use *C**. For (b) lexical expansion, a wholly new meaning *M** should be introduced, represented by a new term ('AI colonialism'), where *M** does not equal colonialism's *M* that it branched out from.

I will defend the soundness of the metaphorical interpretation and that it constitutes a ground for lexical expansion by (1) showing that AI colonialism is a conceptual metaphor that enables a way of *thinking* AI and (2) when AI colonialism nestles itself into our parlance, it could become a neologism, a concept expanding the lexicon. But this task isn't straightforward. To prefigure the connection between (1) and (2): the change from metaphor into neologism has a *temporal dimension*. Over time the metaphor dissipates and the image it communicates becomes a literal meaning. Consider enshittification again. This was itself a metaphor, as online platforms don't *literally* turn into shit. But the new literal meaning of this expression uses the component 'shit' as evaluative of the informativity of platforms' content rather than denotative of feces. Similarly, AI colonialism could become fossilized and entrenched in a literal use that forms its own conceptual domain, rather than connecting two conceptual domains. The lexicalization would mark the end of the metaphor, akin to bottles and rivers literally 'having mouths'. But currently, AI colonialism is an unlexicalized metaphor and the grounds for lexicalization have emerged, because the conceptual metaphor expresses a new, irreducible metaphorical meaning, distinct from the meaning of colonialism qua norm.

In short, regarding AI colonialism as conceptual metaphor entails that the descriptive inferential patterns pertaining to colonialism are mappable onto AI. AI colonialism meaningfully discloses coherent patterns of injustice in the domain of AI. Conversely, as candidate for lexical expansion, it presents a new meaning continuous with its topic of origin, coherent with existing lexical effects of related terms, that is able to anchor new social facts. Correspondingly, I argue in §5.1. that AI colonialism doesn't warrant the conceptual amelioration of colonialism. Subsequently (§5.2.) I will make the case for AI colonialism as conceptual

[74] The American Dialect Society's Word of the Year 2023.

metaphor, and (§5.3.) explain how this enables lexical expansion branching out from 'colonialism'. Lastly, §5.4. makes the case that AI colonialism invites an epistemically successful political perspectival shift.

5.1. Amelioration is not what the concept of colonialism needs

The motivation for conceptual amelioration in general is, as Haslanger (2020, 230) explains, that "we should seek not only to elucidate the concepts we have, but aim to improve them in light of our legitimate purposes". Our target concept is colonialism: its mixed structure is the intentional and institutional dehumanizing subversion of self-determination of one community (colonized) by another (colonizer), such that the colonized non-reciprocally stands as periphery to the metropole that is the colonizer. Colonialism's linguistic descriptivity captures patterns related to the contingent wrongful effects succeeding the colonial frame. Etymologically, *colon* stems from the Latin *colonus*, which roughly translates to 'tiller of the soil' or 'new inhabitant'. Its 'original use' was therefore closely related to a variant of colonialism: *settler* colonialism (*cf.* §2.3.2.), which is understandable, as in the ages before globalization and the European imperial expansion that resulted in the world economy, settler colonialism involving the violent taking of land was the only possible form of colonialism. Corresponding to the historical development of colonialism itself, the 20th-21st century use of colonialism has drifted towards encompassing a distinction between colonialism's normative function (semantic criterium for a moral partitioning of social reality) and its descriptive function (clustering of like cases).

How could the comparisons of AI to colonialism entail colonialism's amelioration? If they warrant conceptual amelioration, they should contain reasons for amending colonialism's current use *C* into a new use *C**. The comparisons would contain the boundaries of a new use *C** that better fit "our legitimate purposes" (Haslanger, 2020, 230). To unpack whether this applies to colonialism, I discuss Haslanger's variants of amelioration: epistemic and semantic amelioration.

By *epistemic* amelioration, she understands the *refinement of a concept* through "broader or deeper knowledge of the phenomenon" or the *improvement of experiential access* to the informational content of the concept. Conceptual refinement can proceed by acquiring knowledge of what a concept is about to obtain a "more fine-grained resolution" of it. Conversely, improving one's experiential access elucidates a concept's informational content to the concept user, via new experiences or the entertainment of new propositions and "different modes of representation" (Haslanger, 2020, 242).

Semantic amelioration is the making of changes to the partitioning of logical space a concept represents. *Alethic* amelioration improves how we account for a concept such that we can better track truth for it. In other words, engaging in social construction can yield a concept more tightly fitting to scientific research or social reality. Alternatively, *pragmatic* amelioration shifts the terms of coordination of a word, i.e., it widens or narrows its boundaries for application because that is useful to us. Lastly, *moral* amelioration involves a shift in informational content because of a word's ethical significance, i.e., to make its partitioning of logical space (no longer) fit a moral wrong (Haslanger, 2020, 242).

My argument *against* colonialism's amelioration is that including cases of AI colonialism *doesn't* "make things better" for users of the word colonialism in epistemic, or semantic ways. While the conclusion that AI colonialism can't be taken literally tempts the counterpoint that this 'already constitutes an argument for changing colonialism', holding this would be fallacious; as I will show, the AI cases fall short of being reasons for semantic or epistemic amelioration. Precisely because the linguistically normative core isn't instantiated in the AI injustices, the idea this points to a malfunction of colonialism is *rendered implausible*: unclassified cases ill-fitted for a concept's current domain would need to come packed with convincing reasons if we

are to change the concept's reach accordingly. Only if it where the other way around – namely that the comparisons would instantiate colonialism's conceptual core but we weren't prepared to call them colonialism – would a form of amelioration apply because of *our* misunderstanding of its domain of applicability. Just like merely calling a state a democracy doesn't make it one or is a pressing reason to change the concept of democracy, AI injustices' non-literal coloniality doesn't make amelioration apply to colonialism.

The amelioration-minded conceptual engineer would change colonialism's *intension* (its formal definition) to better capture its preferable *extension* (cases colonialism should apply to). For *refinement* of colonialism to apply because of AI cases, we would gain a better understanding of colonialism through consideration of our knowledge of these cases. E.g., extractivism would provide a case in point for colonialism's applicability, which we hadn't realized yet. But the refinement of colonialism so construed is misguided. First, the change in intension for including AI cases would require the deletion of colonialism qua norm from its conceptual foundations, because that isn't present in the AI cases. Second, it's important to see that colonialism's core is a moral concept, not an empirical one. Unlike 'spider', where the discovery of a spider with five pairs of legs would displace the criterium of having four pairs of legs from the concept of spider, the emergence of new facts leaves colonialism's core as it is. Instead, colonialism's normativity could be changed only when the people sensitive to it would choose to continue in a different way, e.g., no longer deeming it to be wrong. Therefore, no amassment of knowledge about states of affairs, like epistemic refinement prescribes, would warrant changing colonialism delineation. Empirical reasons are reasons of the wrong kind.

Likewise, the improvement of experiential access is also out of the question, because changing colonialism's intension to better fit AI-related extensions doesn't qualify as gaining a better understanding of colonialism *through experience of it*. The example Haslanger (2020, 242) considers is that "those who have experienced war with "boots on the ground" have a different appreciation of what war is." Now if the AI cases currently don't fit colonialism, how can our consideration of them make us understand *colonialism* better? If it's through their *resemblance* to colonialism, we have drifted from conceptual engineering to the domain of the metaphorical interpretation. Furthermore, if we ameliorate colonialism because this would make experience of the AI cases count as colonialism and hence make them experientially insightful for the latter, then we achieve the former via the latter, while improvement of experiential access requires the converse. This change in intension would no longer make, e.g., Indonesians' experience of Dutch colonial rule experientially insightful for the concept: displacing the importance of the institutional domination from the center of colonialism would partly destroy the knowledge Indonesians have of colonialism that non-colonized peoples don't. The inclusion of the AI cases lowers the moral bar for what counts as colonialism, and this levelling of the experiential playing field is undesirable.

Turning to *alethic* amelioration, this would apply if the AI cases' inclusion in colonialism yields the articulation of previously inarticulable truths about colonialism. A new representation would thus enable the formulation of true propositions that were deemed false on previous representations of colonialism, such as 'Facebook colonizes Sub-Sahara Africa' or 'AI systems colonize the global working class'. If including AI injustices in *C** was only an extension in domain, then this wouldn't be controversial, but the boundary of the colonialism's intension has to be redrawn for them to fit in. Otherwise, 'AI systems colonize the global working class' remains false because, by the Cortés analogy, AI systems lack the actorship of colonial actors and 'Facebook colonizes Sub-Saharan Africa' remains false because of the lack of institutional maintenance of a periphery/metropole relationship, between Facebook and states in Sub-Saharan Africa. The articulation of one specific wrong (*C*) is in conflict with the descriptive pattern inscribed in *C**. Opting for *C** requires abandoning colonialism's linguistically normative use by restricting its normative reach: its normative truth

tracking ('φ is a case of colonialism, therefore what is wrong with φ is A') is traded in for the expansion of the range of phenomena it tracks descriptive truth for ('φ falls within the family resemblance of colonialism, so φ is a case of colonialism'). For such exemplary propositions (or for algorithmic exploitation, data extractivism, etc., to constitute colonialism) some aspect of colonialism's linguistically normative core (domination, intentionality, dehumanization, institutionalization) must be relinquished. Preservation of colonialism's function qua norm for all canonical cases and addition of AI cases are mutually exclusive: a thinner concept of colonialism has a wider extension.

But it can be objected this change isn't undesirable, because the operational definition of colonialism is already too stringent. However, I respond that without this conception, the normative differences between colonialism and imperialism, annexation, or exploitation would remain opaque. Abolishing the linguistically normative core, would yield a broader extension, at the cost of clarity of intension. Likewise, another counterpoint is that colonialism is a family resemblance concept, so the operational definition was bound to run into these limits, that amelioration can in principle amend for. Admittedly, all definitions have limits, but for normative purposes, in line with the mixed structure's motivation, *not all patterns of resemblance are equally important.* Sure, if our purpose was purely descriptive, then an egalitarianism of patterns wouldn't hurt. But changing the distinctive wrong *after* discovering its absence, is committing a foul in conceptual analysis: then we simply abuse amelioration to reach the conclusion we desire. Furthermore, thinning out the linguistically normative core isn't as straightforward as it seems, because after eliminating the conceptual role-play of colonizer/colonized and the idea of colonialism as a dominating institution, what distinctive linguistically normative demarcation is left? Only the dehumanizing subversion of self-determination *per se,* in conjunction with one or more wrongful effects. I contend this trades in too much intension for extension; after all, other concepts are readily available to us for literally defining the AI injustices individually. Furthermore, we should care about preserving colonialism's narrow evaluative scope, as the gravity of the colonialism of the past shouldn't be lightened or forgotten and a distinctive conception of colonialism is necessary for recognizing the obtainment of future instances. Therefore the alethic variant can be dismissed.

Pragmatic amelioration would be supported when the extension of colonialism to new terrain is *useful to us*. Indeed, it *would* be useful to expand colonialism to encompassing the AI injustices. This is the comparisons' point, because they want the normative responses to colonialism to also be applicable to AI cases. It would facilitate the application of retributive concepts that are already part of our arsenal to AI cases, weren't it that this change in intension's utility is a net negative. The pros are (a) using an existing concept to account for what we experience as 'new' injustices, because they pertain to a new technology, and (b) literal extrapolating *to* AI *from* normative responses fitted to colonialism. But these don't weigh up to the con of displacing colonialism's linguistically normative role to widen the terrain of its descriptive role. To see why, I evaluate the gains: (a) only yields a coordinative gain because it enables a clustering of cases that was hitherto unavailable; (b)'s consequences are more weighty, as it is action-guiding and fills the current lacuna in our judgments of how to adequately deal with the AI injustices. But are these worth the risk of losing colonialism's central normative denotation? Colonialism is still an extant possibility, so the need for judging political states of affairs by means of it remains. Therefore, the need for maintaining an easy grouping of injustices we satisfy through (a) is less foundational for the stability of our political interrelations than the ability to distinguish the colonial mode of institutional domination, because there are less costly alternatives for tidying up the grouping of injustices, e.g., lexical expansion, or the relegation of the AI injustices to other existing normative categories (exploitation, violence, etc.). Additionally, (b)'s action-guidingness is clearly desirable, but foregoes that the normative responses it would enable (reparation and decolonization) are primarily

directed against the colonial frame's institutional domination, *not* to the disparate wrongful effects the constitution of the frame makes room for. So, the first gain doesn't outweigh the corresponding loss, while its second gain would warrant the change in intension, weren't it that it saddles us – quite unpragmatically – with the amelioration of normative responses to colonialism as well.

Finally, considering *moral* amelioration, changing *what* is wrong about colonialism *prima facie* seems like a great candidate for aligning colonialism with the AI injustices-extension. The point of extending colonialism to AI systems/D&D is the transposition of colonialism onto the AI cases as *explanans* for their wrongness. Controversy commences with the fact that what's wrong in those cases isn't the distinctive wrong of colonialism. But this doesn't immediately make moral amelioration obsolete, because moral amelioration could yield an important ethical benefit: framing it as colonialism will elicit different reactions from moral agents due to colonialism's negative valence, which can lead to moral improvement, prompting people to behave differently towards the AI cases.[75] What we're left with, is another functional trade-off: tailor colonialism to AI cases to facilitate moral improvement *or* hold on to colonialism's linguistically normative core as distinctive for the concept. But I argue that this trade-off isn't inevitable, because the benefit of changing colonialism can be achieved through other means that leave the linguistically normative core untouched. For example, moral improvement could already be incited by words like exploitation, extractivism, and objectification that elicit negative reactions of their own. The only downside would be their more narrow scope as compared to colonialism, but this con is outweighed by the pro of maintaining the current conceptualization of colonialism. If the reader doesn't follow me there yet, consider the possibility of ameliorating concepts *other* than colonialism to do the job. Maybe it isn't colonialism that stands in need of amelioration, but rather components of that core, such as intentionality or what counts as a political actor, community, or institution, so that AI cases can fit into its linguistically normative core. I won't argue for any of those options here, but the entertainment of those possibilities should dispel the confusion that moral amelioration of colonialism's linguistically normative core is a necessary stop along the route towards moral sensibilization to AI injustices. Therefore amelioration for moral reasons also isn't necessarily applicable.

Because no variant of amelioration turned out to be tenable, I disqualify this variant of the conceptual engineering interpretation.

5.2. Vindicating the metaphorical interpretation

With conceptual amelioration eliminated, I will now defend the metaphorical interpretation. The metaphorical interpretation is validated by showing AI colonialism expresses structural relations between AI and colonialism. The metaphorical articulation of structural relations strikes one as a less impactful result than the literal interpretation being correct, because an interconnection of the two domains for aesthetic or rhetoric purposes would already satisfy this minimal criterium. However, I will show AI colonialism is more than a superfluous metaphor. I will argue that AI colonialism is a conceptual metaphor that provides a *perspectival shift* on AI injustices by projecting colonialism's contingent properties.

Conceptual metaphors facilitate understanding by interrelating concepts, because they "are grounded in systematic correlations within our experience" (Lakoff & Johnson, 2003, 61). In line with the 'irreducibility theorists' of metaphor, I hold that metaphor expresses meaning over and above its constitutive terms' literal meaning (Johnson, 1981; Black, 1955; Ricoeur, 1978). Metaphors so construed are 'devices for reorganizing conceptual structures' (Johnson, 1981, 31). Conceptual metaphors such as 'time *is like* money', 'rational

[75] I am indebted to Uğur Aytaç for pointing me in this direction.

argument *is like* war', 'the mind *is like* a machine', 'theories *are like* buildings' (Lakoff & Johnson, 2003) or 'language *is like* an organism' (Gill, 1979) are ubiquitous in everyday language. I argue AI colonialism likewise aligns two concepts, showing that 'AI *is like* colonialism'. The designation of AI in terms of colonialism "reduces [AI] to those features" which colonialism makes "salient" (Danto, 1993, 30). For the resemblance to be structural – i.e., there being correlations between the respective family resemblances of colonialism and AI injustices – properties of colonialism's periphery should correspond to the cluster of AI injustices' properties. This metaphor in particular doesn't structure everyday experiences, but communicates a *new way of thinking* AI (Lakoff & Johnson, 1980, 308): thinking AI *as* colonialism.

The following further unpacks the theory of conceptual metaphor and provides an explanation for why AI colonialism is one.

Black's 'interaction view' of metaphor is foundational for the idea that AI colonialism is a conceptual metaphor. The interaction view holds that metaphor introduces a new meaning that is distinct from the literal meaning. It does so by imposing a "focus" on the literal principal subject of the metaphor so that the focus – the nonliteral subsidiary subject – "selects" or "filters" aspects from the principal subject. Applied to AI colonialism, AI is the principal subject (or 'target domain') and colonialism is the subsidiary subject (or 'source domain') and this application evokes the colonialism-"system of commonplaces" on AI (Black, 1955, 73-75).[76]

However, the interactionist view requires substantiation. Why is it about *structure* rather than aesthetics? 'AI *is like* colonialism' is, on Lakoff's (1993, 5) theory of metaphor, not the mapping *itself*. Instead, the conceptual metaphor is the set of *pairings* of AI's cultural imposition, racialization, etc., with their equivalents in colonialism. The *structural cohesion* of these pairings strongly 'emphasises and resonates', by which Black (1977, 439-440) means that its focus is hardly substitutable and supports a "high degree of implicative elaboration". This is to say that much of colonialism's internal semantic and inferential structure is projected successfully onto AI: there is a descriptive 'invariance' in AI colonialism with respect to colonialism, such that the conceptual domain of AI becomes connected to that of colonialism (Lakoff, 1993, 6-10). It's an *analogue model* that communicates "the same structure or pattern of relationships", a "sustained and systematic metaphor" that is a "distinctive mode of achieving insight", i.e., that helps us "to *see new connections*" (Black, 1962, 223; 236-237).

Let's further unpack this qualification by discussing three reasons why it succeeds.

[76] We needn't however uncritically accept Black's interaction theory. Here I present the elements of Black's thought that are congruent with and inherited by CM-theory, the dominant interdisciplinary (linguistics, philosophy, cognitive science) theorization of metaphor over the past thirty years. What for example Mary Hesse criticizes Black for is his employment of the literal/metaphorical distinction. According to Hesse (1988, 6), Black "maintains the literal/metaphoric distinction. Indeed he appears to think that without some literal meanings in the background of the interaction process, for example in the understanding of associated commonplaces, its consequences for meaning would be viciously circular. But I have indicated how with certain constraints upon classification built into natural language this need not be so. There is no reason for the interaction theory to maintain the literal/metaphoric distinction, and every reason why it should be regarded as a quite general theory of meaning change, in which case it would be inconsistent to maintain the distinction, since mutual interactions of meaning are pervasive throughout the language". I will not settle this controversy here, but I agree with Hesse that language in general relies more on metaphors than is often admitted. However, from this point it's an overgeneralization to make *everything* into a metaphor: the word 'literal' also has its use, namely *in contrast* with metaphor itself. In a way, the metaphorical needs the literal in order to provide its change of perspective (focus) on its principal subject.

(i) *Availability of domain-to-domain pairings.* AI colonialism presents domain-to-domain pairings of injustices for AI and colonialism: the correspondences between the two domains (see §3). AI D&D's exploitation and extractivism strongly resemble that of historical colonialism and the systematic disrespect of humans and resources involved can be likened to the appropriation of the people and resources of the colonized by the colonizer. Furthermore, AI systems' cultural imposition, racialization and epistemic violence can be paired with the forced imposition of exogenous norms, values, standards of justification and racial categories by the colonizer on the colonized. This closely fits §4's explanation of the conceptual deviation underlying the literal interpretation: the descriptive similarities between colonialism and AI injustices tempt one to move *from* colonialism's usage as description of AI *to* using semantic identifying them as colonialism. So there is ample evidence the conceptual metaphor is rich in content over and above the aesthetic likening or rhetorical abuse of both domains. By way of projecting not colonialism's *defining* properties, but its *contingent* ones onto AI, the domain-to-domain pairings undergird AI colonialism's qualification as conceptual metaphor.

(ii) *Extrapolation of organization.* Apart from pairings of content, the conceptual metaphor communicates the conceptual domain of AI injustice is organized similarly to colonialism. The normative evaluation of AI is *structured in terms of* the normative evaluation of colonialism's contingent periphery, because of the connections of resemblance between this periphery and AI injustices. The conceptual metaphor facilitates a *perspectival shift* or 'thinking-as' because it articulates that AI systems/D&D involve injustices that hang together analogously to how they do in colonialism. For example, the triangle of injustice of cultural imposition, epistemic violence and racialization is implied by AI systems/D&D analogously to how it is implied by colonialism: they share an "inferential complex" (to use Black's words). Unpacking this example further, in colonialism, these three injustices were possibilized by the institution of the colonial frame, but have an internal dynamic of their own. The creation of race entailed ideas of cultural superiority and the testimonial disqualification of, and hermeneutical prejudices towards, racially identified cultural groups. Given the racial dynamics at the heart of modern Western society, this interrelation has seeped into applications of AI systems, as the findings of Katz (2020), McQuillan (2022) and Cave & Dihal (2020) suggest: the capabilities of AI systems are crafted out of predominantly Western and white texts, images, and sounds, but are construed as universal feats of intelligence. There is a degree of insensitivity towards cultural and epistemic particulars (concepts, values, ways of formulation) outside of the training domain *as if* they are invalid, leading to the displacement of those 'exogenous' (from the perspective of the AI system) particulars from the domain of making sense, being true, valuable, sacrosanct, etc. The interrelation of racial, cultural and epistemic biases resembles a diffused picture of the interplay between racialization, cultural imposition and epistemic violence in colonial hegemony. Bringing this back to the theory of metaphor, the metaphorical evocation of colonialism in AI colonialism functions as 'filter' that "selects, emphasises, suppresses and organises features" (Black, 1977, 442) of AI cases to bring the pattern of resemblance with colonialism *to the fore*, while the resemblances to other injustices remain *explicitly unarticulated.* In this 'extrapolation of organization', the balance between expression and suppression is delicate, as metaphors

> …highlight certain aspects of a phenomenon – make them noticeable. They make us perceive certain gestalts, make us recognize certain patterns, make us liable to draw certain inferences and ask certain questions but not others. They may thereby guide research. Which metaphors are chosen in order to talk about (or frame) a particular object is driven by what we know and what we try to explain. Yet they may also constrain thinking. By highlighting certain aspects, they are downplaying others; they make us see certain aspects of a phenomenon but blind us to others. (Kompa, 2021, 40)

Besides that the extrapolation of organization from colonialism to AI works, this contrastive 'thinking-as' explains the *discourse* on AI colonialism, which has produced the set of domain-pairings in response to AI's impact on society. It's a discursive trope that hasn't caught on outside of media and technology scholarship, decolonial theory and non-academic organizations subscribing to some form of dependency theory, but it communicates that AI is neither neutral nor harmless and that AI involves injustices that are all too familiar to both the global North and South, *because of their familiar resemblance to colonialism*.

(iii) *Competence in use.* Regardless of the metaphor's popularity, AI colonialism is a conceptual metaphor because there is *competence* in its use without the requisite of it being actively *comprehended* in its usage. In everyday parlance, people's evocations of pairings captured in 'the mind *is like* a machine' or 'time *is like* money' are ubiquitous. The evocations manage to be useful because they bring across a readily understandable metaphorical meaning, without *actively* overseeing the embedding of their utterance in a network of similarities and dissimilarities between two conceptual domains. 'AI *is like* colonialism' is useable in the same way, because colonialism's linguistically descriptive inferential structure is applicable to AI: AI systems/D&D are metaphorical colonizers via their 'constitution' of interrelated processes of extraction, exploitation, objectification, cultural imposition, etc. Comprehension isn't needed for fruitful use of the metaphor, and it's currently partly lacking, as is illustrated by the literal interpretation's unwarranted temptation (but which isn't recognized by the comparers themselves) of moving from 'it looks like colonialism' to 'it is colonialism'. Although their identifications of colonialism fail, the calls for these associations by academics like Birhane, Mbembe, McQuillan, Muldoon & Wu and non-academics or organizations like Chatham House, Ndiaye, Dobrin, Khan and Sahbaz are thus *descriptively justified* if viewed as being insightful metaphors. However, if their comparisons between AI and colonialism were intended metaphorically – as most don't seem to be – backing this adjecting up with substantiations *as if* it were literal remains a puzzling endeavor for academics and civil organizations to undertake. The non-literal results of the previous two chapters corroborate that there is a sense in which AI systems 'invade our minds', degradingly 'objectify and racialize some of us' and 'propagate exogenous worldviews'. These injustices all *function like* they would do in actual colonialism, but without ties to an actually instituted colonial frame. Similarly, the analogy that the exploitation and extractivism involved in AI D&D converge with colonialism's editions of them is also stable and readily understandable. It makes us think 'extractivism and exploitation are involved in the AI economy *just like* they were involved in colonialism', which isn't *the same* thought as 'extractivism and exploitation are involved in the AI economy and they were involved in colonialism, *therefore they constitute colonialism*'. Subsequently we can move from the thought that 'extractivism and exploitation were part of why colonialism was wrong' to the conversely structured proposition that 'the AI economy is wrong because of its extractivism and exploitation'. AI colonialism thus enables a way of looking, articulating and expressing the relation between the cluster of AI injustices and instantiations of colonialism that we can competently employ, *because it works*, without, in the act of employing it, having to be conscious of *why exactly* it works – just like the insights presented by conceptual metaphors do cognitive work for us and present "a claim to truth or […] an opinion to be tested" (Miller, 1979, 166).

To recapitulate, I have explained what a conceptual metaphor is and I have presented three reasons for why AI colonialism is one: (i) it expresses readily available colonialism-to-AI pairings; (ii) it extrapolates the inferential organization of colonialism's peripherical injustices to injustices pertaining to AI systems/D&D; and (iii) it can be competently used to convey a metaphorical meaning to incite perspectival shifts without previous ratiocination of all domain-to-domain pairings.

5.3. Paving the way for lexical expansion

The remaining task is showing that the conceptual metaphor paves the way for lexical expansion. Recall that lexical expansion requires the introduction of a new expression with a new meaning (Cappelen, 2020, 142) and that if the lexicon is expanded with a metaphor, it becomes a neologism with the formerly metaphorical meaning as content of a *literal* use.

Since it was shown that AI colonialism, as it stands,

i. cannot be taken literally;
ii. doesn't involve conceptual amelioration; and
iii. is a conceptual metaphor;

it follows that:

iv. if the conceptual arsenal is expanded with 'AI colonialism' with meaning *M**, then *M** must be distinct from but related to colonialism's descriptive meaning *M*.

Both (iv)'s distinction and relation require further unpacking. To paraphrase, *M** is *related* to *M* because the latter is the source domain the conceptual metaphor draws from, but it has to be *distinct* from it because otherwise there would be no 'new meaning' to be lexicalized. *M** doesn't allow changing colonialism's use *C* to *C** (amelioration's approach), but rather an emergence of a *C′* that specifically pertains to *M**. *C′* elevates AI colonialism to our range of standard expressions such that it is a domesticated, literally evoked image, no longer an interrelation of separate domains inciting a perspectival change. Here it's helpful to turn to Kompa's (2021, 28) distinction between *diachronic* metaphors that "we use today in a non-figurative manner" but "were metaphors at some point in their history" and *synchronic* metaphors that "are alive and kicking" and "still come with the feeling of displacement". AI colonialism is a synchronic conceptual metaphor, but the temporal process of cementation into society's parlance on AI could in principle lexicalize it, turning it into a diachronic metaphor.

Succeeding this schematic start, the constraints on the conceptual metaphor's *M** need to be further elucidated to show whether lexical expansion applies. Cappelen introduces four caveats that need to be dispelled before lexical expansion is justified – otherwise one could expand the lexicon with new terms representing arbitrary meanings *ad nauseam*. Two certainties are however that (1) *M** encompasses aspects of the meaning of colonialism qua description – the articulation of structural resemblance between AI and colonialism's wrongful effects – and (2) that *M** doesn't express a correspondence between colonialism qua norm and the wrongness of AI injustices. As such, *M** doesn't necessitate the mirroring of normative responses to colonialism, such as reparations and formal decolonization.

The merit of lexical expansion according to Cappelen (2020, 143-146) is that it enables a way out of discussions over what one '*really* means' by a word, e.g., 'freedom'; quibbles over the primacy of different, equally tenable, uses of a word, masquerading as discussions. Lexical expansion is a way in which Cappelen thinks we can 'fix language' by clearly allocating new meanings with new terms. The corresponding maxim of 'always expand' says: 'if it has a new meaning, lexicalize it'. The string of letters we choose to lexicalize doesn't matter that much, but can possibly evoke (un)desirable reactions of its own. Equally, there are cases for which lexical expansion is undesirable and Cappelen's constraints prevent the undesirable proliferation of lexical terms. Therefore, to make the case that lexical expansion can follow from AI colonialism being a conceptual metaphor, I discuss whether his four constraints apply to AI colonialism (*M**).

(i) *Preserving the lexical effect* of the original term can ground rejections of lexical expansion because leaving the original term's "massive effect in social, political, legal, medical, and interpersonal contexts" untainted

is desired. As Cappelen (2018, 123-128) explains: "An expression can have cognitive and emotive effects over and beyond (and in some sense independently of) any of its semantic and pragmatic properties", such as is the case with brand names, children's names, pejoratives, heavily contested concepts such as marriage, and, unsurprisingly, metaphors. The lexical effect is the connotation the use of a word evokes for ordinary language-users, which can be stable across populations, temporally unstable, or vary per individual depending on the effect (Cappelen, 2018, 143). If some lexical effect is somehow important, then its preservation is desirable in words that expand from what tokens this effect. Conversely, if a new term evokes the same reactions as an existing term, but this isn't in accordance with the former's intended use, one should refrain from lexical expansion.

Does AI colonialism elicit reactions inherited from its parent concept ill-fitted for its intended use, or distort its parent concept's use? Calling something 'colonial' is a grave indictment, which, if justified, leads to strong social and political condemnation and possibly warrants legal responses. The debate on reparations for colonialism illustrates this, but simultaneously shows the controversiality of social, political and legal responses to (historical) colonialism (Badru, 2010; Butt, 2013b; Forrester, 2019; Howard-Hassman, 2004). Given that public and scholarly discourses on responses to colonialism alike are in disarray because of controversies over questions like 'Are reparations required or not?' or 'Is formal decolonization sufficient?', I argue that colonialism has a lexical effect, but not an homogenous one. It's plausible, partly because of the heated controversy, that colonialism's lexical effect isn't stable across communities, such that it varies per person, depending on the individual's socio-political background. Therefore, functional homogeneity in emotive or cognitive reactions that we should protect from being distorted is lacking. Alternatively: does the introduction of AI colonialism as lexical item displace colonialism's heterogenous lexical effect in a detrimental way? I contend not, since AI colonialism covers a different domain, but carries over inferential patterns to the domain of AI *and* some of the feelings of outrage, indignity and wrongdoing that belong to colonialism's lexical constellation. Additionally, evoking the same lexical effect – i.e., moral controversy – as colonialism will probably be beneficial to those who want to right the AI injustices AI colonialism is concerned with. So I conclude that in absence of a detrimentally distorted lexical effect, this caveat doesn't obstruct AI colonialism's lexical expansion.

(ii) *Topic continuity* is the idea that people use words intertemporally to "*say the same thing*"(Capellen, 2021, 140), the preservation of which is important to retain the capacity of interrelating statements about 'the same thing' diachronically. It motivates resisting lexical expansion when the new lexical item introduces a discontinuity with the original it's important to see that people at different times are using interrelated concepts. For example, Plato's conception of justice differs from how Rawls uses the word, but recognizing their circumscription of the same concept is imperative. With this in mind, it's easy to see that topic continuity isn't a problem for AI colonialism, because continuity with its stem concept colonialism is retained, despite having distinct meanings. Unlike different conceptions of justice or freedom, they don't circumscribe an identical conceptual domain, but their interrelation is the mirroring of the connections between the injustices they both map onto social reality. The relatedness of *M** to *M* trumps their distinctness in strength: they're not distinct enough to erect a topical separation. After all, lexicalized AI colonialism wouldn't be a subclass of colonialism, but rather a *commonplace* communicating 'AI systems/D&D *are like* colonialism'. If this weren't the case, then the topic continuity constraint would obstruct the lexical expansion of *M** into the specific string that is 'AI colonialism', because the two terms are *prima facie* explicit linguistic kin due to their tokenization.

(iii) Appealing to *the anchoring role* of a lexical item is a variant of preserving topic continuity. It opposes lexical expansion if a term has the role of pinpointing a cluster of like cases. Lexical expansion starts *somewhere*, after all. Cappelen discusses the example of freedom and certain conceptions of freedom that belong to its cluster. The cluster is determined by what properties are in its neighborhood: for example, 'red' isn't, while 'non-domination' is. The appeal to the original term "simpliciter" "might be theoretically *indispensable*" for how we use the existing word. For example, the word 'quirkdung' can be used to denote a conceptual child of freedom where being free is compatible with being dominated, but what use would the introduction of a lexically unrelated token with a technical meaning have to us? If we care about distinguishing this meaning from 'freedom simpliciter', then we can solve this problem by preserving freedom's anchoring role and choosing a token that expresses kinship with freedom, e.g., 'freedomination'. Applied to AI colonialism, this clearly remains closely related to its lexical parent. It isn't a subset of the cluster of colonialism, but it *is* in its neighborhood. What's more: AI colonialism denotes a cluster of like cases that are in the conceptual periphery of colonialism. The question that remains to be answered then is: does the situating of AI colonialism in colonialism's periphery distort the family resemblance function of colonialism? I argue that this isn't the case. AI colonialism namely enables the formation of a *new* cluster of instances: the AI injustices of §3. The formation of this cluster doesn't distort the original cluster anchored by colonialism, because the former's relation to the latter is one of *derived resemblance* and of substitution. Simultaneously, by (ii), basic topic continuity remains in effect. Therefore AI colonialism anchoring its own cluster is non-exclusive with colonialism's anchoring role.

The last caveat is: (iv) a lexical item's *role in social ontology*. As Cappelen (2020, 146) writes: "According to many views of social ontology, language plays an important role in the creation and preservations of social facts. There isn't much agreement on just what that role is, but there is fairly broad agreement that it plays a role". Lexical items can be valuable for forming propositions about social reality. Cappelen (2020, 146) suggests that "changing the meaning of a lexical item might contribute to a change in social reality." An improved version of the original lexical item but differently tokenized might not have the same lexical effect on social reality as the original item, but were that original ameliorated in epistemic or semantic content. Cappelen's point is that the relation of a term to bringing about or shaping a social state of affairs can be a reason to opt for amelioration instead of for expansion. This is admittedly a problem, since if we care about lexical effect or topic continuity (see (i) and (ii) above), we are engaged in conceptual engineering in the first place to make our concepts *better capture or serve certain social facts*. But applied to AI colonialism, the worry that its role in social ontology would make amelioration of colonialism preferable, doesn't obstruct the expansion. Since we are in the game of lexical expansion for improving our conceptual arsenal, we care about the continuity of AI colonialism and colonialism's lexical effects and topics. Furthermore, because AI colonialism wants to express social facts pertaining to a very specific social domain – the development and deployment of AI systems – it isn't that we want to affect social facts *about colonialism* in using the concept. The conclusions one can draw from valid descriptions by way of AI colonialism are analogous to *some* facets, but *distinct from others* of what colonialism articulates, but the social facts they pertain to aren't *interchangeable*, e.g., we cannot describe Cortés' colonization of Mexico using 'AI colonialism', nor can we describe the worker exploitation involved in the AI economy with 'colonialism *tout court*'. The social facts expressed by AI colonialism *intentionally pertain* to a different domain than that of colonialism. Therefore the final caveat also doesn't apply to AI colonialism.

Something always remains a metaphor, but once turned into a lexical item, the enlightenment of its disruptive function might be dimmed: "the entry in the lexicon is a metaphor's obituary" (Kompa, 2021,

28) and its corresponding perspectival shift, its 'sense of displacement', is domesticated. But having overcome Cappelen's obstacles, I conclude AI colonialism's constituency as conceptual metaphor paves the way for lexical expansion. The final section will argue that AI colonialism is epistemically successful as well, via an investigation of metaphor's epistemic benefits.

5.4. Epistemic hurdles for AI colonialism

As conceptual metaphor, AI colonialism cannot be entirely epistemically unilluminating. But it's true that, as Ervas & Gola (2013, 33) write, "literal meaning is never completely suppressed. There is simply a difference of degree", because something of the meaning of the literal vehicle not only is *still active* in metaphor, but also *has to be* for conceptual metaphors to communicate analogies. Likewise, Taylor & Dewsbury (2018, 4) rightly state that: "No metaphor is perfect, and incongruities between target and source meanings are unavoidable. Some metaphors, however, may be more (or less) constraining […]." One main incongruity has already emerged: colonialism's conceptual core doesn't survive the metaphorical projection of contingent, rather than defining properties. If we combine what Ervas & Gola, Taylor & Dewsbury and especially Kompa say, it becomes clear that *in potentia* metaphors can be epistemically illuminating as well as obfuscating. Therefore it's worthwhile to investigate on which side of this epistemic spectrum we should place AI colonialism. As the previous section showed, AI colonialism is constituted by the successful extrapolation of the organization of colonialism's periphery onto AI injustices in the form of domain-to-domain parings. Therefore, the selection, emphasis, suppression and organization of features of AI systems/D&D in terms of colonialism bear epistemic fruits – but determining *what* those fruits are invites further investigation.

Before we can grapple with metaphors' conditions for epistemic success, it's important to situate the notion of conceptual metaphor in political theory and ask what metaphors could hope to achieve there. Arthur Danto (1993, 22) reminds us that metaphors pertain to the "politics of the mind": a metaphor is nothing without "manipulation" of the receiver into simulating a certain perspective. "If we lose sight of the psychological power of the metaphor", Danto (1993, 26) writes, "and think of it merely as a figure of speech – or, in the case of pictures as a kind of manneristic conceit – we have lost sight of something central in metaphors and in ourselves".

Corresponding to metaphor's political effect, we can, firstly, identify that qua being a political metaphor, AI colonialism politicizes a technological domain and sociotechnical system. Frank Ankersmit (1993, 162) corroborates this idea of political metaphor disclosing reality in a politicized way. The perspectival shifts they incite "predispose us in favor of a specific line of action and it is because metaphors embody proposals that they produce this effect." According to Eugene Miller, "Metaphor discloses the meaning or reality of political things not by constituting them, however, but by *manifesting their intelligible structure*" (1979, 162; italics supplied). Accordingly, AI colonialism pushes a certain view of political reality onto those who follow its perspectival shift.

Secondly, with Ankersmit's (1993) study of political metaphor from Plato and onwards, we can identify political metaphor's three main features: "goaldirectedness, unity and distance". The political metaphor *unites* source and target domain with the *goal* of inviting a particular political perspective on the target domain at the *distance* that every analogy presupposes. It provides a 'political model' of a target that facilitates understanding the target in terms of the source that, by glossing over their dissimilarities to emphasize their similarities, is necessarily distanced from the empirical "rough ground" (Wittgenstein, 2009, §107). Political metaphor incites political action and gives meaning at a "[d]istance [that] stabilizes political reality" (Ankersmit, 1993, 193-194), which places the political in between what Ankersmit calls reality and "externality":

the interplay between factual states of affairs and (metaphorical) narrativization. AI colonialism unites AI and colonialism with the goal of disclosing the detriments of AI injustices at a distance that stresses the similarities with colonialism's contingent periphery.

With these features of political metaphor in mind, Kompa's (2021) claim that a metaphor can be theory-constitutive can be extended to AI colonialism. AI colonialism invites political theory and postcolonial studies *into* the domain of AI to conceptualize what AI ethics has failed to do up until now: AI's embedding in the colonial matrix of power and its political, ecological and epistemic interrelations with Western colonial heritage. AI colonialism is a *manifestation of an intelligible structure* in between reality and externality that, through the communication of an irreducible meaning, invites taking a perspective. It lays down "a claim to truth" or "an opinion to be tested" (Miller, 1979). Rejecting or fortifying this claim on political reality (with its consequences for action) is the challenge the metaphorical meaning of AI colonialism poses to political theory. Given the domain-to-domain pairings that constitute AI colonialism, *should* we theorize the political reality of AI systems/D&D as it invites us to do? Like Ankersmit, Lakoff & Johnson (1980, 321) state that metaphors guide future action by helping cohere human experience. AI colonialism in particular does so by communicating AI's refraction through colonialism, such that certain normative inferences are invited, particularly that the AI economy is no politically neutral marketplace, that AI is related to global power asymmetries and that AI's promise of welfare comes at the expense of others' wellbeing. Now if it's misunderstood as asserting *literal* colonialism in the domain of AI, the risk of reading other normative inferences and actions than those the metaphor warrants emerges. To illustrate, a statement like 'That man is a wolf' clearly doesn't warrant bringing the man to a zoo or a nature reserve. Rather, it invites us to talk about the man's hostile and feral behavior (which is a commonplace about wolves that is itself *not defining* for wolves) (Black, 1955).

Having analyzed the political invitation the meaning of the conceptual metaphor communicates, we need to evaluate whether it does so *effectively*. Is AI colonialism epistemically successful in this respect? To answer this question, I review Kompa's (2021) epistemic success criteria for metaphors and simultaneously translate them from philosophy of science to political theory.[77]

(i) *Can AI colonialism be shown to exist?* The AI injustices associated with colonialism are grave and actual and surely exist (see §3). The metaphor doesn't conjure up non-existing objects or hypothetical social phenomena in order to communicate its meaning. This places AI colonialism at a smaller distance from political reality than, e.g., Plato's 'allegory of the cave' or 'helmsman of the state' metaphor (Ankersmit, 1993), since, given the detailed argumentation necessary for separating the metaphorical and literal interpretations, AI injustices come pretty close to colonialism, while it remains a grander imaginary stretch to conceive of society as a cave or the state as a ship. This question is straightforwardly answerable for metaphors pertaining to observable objects, but is more difficult for the political that stands between reality and externality.[78] Without wading into discussions of the observability of the political, I suggest that AI colonialism's synthesis of domains invites, depending on the receiver of the metaphor, taking a *previously unknown perspective*, which

[77] I refrain from discussing the seventh condition of literal explicability, as §4 already showed that AI colonialism is at best partly literally explicable. If a complete literal explication of AI colonialism is attempted, then the fact that colonialism's conceptual core isn't instantiated obstructs the literal explication. However, AI colonialism can be literally decomposed into its domain-to-domain pairings: six injustices that §3 showed form the contours of AI colonialism.

[78] For political things, Miller (1979, 165) writes, their "discovery" "is a case of coming to know something of which we had no prior experience or awareness. At other times, it is a case of discerning more clearly something of which we had already some impression, some vague and fragmentary experience."

would lead people to either 'discover' that AI's political reality could be like this, *or* make the already known downsides to modern capitalism *more clearly discernable* with respect to AI. Therefore we can answer that the relations and social processes the metaphor discloses aren't fictional and are as real as political theorization requires.

(ii) *Is AI colonialism's metaphorical connection inference preserving?* Colonialism's injustices are interconnected. As was shown in §5.2., parts of colonialism's inferential structure are preserved in their mapping onto AI by virtue of the availability of domain-to-domain pairings and extrapolation of organization. In colonialism, elements of the periphery are inferentially related[79] and these interrelationships are preserved in their mapping to AI. D&D's extractivism is entangled with dehumanizing objectification and several forms of exploitation and in turn the racialization, cultural imposition and epistemic violence related to AI systems and D&D form an interrelated triangle by virtue of the entanglement of ethnicity and racial categories with geopolitical power and user/developer asymmetries (see §3). §5.2. already developed the preservation of the interrelated triangle of injustice that is cultural imposition, epistemic violence and racialization in AI systems/D&D. Two other examples of the preservation of inferences are (a) *the interrelation of extractivism and exploitation* and (b) *the interrelation of cultural imposition and epistemic violence to global power asymmetries*. For (a), we see that in colonialism the possibility of extractivism was rooted in the exploitation of the colonized, while AI colonialism, in absence of literally colonized communities, invites one to see an extractivism based on large scale exploitation, as a mirror of colonialism's wrongful effects. For (b) then, the part of colonialism's inferential complex that is replicated in AI colonialism is the imposition of cultural particulars and patterns of testimonial and hermeneutical injustice. In trajectory it follows how the power distribution between the global North and the global South is North-skewed. Since yet other connections between the six constituent injustices of AI colonialism that mirror the organization of colonialism's periphery can be discerned, it's clear that AI colonialism is inference preserving.

(iii) *Are there strong disanalogies?* The strongest disanalogy is that colonialism's conceptual core isn't part of the projection from source to target. According to Miller (1979, 167) this is however unavoidable, as in political metaphors "[s]ort crossing always presupposes fundamental differences as well as likenesses". Specifically, the disanalogy is constituted by the absence of a clear dynamic between colonizer and colonized and the lack of institutionality of the domination in question. This is thus where the externality of AI colonialism's political perspective stands at a distance from reality: only in metaphor can the possibility of 'colonialism without a colonizer' be entertained. In this perspective, colonialism functions as a model for AI qua sociotechnical system and, as is the case with all models, enables the analyses of some features through the suppression of others. As a result, the cost of the disanalogy comes with the benefit that the meaning of the resemblance captured in AI colonialism is the *structural emulation* of the effects of colonialism by societal and geopolitical dynamics of dependency in the AI economy, driven by the development and deployment of AI systems. This is the way in which AI colonialism, as political metaphor, "can be a way after knowledge that leads somewhere, a path that takes us beyond the familiar experience of everyday life to knowledge of political things as they are" (Miller, 1979, 169): an invitation for political theory and further inquiry into its claims to truth.

[79] For example, Okeja (2022) and Wiredu (2002) have shown, cultural imposition goes hand in hand with epistemic violence; Fanon (2008) and Mbembe (2019) showed epistemic violence is but one form of colonial violence, which is itself tied up with dehumanization and exploitation.

(iv) *Are AI colonialism's gestalts theoretically fruitful?* Apart from the preservation of inferences already discussed above, which I consider to be a theoretical fruit, AI colonialism lets colonial wrongs loom like a specter over the otherwise techno-utopian global deployment of AI for Western profits. This illustrates its fruitfulness: it provides Western techno-utopianism with a much needed mirror. The moral perspective it invites encompasses that what is wrong with the current asymmetry of AI D&D and the one-sided universalism of AI systems resembles the ways in which colonialism's wrongful effects are wrong. The metaphor's evocation communicates that AI welfare and capabilities are built upon extractivism, exploitation and objectification. It also communicates AI's spreading of the global North's racial, cultural and epistemic dominance. Therefore, what this opinion to be tested suggests, is that there are substantial wrongs involved with this dual extension of capitalism and Western digital technologies. The theoretical merit for a political theory of incorporating AI colonialism therefore is twofold. First, it can function as a *target*: it lays down a claim to truth about the political reality of AI to be validated or debunked. Second, it can function as *postulate*: if one takes the offered political stance and accept its vision, then AI colonialism becomes a theoretical tool to disclose further relations in the political reality of AI. Now one can ask: could these things not be done without AI colonialism? It must of course be admitted that one could arrive at the same conclusions through some other way, much like it's plausible that if Harvey hadn't likened the workings of the heart to that of a pump, or if McCulloch and Pitts' perceptron wasn't modelled on a neuron, some other analogies could have emerged leading to an understanding of the heart or the advancement of connectionist AI. But this doesn't diminish the salience a particular metaphor can have for our current epistemic background – much like a particular gestalt can only emerge against a backdrop that is somehow homogenous in contrast to that gestalt (Ankersmit, 1993).

Without AI colonialism, a political theory would have to do a lot of work in explaining, interrelating and identifying the six types of contemporary AI injustices along different lines. This however comes easily with the projection of colonialism onto AI; call this the metaphor's *pragmatic benefit*. Besides that, there is the metaphor's integration of what are normally distinct disciplines: postcolonial studies, computer science and political philosophy/theory can work together in theorizing AI injustices from a shared starting point, namely that AI colonialism places a claim to truth that is far from outright wrong. Call this the *unifying benefit*. Finally, there is the *generative benefit*: AI colonialism generates new avenues for theorizing in the sense that it can be examined if the historical forms of redress for colonialism's wrongful effects can be transposed onto AI. So I conclude that AI colonialism's theoretical fruits are at minimum the embedding of AI into political theories of AI injustices, and at maximum the functioning of target for validation or postulate for further theorizing with pragmatic, unifying and generative benefits.

(v) *Does AI colonialism constrain thinking in important ways?* AI colonialism constrains thinking in at least one sense pertaining to its aggregation of different AI injustices into an organized whole. The aggregation is constraining because it runs into the metaphor's limits: a literal colonizer is absent in AI colonialism. The interconnections between the particular AI injustices are emphasized, but what doesn't resonate through in the metaphorical perspective is (a) that these injustices involve many disparate actors, none of which are in control of the 'colonial' supply chain; and (b) that AI systems, despite being part of a surge of technological and economic development, are likewise heterogenous and related to, e.g., cultural imposition and epistemic violence in diverse ways, if related at all. The projection of colonialism onto AI risks suggesting that these injustices hang together in a holistic, quasi-teleological fashion, while further inquiry may point out that many of AI's societal effects are unintended by most of the perpetrators involved. My point is that the metaphor of AI colonialism might be a great starting point for theorizing AI on a geopolitical level, but

transposing it as *explanans* on some of the individual cases it aggregates, might be overly formalistic and obscure the explanatory success of less complex normative concepts for that particular case. Just like decolonial theory's notion of 'coloniality' works great on an abstract level – as a counternarrative to that of Western modernity – AI colonialism stimulates counterreflections to Western AI utopias and dystopias, but might bring in too much conceptual baggage for explaining single injustices in isolation.

(vi) *Does AI colonialism give rise to new productive metaphors?* The generative benefit discerned in condition (iv) above prefigured this question. AI colonialism, with the current semantic and epistemic content of its projection, can engender new metaphorical perspectives at its limits. One metaphor it already produced is that of 'decolonial AI', the idea that AI has to be undone of the traces of colonial heritage (Mohamed, Png & Isaac, 2020). Similarly, another metaphor generated by AI colonialism is that of the 'wretched of AI', denoting the citizens of the global South disproportionately bearing AI's burdens, analogous to Frantz Fanon's 'wretched of the Earth' (Fanon, 1963; Mollema, 2024a). Also, in conjunction with the approach of 'algorithmic reparation', a decolonial approach from computer science (Davis, Williams & Yang, 2021), it could for example lead to the development of the complex metaphor 'algorithmic reparations for AI colonialism': the enactment of reparative justice *by means of AI* for the harms *specific to AI colonialism*. Another idea that was previously touched upon is AI colonialism's 'colonization without a colonizer'. This produces the metaphor of 'emergent colonialism', which incorporates an analogy to biological emergence by transposing it onto a socio-political analogous and bears resemblance to actor-network theory and process ontology's emergence of aspects of agency from the dynamics of processes which can't themselves be deemed 'intentional' in any ordinary sense. This brief exposition showed that AI colonialism has potential for generating new metaphors.

What is the takeaway from this discussion of the metaphor's epistemic success? The answers to Kompa's desiderata show AI colonialism's epistemic value is multifaceted. On the one hand, multiple disanalogies and constraining aspects where identified, but those are outweighed by the conclusions that AI colonialism can generate other productive metaphors, is greatly inference preserving, is in distance quite close to political reality, and can function as target and postulate for political theory with pragmatic, unifying and generative benefits. This leads me to conclude that AI colonialism is far from a superfluous metaphor. The claim to truth its perspectival change lays down is worth investigating and discloses relations between AI injustices that would otherwise escape our attention.

To recapitulate, I have argued that conceptual amelioration doesn't apply to AI colonialism because reasons for semantic or epistemic adjustment of colonialism's current use are absent. Subsequently the metaphorical interpretation was vindicated through the argument that AI colonialism is a conceptual metaphor: it enables the thinking of structural and analogical connections between the domains of AI and colonialism. Its meaning descriptively overlaps with the conceptual peripheries of colonialism and facilitates a *perspectival shift* on AI injustices, expressing how they involve harms that colonialism also involved. Based on the conceptual metaphor view, I argued why the lexicon could be expanded with AI colonialism. Finally, I showed how AI colonialism is an *epistemically successful* politico-conceptual metaphor.

6
Conclusion

The previous chapters' wide-ranging investigations led to the conclusion that AI colonialism is an epistemically successful conceptual metaphor. §2 delved into debates in political philosophy to construct operational definitions of domination, exploitation and colonialism. The 'institutional dehumanizing subversion of self-determination' was used as shorthand for the linguistically normative core of colonialism's 'mixed structure'. Colonialism's wrongful effects were found to constitute its use qua description. §3 examined the content of the charge of AI colonialism and concluded it has six faces: extractivism, objectification, exploitation, epistemic violence, cultural imposition and racialization. Based upon §2's operational definitions, §4 argued that the literal interpretation of AI colonialism is mistaken: AI systems cannot literally colonize and AI D&D doesn't instantiate colonialism's linguistically normative core. Finally, §5 defended the metaphorical interpretation, because AI colonialism is a conceptual metaphor that expresses insightful interrelations between AI and colonialism's contingent properties: its projection 'stretches and twists' colonialism to disclose a meaningful coherence in AI injustices.

The meaning of AI colonialism elucidated here isn't insignificant. By virtue of its epistemic success, AI colonialism's perspectival shift can inform future work on developing theories of AI in both political philosophy and technology studies. Also, the theorization of AI colonialism itself has only just begun. Especially the applicability of the forms of redress analogically linked to colonialism are worthwhile to research: what would it mean to 'decolonize AI' or 'repair for or through AI'? Outside of academia, AI colonialism could also prove useful. It expresses seriously harmful and deplorable states of affairs and the utility of its evocation of associations with colonialism could help make citizens and corporations of the global North notice how their actions contribute to systematic injustices related to AI in both the global South and North. Likewise, AI colonialism's patterns of inference could inform responses from civil and governmental initiatives because they give them *targets* to address uniformly and coherently: AI's economic and ecological harms are directly related to its cultural, racial, and epistemic harms. Therefore distributive countermeasures should simultaneously be thought of as recognitive countermeasures.

This impure and synthetic philosophical investigation into AI's resemblance of colonialism thereby contributes to the growing scholarly literature on identifying and theorizing responses to the injustices and harms related to AI. Staying true to its metaphorical conclusions, the foregoing philosophical expeditions may have 'left things as they are', but have hopefully made this way of viewing things both plausible and relevant. After all, without being able to *think* the coherence between AI injustices' and colonialism's respective interrelations, responses to that reality will inevitably fall short.

Bibliography

Achilleos, A. (2023). The AI Colonialism Board Game [Cyprus]. https://alexiaachilleos.com/the-ai-colonialism-board-game-cyprus/.

Adams, R. (2021). Can Artificial Intelligence Be Decolonized? *Interdisciplinary Science Reviews* 46, no.1-2: 176-197. https://doi.org/10.1080/03080188.2020.1840225.

Alammar, J. (November 2022). The Illustrated Stable Diffusion. https://jalammar.github.io/illustrated-stable-diffusion/.

Amnesty International (May 2 2023). *Israel and Occupied Palestinian Territories: Automated Apartheid: How Facial Recognition Fragments, Segregates and Controls Palestinians in the OPT*. https://www.amnesty.org/en/documents/mde15/6701/2023/en/.

Ankermit, F. R. (1993). Metaphor in Political Theory. In *Knowledge and Language: Volume III: Metaphor and Knowledge*, Ankersmit, F. R. & Mooij, J. J. A. (eds.), 155-202. Springer.

Arun, C. (2020). AI and the Global South: Designing for Other Worlds. In *The Oxford Handbook of Ethics of AI*, Dubber, M. D. et al. (eds.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.38.

Aytaç, U. (2021). On the Limits of the Political: The Problem of Overly Permissive Pluralism in Mouffe's Agonism. *Constellations* 28: 417-431. https://doi.org//10.1111/1467-8675.12525.

Aytaç, U. (2024). Digital Domination: Social Media and Contestatory Democracy. *Political Studies* 72, no.1: 6-25. https://doi.org/10.1177/00323217221096564.

Badru, R. O. (2010). Reparations for Africa: Providing Metaphysical and Epistemological Grounds of Justice to the Descendants of Dehumanised Generation. *Cultura. International Journal of Philosophy of Culture and Axiology* 7, no.2: 67-81. https://doi.org/10.5840/cultura2010725.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, 610–623. https://doi.org/10.1145/3442188.3445922.

Bennett, M. R. & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience.* Blackwell.

Barrenechea, R. & Castillo, I. (2019). The Many Roads To Rome: Family Resemblance Concepts In The Social Sciences. *Qual Quant* 53: 107-130. https://doi.org/10.1007/s11135-018-0732-7.

Birhane, A. (2020). Algorithmic Colonization of Africa. *SCRIPTed* 17, no.2: 389-409. https://scripted.org/?p=3888.

Birhane, A. (2022). The Unseen Black Faces of AI Algorithms. *Nature* 610: 451-452. https://www.nature.com/articles/d41586-022-03050-7.

Black, M. (1955). Metaphor. In *Philosophical Perspectives on Metaphor* (1981), Johnson, M. (ed.), 63-82. University of Minnesota Press.

Black, M. (1962). *Models and Metaphors.* Cornell University Press.

Black, M. (1977). More About Metaphor. *Dialectica* 31, no.3/4: 431-457. https://www.jstor.org/stable/42969757.

Brennan, J. & Jaworski, M. (2015). Markets Without Symbolic Limits. *Ethics* 125, no.4: 1053–1077. https://doi.org/10.1086/680907.

Brownlee, J. (July 19 2019). *A Gentle Introduction to Generative Adversarial Networks (GANs)*. Machine Learning Mastery. https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/.

Buolamwini, J. (June 26 2018). Fighting "the coded gaze". *Ford Foundation*. https://www.fordfoundation.org/news-and-stories/stories/fighting-the-coded-gaze/.

Butt, D. (2013a). Colonialism and Postcolonialism. In *The International Encyclopaedia of Ethics*, LaFollete, H. (ed.). Wiley/Blackwell. https://doi.org/10.1002/9781444367072.wbiee763.

Butt, D. (2013b). Inheriting Rights to Reparation: Compensatory Justice and the Passage of Time. *Ethical Perspectives* 20, no.2: 245-269. https://doi.org/10.2143/EP.20.2.2979753.

CA Global Finance. (August 23 2023). What Is AI 'Colonialism', And How Are African Tech Startups Fighting Against It? *CA Global Finance Blog*. https://www.banking-recruitment-jobs.com/en/what-is-ai-colonialism-and-how-are-african-tech-startups-fighting-against-it/.

Calvino, I. (2013). *Letters 1941-1985*. Trans. McLaughlin, M. Originally published 2000. Penguin.

Calvino, I. (2023). Montezuma and Cortés. In *The Written World and the Unwritten World*, 264-278. Trans. Goldstein, A. Originally published 1974. Penguin.

Cant, C. (November 23 2023). Be Warned: Deliveroo's Victory Over Its Riders Shows Just How Vulnerable British Workers Are. *The Guardian*. https://www.theguardian.com/commentisfree/2023/nov/23/deliveroo-victory-riders-british-workers-gig-economy-europe.

Cappelen, H. (2018). *Fixing Language: an Essay on Conceptual Engineering*. Oxford University Press.

Cappelen, H. (2020). Conceptual Engineering: The Master Argument. In *Conceptual Engineering and Conceptual Ethics*, Burgess, A., Cappelen, H. & Plunkett, D. (eds.), 132-151. Oxford University Press. https://doi.org/10.1093/oso/9780198801856.003.0007.

Cave, S. & Dihal, K. (2020). The Whiteness of AI. *Philosophy & Technology* 33: 685–703 https://doi.org/10.1007/s13347-020-00415-6.

Chekhov, A. P. (1958). *Verzamelde werken – Deel V: Verhalen 1880-1903*. G.A. van Oorschot.

Cicerchia, L. (2022). Structural Domination In the Labor Market. *European Journal of Political Theory* 21, no.1: 4-24. https://doi.org/10.1177/1474885119851094.

Couldry, N. & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*. Stanford University Press. https://www.sup.org/books/title/?id=28816.

Couldry, N. & Mejias, U. A. (2023). The Decolonial Turn in Data and Technology Research: What Is at Stake and Where is It Heading? *Information, Communication & Society* 26, no.4: 786-802. https://doi.org/10.1080/1369118X.2021.1986102.

Couldry, N. & Mejias, U. A. (May 1 2024). Today's Colonial "Data Grab" is Deepening Global Inequalities. *London School of Economics Inequalities*. https://blogs.lse.ac.uk/inequalities/2024/05/01/todays-colonial-data-grab-is-deepening-global-inequalities/.

Crawford, K. & Joler, V. (2018) Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources. *AI Now Institute and Share Lab*. https://anatomyof.ai/.

Crawford, K. (2021). *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence*. Yale University Press. https://yalebooks.yale.edu/book/9780300264630/atlas-of-ai/.

Crawford, K. (February 22 2024). Generative AI's Environmental Costs are Soaring — and Mostly Secret. *Nature* 626: 693. https://www.nature.com/articles/d41586-024-00478-x.

Cyphert, A. B. (2024). Generative AI, Plagiarism, and Copyright Infringement in Legal Documents. *Minnesota Journal of Law, Science & Technology* 25, no.2: 49-65. https://scholarship.law.umn.edu/mjlst/vol25/iss2.

Danto, A. C. (1993). Metaphor and Cognition. In *Knowledge and Language: Volume III: Metaphor and Knowledge*, Ankersmit, F. R. & Mooij, J. J. A. (eds.), 21-35. Springer. https://doi.org/10.1007/978-94-011-1844-6_2.

Davis, J. L., Williams, A. & Yang, Y. W. (2021). Algorithmic Reparation. *Big Data & Society* 8, no.2: 1-12. https://doi.org/10.1177/20539517211044808.

De Dijn, A. (2020). *Freedom: An Unruly History*. Harvard University Press.

De Man, P. (1978). The Epistemology of Metaphor. *Critical Inquiry* 5, no.1: 13-30. https://www.jstor.org/stable/1342975.

De Pina Cabral, J. (2011). What Is An Institution? *Social Anthropology/Anthropologie Sociale* 19, no.4: 477–494: https://doi.org/10.1111/j.1469-8676.2011.00173.x.

De Volkskrant. (May 31 2024). Google reageert op lijm-op-pizza-gate en belooft beterschap. https://www.volkskrant.nl/tech/liveblog-nieuws-ai-chatgpt-kunstmatige-intelligentie~bf533c99/?referrer=https://www.google.nl/.

Dennett, D. C. (2017). *From Bacteria to Bach and Back*. Allen Lane.

Dobrin, S. (February 6 2024). AI and the Risk of Technological Colonialism. *AI Business*. https://aibusiness.com/responsible-ai/ai-and-the-risk-of-technological-colonialism#close-modal.

Dussel, E. (1985). *The Philosophy of Liberation*. Orbis Books.

Eke, D. E., Wakunuma, K. & Akintoye, S. (eds.) (2023). *Responsible AI in Africa: Challenges and Opportunities*. Palgrave Macmillan. https://link.springer.com/book/10.1007/978-3-031-08215-3.

El Naqa, I. & Murphy, M. J. (2015). What Is Machine Learning? In *Machine Learning in Radiation Oncology*, El Naqa, I., Li, R. & Murphy, M. (eds.). Springer. https://doi.org/10.1007/978-3-319-18305-3_1.

Ervas, F. & Gola, E. (2013). The Pragmatics Of Metaphor Use From The Conceptual View To The Relevance-Theoretic Perspective. In *Metaphor in Focus: Philosophical Perspectives on Metaphor Use*, Gola, E. & Ervas, F. (eds.) Cambridge Scholars Publishing.

Fanon, F. (1963). *The Wretched of the Earth*. Trans. Farrington, C. Grove Press.

Fanon, F. (2008). *Black Skin, White Masks*. Trans. Markmann, C. L. Pluto Press.

Forrester, K. (2019). Reparations, History and the Origins of Global Justice. In *Empire, Race and Global Justice*, Bell, D. (ed.), 22-51. Cambridge University Press. https://scholar.harvard.edu/files/katrinaforrester/files/forrester_reparations_history_and_the_origins_of_global_justice.pdf.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001.

Gädeke, D. (2020). Does a Mugger Dominate? Episodic Power and the Structural Dimension of Domination. *Journal of Political Philosophy*, 28: 199-221. https://doi.org/10.1111/jopp.12202.

Gausen, A. (2023). Colonialism and AI. *Accessible AI*. www.accessible-ai.co.uk.

Getachew, A. (2019). *Worldmaking after Empire: The Rise and Fall of Self-Determination*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691179155/worldmaking-after-empire.

Gill, J. H. (1979). Wittgenstein and Metaphor. *Philosophy and Phenomenological Research* 40, no.2: 272-284. https://www.jstor.org/stable/2106323.

Glissant, É. (1989). *Caribbean Discourse*. Trans. Dash, J. M. University Press of Virginia.

Glissant, É. (1997). *Poetics of Relation*. Trans. Wing, B. The University of Michigan Press.

Grynbaum, M. M. & Mac, R. (December 27 2023). The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. *The New York Times*. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

Gwagwa, A., Kazim, E., & Hilliard, A. (2022). The Role of the African Value of Ubuntu in Global AI Inclusion Discourse: A Normative Ethics Perspective. *Patterns* 3: 1-7. https://doi.org/10.1016/j.patter.2022.100462.

Gwagwa, A. & Mollema, W. J. T. (2023). Afrikas Digitale Dilemma. *Udenrigs* IIII. https://www.udenrigs.dk/afrikas-digitale-dilemma/.

Gwagwa, A. & Townsend, B. (May 10 2023). Re-imagining Africa's Sovereignty in a Digitally Interdependent World. *Global Policy Journal*. https://www.globalpolicyjournal.com/blog/10/05/2023/re-imagining-africas-sovereignty-digitally-interdependent-world.

Hao, K. (July 31 2020). The Problems AI has Today Go back Centuries. *MIT Technology Review*. https://www.technologyreview.com/2020/07/31/1005824/decolonial-ai-for-everyone/amp/.

Hao, K. (April 19 2022). Artificial Intelligence Is Creating a New Colonial World Order. *MIT Technology Review*. https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/.

Harari, Y. N. (October 2018). Why Technology Favors Tyranny. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/.

Harari, Y. N. (May 6 2023). Yuval Noah Harari Argues that AI has Hacked the Operating System of Human Civilisation. *The Economist*. https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation.

Harrisberg, K. (August 9 2023). African Tech Startups Take Aim at AI 'Colonialism'. *The Star* https://www.thestar.com.my/tech/tech-news/2023/08/09/african-tech-startups-take-aim-at-ai-colonialism.

Haslanger, S. (2020). Going On, Not in the Same Way. In *Conceptual Engineering and Conceptual Ethics*, Burgess, A., Cappelen, H. & Plunkett, D. (eds.), 230-260. Oxford University Press. https://doi.org/10.1093/oso/9780198801856.003.0012.

Haugaard, M. (2010). Power: A 'Family Resemblance' Concept. *European Journal of Cultural Studies* 13, no.4: 419-438. https://doi.org//10.1177/1367549410377152.

Hesse, M. (1988). The Cognitive Claims of Metaphor. *The Journal of Speculative Philosophy* 2, no.1: 1-16. https://www.jstor.org/stable/25668224.

Hoeksema, B. (2023). Digital Domination and the Promise of Radical Republicanism. *Philosophy & Technology* 36, no.17: 1-20. https://doi.org/10.1007/s13347-023-00618-7.

Hopster, J. K. G., Brey, P., Klenk, M., Löhr, G., Marchiori, S., Lundgren, B. & Scharp, K. (2023). Conceptual Disruption and the Ethics of Technology. In *Ethics of Socially Disruptive Technologies*, Van de Poel, I., Frank, L., Hermann, J., Hopster, J., Lenzi, D., Nyholm, S., Taebi, B., & Ziliotti, E. (eds.), 141-162. Open Book Publishers. https://doi.org/10.11647/OBP.0366#resources.

Hopster, J. K. G. & Löhr, G. (2023). Conceptual Engineering and Philosophy of Technology: Amelioration or Adaptation? *Philosophy & Technology* 36: 70. https://doi.org/10.1007/s13347-023-00670-3.

Howard-Hassman, R. E. (2004). Reparations to Africa and the Group of Eminent Persons. *Cahiers d'études africaines* 173-174: 81-97. https://doi.org/10.4000/etudesafricaines.4543.

Johnson, M. (1981). Introduction: Metaphor in the Philosophical Tradition. In *Philosophical Perspectives on Metaphor* (1981), Johnson, M. (ed.), 3-47. University of Minnesota Press.

Joler, V. (2020). New Extractivism. extractivism.online.

Jung, M. (2023). Digital Capitalism Is a Mine not a Cloud. *Transnational Institute*. State of Power 2023. https://www.tni.org/en/article/digital-capitalism-is-a-mine-not-a-cloud.

Karpathy, A. (December 23 2023). [1hr Talk] Intro to Large Language Models. YouTube. https://www.youtube.com/watch?v=zjkBMFhNj_g.

Katz, Y. (2020). *Artificial Whiteness*. Columbia University Press. https://cup.columbia.edu/book/artificial-whiteness/9780231194914.

Kelly, J. (May 31 2024). Google's AI Recommended Adding Glue To Pizza And Other Misinformation—What Caused The Viral Blunders? *Forbes*. https://www.forbes.com/sites/jackkelly/2024/05/31/google-ai-glue-to-pizza-viral-blunders/.

Khan, D. (July 30 2022). AI Colonialism: The Quiet Invasion of Our Minds. *The Business Standard*. https://www.tbsnews.net/tech/ai-colonialism-quiet-invasion-our-minds-467798.

Kidd, C. & Birhane, A. (2023). How AI Can Distort Human Beliefs. *Science* 380, no.6651: 1222-1223. https://www.science.org/doi/10.1126/science.adi0248.

Klein, E. (April 12 2024). What if Dario Amodei Is Right About A.I.? *The New York Times*. https://www.nytimes.com/2024/04/12/opinion/ezra-klein-podcast-dario-amodei.html?showTranscript=1.

Kövecses, Z. (2016). Conceptual Metaphor Theory. In *The Routledge Handbook of Metaphor and Language*, Semino, E., & Demjén, Z. (eds.). Routledge. https://doi.org/10.4324/9781315672953.

Kövecses, Z. (2017). Conceptual Metaphor Theory: Some New Proposals. *LaMiCuS* 1, no.1: 16-32.

Krishnan, A., Abdilla, A., Moon, A. J., Souza, A. S., Adamson, C., Lach, E. M., Ghazal, F., Fjeld, J., Taylor, J., et al. (2023). Decolonial AI Manyfesto. https://manyfesto.ai/.

Kompa, N. (2021). Insight by Metaphor – The Epistemic Role of Metaphor in Science. In *Physics and Literature*, Heydenreich, A. & Mecke, C. (eds.), 23-48. De Gruyter. https://doi.org/10.1515/9783110481112-002.

Kwet, M. (2019). Digital Colonialism: US Empire and the New Imperialism in the Global South. *Race & Class* 60, no.4: 1-20. https://doi.org/10.1177/0306396818823172.

Lakoff, G. (1993). The Contemporary Theory of Metaphor. In *Metaphor and Thought*, Ortony, A. (ed.), 2nd ed., 202-251. Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.013.

Lakoff, G.& Johnson, M. (1980). Conceptual Metaphor in Everyday Language. In *Philosophical Perspectives on Metaphor* (1981), Johnson, M. (ed.), 286-325. University of Minnesota Press.

Lakoff, G. & Johnson, M. (2003). *Metaphors We Live By*. University of Chicago Press.

Lambrechts, W., Sinha, S. & Mosoetsa, S. (2022). Colonization by Algorithms in the Fourth Industrial Revolution. *Institute of Electrical and Electronics Engineers*: 11057-11064. https://doi.org/10.1109/access.2022.3145236.

Latour, B. (2004). Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical Inquiry* 30: 225-248. http://www.bruno-latour.fr/sites/default/files/89-CRITICAL-INQUIRY-GB.pdf.

Lehuedé, S. (March 19 2024).What is Colonial about AI and the Climate Crisis? *AI & Climate Crisis*. Talk. Goldsmiths, University of London. https://goldsmiths.cloud.panopto.eu/Panopto/Pages/Viewer.aspx?id=609824f6-5e69-4241-9bd4-b13a00d3d05b.

Madianou, M. (2019). Technocolonialism: Digital Innovation and Data Practices in the Humanitarian Response to Refugee Crises. *Social Media + Society* 5, no.3: 1-13. https://doi.org/10.1177/2056305119863146.

Marx, K. (2015). *Grundrisse*. Trans. Nicolaus, M. Original work published in 1939–41 [1857-61]. Marxists Internet Archive. https://www.marxists.org/archive/marx/works/1857/grundrisse/index.html.

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). The AI Index 2023 Annual Report. AI Index Steering Committee. *Institute for Human-Centered AI*. Stanford University. https://arxiv.org/pdf/2310.03715.

Matthias, A. (2004). The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata. *Ethics and information technology* 6, no.3: 175–183. https://doi-org.proxy.library.uu.nl/10.1007/s10676-004-3422-1.

Mbembe, A. (1992). Provisional Notes on the Postcolony. *Journal of the International African Institute* 62, no.1: 3-37. https://doi.org/10.2307/1160062.

Mbembe, A. (2015). Decolonizing Knowledge and the Question of the Archive. *Platform for Experimental Collaborative Ethnography (PECE)*. https://worldpece.org/content/mbembe-achille-2015-%E2%80%9Cdecolonizing-knowledge-and-question-archive%E2%80%9D-africa-country.

Mbembe, A. (2017). *Critique of Black Reason*. Trans. Dubois, L. Duke University Press.

Mbembe, A. (2019). *Out of the Dark Night: Essays on Decolonization*. Columbia University Press. https://cup.columbia.edu/book/out-of-the-dark-night/9780231160285.

Mbembe, A. (2020). Futures of Life and Futures of Reason. *Public Culture* 33, no.1. https://doi.org/10.1215/08992363-8742136.

Mbembe, A. (2022). *The Earthly Community: Reflections on the Last Utopia*. Trans. Corcoran, S. V2_ Lab for the Unstable Media, 2022. https://store.v2.nl/products/the-earthly-community-reflections-on-the-last-utopia.

McQuillan, D. (2015). Algorithmic States of Exception. *European Journal of Cultural Studies* 18, no.4-5: 1-11. https://doi.org/10.1177/1367549415577389.

McQuillan, D. (2022). *Resisting AI: An Anti-fascist Approach to Artificial Intelligence*. Bristol University Press, 2022. https://bristoluniversitypress.co.uk/resisting-ai.

McQuillan, D. (2023a). We Come to Bury ChatGPT, not to Praise It. https://www.danmcquillan.org/chatgpt.html.

McQuillan, D. (2023b). House of Lords Communications and Digital Select Committee Inquiry: Large Language Models. LLM0015: 1-10. https://committees.parliament.uk/writtenevidence/124038/pdf/.

Mhlambi, S. & Tiribelli, S. (2023). Decolonizing AI Ethics: Relational Autonomy as a Means to Counter AI Harms. *Topoi* 42: 867–880. https://doi.org/10.1007/s11245-022-09874-2.

Mignolo, W. D. (2020). Colonial Semiosis and Decolonial Reconstitutions. *Echo* 2: 8-15. https://doi.org/10.15162/2704-8659/1204.

Miller, E. F. (1979). Metaphor and Political Knowledge. *The American Political Science Review* 73, no.1: 155-170. https://www.jstor.org/stable/1954738.

Mills, C. W. (1997). *The Racial Contract*. Cornell University Press.

Mills, C. W. (1998). *Blackness Visible: Essays on Philosophy and Race*. Cornell University Press. https://www.jstor.org/stable/10.7591/j.ctt1tm7j79.

Mills, C. W. (2017). *Black Rights/White Wrongs: The Critique of Racial Liberalism*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190245412.001.0001.

Mohamed, S., Png, M. T. & Isaac, W. (2020). Decolonial AI: Decolonial theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology* 33, no.4: 659–684. https://doi.org/10.1007/s13347-020-00405-8.

Mollema, W. J. T. (2024a). Decolonial AI as Disenclosure. *Open Journal of Social Sciences* 12, no.2: 574-603. https://doi.org/10.4236/jss.2024.122032.

Mollema, W. J. T. (2024b). Responding to the Watson-Sterkenburg Debate on Clustering Algorithms and Natural Kinds. *Unpublished*. https://philpapers.org/rec/THORTT-4.

Moore, M. (2016). Justice and Colonialism. *Philosophy Compass* 11: 447–461. https://doi.org/10.1111/phc3.12337.

Muldoon, J. & Wu, B. A. (2023). Artificial Intelligence in the Colonial Matrix of Power. *Philosophy & Technology* 36: 80. https://doi.org/10.1007/s13347-023-00687-8.

Ndlovu-Gatsheni, S. J. (2013). *Coloniality of Power in Postcolonial Africa*. CODESRIA. https://muse.jhu.edu/book/25053.

Ndlovu-Gatsheni, S. J. (2019). Discourses of Decolonization/Decoloniality. *Papers on Language and Literature* 55, 3: 201-226. https://eref.uni-bayreuth.de/id/eprint/69213.

Nersessian, N. J. (2015) The Cognitive Work of Metaphor and Analogy in Scientific Practice. *Philosophical Inquiries* 3, no.1: 133-156. https://www.philinq.it/index.php/philinq/article/download/118/65.

Norberg, J. (2015). Concepts, Political. In *The Encyclopedia of Political Thought*, Gibbons, M. (ed.), 647-658. Wiley. https://doi.org//10.1002/9781118474396.

NWO. (November 1 2019). Artificial Intelligence Research Agenda for the Netherlands. https://www.nwo.nl/sites/nwo/files/documents/AIREA-NL%20AI%20Research%20Agenda%20for%20the%20Netherlands.pdf.

Okeja, U. (2022). *Deliberative Agency: A Study in Modern African Political Philosophy*. Indiana University Press.

Pauwels, E. (2020). *The Anatomy of Information Disorders in Africa*. Konrad-Adenauer-Stiftung. https://www.kas.de/en/web/newyork/single-title/-/content/the-anatomy-of-information-disorders-in-africa.

Pasquinelli, M. (2019). Three Thousand Years of Algorithmic Rituals: The Emergence of AI from the Computation of Space. *E-flux Journal* 101. https://www.e-flux.com/journal/101/273221/three-thousand-years-of-algorithmic-rituals-the-emergence-of-ai-from-the-computation-of-space/.

Pasquinelli, M. & Joler, V. (2021). The Nooscope Manifested: AI as an Instrument of Knowledge Extractivism. *AI & Society* 36: 1263–1280. https://doi.org/10.1007/s00146-020-01097-6.

Perrigo, B. (January 18 2023). Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic. *Time*. https://time.com/6247678/openai-chatgpt-kenya-workers/.

Pettit, P. (1999). *Republicanism: A Theory of Freedom and Government*. Oxford University Press. https://doi.org/10.1093/0198296428.001.0001.

Pettit, P. (2007). Responsibility Incorporated. *Ethics*, 117, no.2: 171–201. https://doi.org/10.1086/510695.

Pettit, P. (2013). *On the People's Terms: A Republican Theory and Model of Democracy*. https://doi.org/10.1017/CBO9781139017428.

Pitkin, H. P. (1972). *Wittgenstein and Justice.* University of California Press.

Queloz, M. & Cueni, D. (2021). Left Wittgensteinianism. *European Journal of Philosophy* 29, no.4: 758-77. https://doi.org/10.1111/ejop.12603.

Quijano, A. (2000). Power, Eurocentrism and Latin America. Trans. Ennis, M. *Nepantla: Views from South* 1, 3: 533-580. https://muse.jhu.edu/article/23906.

Rachovitsa, A. & Johann, N. (2022). The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case. *Human Rights Law Review* 22, no.2. https://doi.org/10.1093/hrlr/ngac010.

Renzo, M. (2019). Why Colonialism is Wrong. *Current Legal Problems*, 72, no.1: 347–373. https://doi.otg/10.1093/clp/cuz011.

Ricaurte, P. (2019). Data Epistemologies, The Coloniality of Power, and Resistance. *Television & New Media* 20, no.4: 350-365. https://doi.org/10.1177/1527476419831640.

Ricaurte, P. (2022). Ethics for the Majority World: AI and the Question of Violence at Scale. *Media, Culture & Society* 44, no.4: 726–745. https://doi.org/10.1177/01634437221099612.

Ricoeur, P. (1979). The Metaphorical Prosses as Cognition, Imagination and Feeling. In *Philosophical Perspectives on Metaphor* (1981), Johnson, M. (ed.), 228-247. University of Minnesota Press.

Roemer, J. E. (1982). *A General Theory of Exploitation and Class.* Harvard University Press.

Roemer, J. E. (1985). Should Marxists be Interested in Exploitation? *Philosophy & Public Affairs* 14, no.1: 30-65. https://www.jstor.org/stable/2265236.

Rosenthal, D. C. (1982). Metaphors, Models, and Analogies in Social Science and Public Policy. *Political Behavior* 4, no.3: 283-301. https://www.jstor.org/stable/586442.

Roy, A. (December 12 2020). *Introduction To Autoencoders*. Towards Data Science. https://towardsdatascience.com/introduction-to-autoencoders-7a47cf4ef14b.

Sabera, T. (2021). The Moral Wrong of Colonialism: A Non-Instrumentalist Approach. *Global Journal Of Human-Social Science: Arts & Humanities – Psychology* 12, no.14: 61-65. https://doi.org/10.34257/GJHSSAVOL21IS14PG61.

Sahbaz, U. (2019). Artificial Intelligence and the Risk of New Colonialism. *Horizons* 14: 58-70. https://www.cirsd.org/en/horizons/horizons-summer-2019-issue-no-14/artificial-intelligence-and-the-risk-of-new-colonialism.

Santos, B. de Sousa. (2014). *Epistemologies of the South: Justice Against Epistemicide.* Routledge. https://unesco-chair-cbrsr.org/pdf/resource/Epistemologies_of_the_South.pdf.

Sharma, R., Sharma, K. & Khanna, A. (2020). Study of Supervised Learning and Unsupervised Learning. *International Journal for Research in Applied Science & Engineering Technology* 8, no.6: 588-593. http://doi.org/10.22214/ijraset.2020.6095.

Stilz, A. (2015). Decolonization and Self-determination. *Social Philosophy and Policy* 32, no.1: 1-24. https://doi.org/10.1017/S0265052515000059.

Suchman, L. & Weber, J. (2016) Human-Machine Autonomies. In *Autonomous Weapons Systems: Law, Ethics, Policy*, Bhuta, N., Beck, S., Geiβ, R., Liu, H.-Y. & Kreβ, C. (eds.), 75–102. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781316597873.004.

Taylor, J. (January 23 2024). Precarious Conditions of AI 'Ghost Workers' Revealed by Google Termination of Appen Contract, Union Says. *The Guardian.* https://www.theguardian.com/australia-news/2024/jan/23/precarious-conditions-of-ai-ghost-workers-revealed-by-google-termination-of-appen-contract-union-says.

Taylor, C. & Dewsbury, B. M. (2018). On the Problem and Promise of Metaphor Use in Science and Science Communication. *Journal of Microbiology & Biology Education* 19, no.1: 1-5. https://doi.org/10.1128/jmbe.v19i1.1538.

Taiwo Afisi, O. (2024). Neocolonialism. *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, https://iep.utm.edu/.

Timcke, S. (July 11 2024). AI and the Digital Scramble For Africa. *Review of African Political Economy.* https://roape.net/2024/07/11/ai-and-the-digital-scramble-for-africa/.

Ugar, E. T. (2023). The Fourth Industrial Revolution, Techno-Colonialism, and the Sub-Saharan Africa Response. *Filosofia Theoretica: Journal of African Philosophy, Culture and Religions* 12, no.1: https://doi.org/10.4314/ft.v12i1.3.

UN News (January 2 2024). INTERVIEW: AI expert warns of 'digital colonization' in Africa. *UN News.* https://news.un.org/en/story/2024/01/1144342.

Valentini, L. (2015). On the Distinctive Procedural Wrong of Colonialism. *Philosophy & Public Affairs* 43, no.4: 312-331. https://doi.org/10.1111/papa.12057.

Van Wietmarschen, H. (2018). The Colonized and the Wrong of Colonialism. *Thought* 7: 170-178. https://doi.org/10.1002/tht3.381.

Van de Poel, I. (2020). Embedding Values in Artificial Intelligence (AI) Systems. *Minds & Machines*, 30 (2020): 385-409. https://doi.org/10.1007/s11023-020-09537-4.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention Is All You Need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*: 1-15. https://arxiv.org/abs/1706.03762.

Vrousalis, N. (2013). Exploitation, Vulnerability and Social Domination. *Philosophy & Public Affairs* 41, no.2: 131-157. https://doi.org/10.1111/papa.12013.

Vrousalis, N. (2018). Exploitation: A Primer. *Philosophy Compass* 13, no.2: e12486. https://doi.org/10.1111/phc3.12486C.

Watson, D. S. (2023). On the Philosophy of Unsupervised Learning. *Philosophy & Technology* 36: 28. https://doi.org/10.1007/s13347-023-00635-6.

Weirich, K. (April 11 2024). Military AI and the Illusion of Authority. *The Prindle Institute for Ethics.* https://www.prindleinstitute.org/2024/04/military-ai-and-the-illusion-of-authority/.

Williams, B. (2005). Pluralism, Community and Left Wittgensteinianism. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*, Hawthorn, G. (ed.), 25-33. Princeton University Press. https://www.jstor.org/stable/j.ctt7rhmz.

Williams, A., Miceli, M. & Gebru, T. (October 31 2022). The Exploited Labor Behind Artificial Intelligence. *Noema Magazine*. https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence/.

Wiredu, K. (2002). Conceptual Decolonization as an Imperative in Contemporary African Philosophy: Some Personal Reflections. *Rue Descartes* 36, no.2: 53-64. https://doi.org/10.3917/rdes.036.0053.

Wittgenstein, L. (1975). *On Certainty*. Anscombe, G. E. M. & von Wright, G. H. (eds.). Trans. Paul, D. Originally published in 1969. Blackwell.

Wittgenstein, L. (2009). *Philosophical Investigations*. Trans. Anscombe, G. E. M., Hacker, P. M. S. & Schulte, J. Originally published in 1953. Wiley-Blackwell.

Woolridge, M. (December 23 2023). The Turing Lectures: The Future of Generative AI. *The Alan Turing Institute*. Lecture. https://www.youtube.com/watch?v=2kSl0xkq2lM.

Ylikoski, P. & Kuorikoski, J. (2010). Dissecting Explanatory Power. *Philosophical Studies* 148: 201-219. https://doi.org.10.1007/s11098-008-9324-z.

Yoon, S. (April 25 2023). Artificial General Intelligence Risks a Return to Cultural Colonialism. *VentureBeat*. https://venturebeat.com/games/artificial-generative-intelligence-risks-a-return-to-cultural-colonialism/.

Young, I. M. (1990). *Justice and the Politics of Difference*. Princeton University Press, https://doi.org/10.2307/j.ctvcm4g4q.

Young, I. M. (2005). Responsibility and Global Justice: A Social Connection Model. *Anales de la Cátedra Fransisco Suárez* 39: 709-726. https://revistaseug.ugr.es/index.php/acfs/article/download/1040/1235/1767.

Ypi, L. (2013). What's Wrong with Colonialism. *Philosophy & Public Affairs* 41, no.2: 158-191. https://doi.org/10.1111/papa.12014.

Zimeta, M. (September 29 2023). Why AI Must be Decolonized to Fulfill its True Potential. *Chatham House*. https://www.chathamhouse.org/publications/the-world-today/2023-10/why-ai-must-be-decolonized-fulfill-its-true-potential.

Zimmermann, A. & Lee-Stronach, C. (2022). Proceed with Caution. *Canadian Journal of Philosophy* 52: 6-25. https://doi.org/10.1017/can.2021.17.