

Seeing economic development like a large language model. A methodological approach to the exploration of geographical imaginaries in generative AI

Boris Michel^{*}, Yannick Ecker

Martin-Luther-Universität Halle-Wittenberg, Germany

ABSTRACT

The recent hype surrounding the disruptive potential of AI technologies in the form of large language models or text to image generators also raises questions for geographical research and practice. These questions include the power relations and inequalities inscribed in these systems, their significance for work and labor relations, their ecological and economic impact, but also the geographical and spatial imaginaries they reproduce. This article focuses on the latter and formulates a series of theoretical and methodological considerations for dealing with the output of these systems. As we assume that outputs generated by large language models will play an increasing role in the future, both in public and media discourses as well as in the discourses and practices of spatial planning and economic policy making, we consider it important to gain a critical understanding of these socio-technical systems. The empirical object of investigation of this paper is generated output that deals with questions of regional development and economic challenges in three European regions that are currently particularly affected by the transition to a climate-neutral economy and are designated by the European Union as Just Transition Fund Territories. We are particularly interested in how geographical imaginaries about these regions are formulated, how economic and social problems of these regions are presented and how this is translated into planning advice and development plans.

1. Introduction

“I can deliver comprehensive data analysis, economic modeling, and strategic recommendations to support informed decision-making and sustainable development in regional planning and economic policy making.”¹

The recent hype surrounding the disruptive potential of AI technologies in the form of large language models (LLM) or text-to-image generators (TTI) also raises questions and challenges for geographical research and practice. These include questions about the power relations, dispossessions and inequalities inscribed in the data used to train these systems, the way these systems change geographies of labor exploitation and capital accumulation as well as their ecological, infrastructural and economic impact (Roberge and Castelle 2021; Bode and Goodlad 2023; Raley and Rhee 2023; Mackenzie 2017; Morreale et al. 2023; Lehuédé 2024; Thatcher, O’Sullivan, and Mahmoudi 2016). Assuming a growing role of these models in everyday life, academic and professional labor as well as its integration as “infrastructural platforms” (van Dijck et al. 2018) within third party services and devices, we expect that LLMs and the outputs they generate will gain importance in the

cultural political economy of digital capitalism. From a geographical standpoint, the question thus arises as to the nature of the geographic “knowledge” being generated and the inherent reproduction of spatial and economic imaginaries.

While we are sympathetic towards calls to reject these systems in their entirety or at least to plead for a very cautious use, we consider it important to gain a critical understanding of these socio-technical systems. We believe it is important to look at the way these statistical models generate texts that appear as plausible outputs on geographical topics. It is not just a matter of accusing these models of occasionally “getting it wrong” or of hallucinating. Nor is it sufficient to hastily adopt the belief in their disruptive potential. Rather, it is imperative to gain a more profound understanding of these systems, their modes of text production and the data embedded in these models. It is therefore our aim to contribute to a more empirical and methodological debate in the hype surrounding AI and its utopian and dystopian extremes. For this, we connect the current discussion about CriticalAI (Raley and Rhee 2023; Bode and Goodlad 2023; Roberge and Castelle 2021; Broussard 2019; Crawford 2021) with a cultural political economy perspective to complement existing research on the making of LLMs with an analysis of ideological implications of their outputs (Sum and Jessop 2013; Jessop

^{*} Corresponding author.

E-mail addresses: boris.michel@geo.uni-halle.de (B. Michel), yannick.ecker@geo.uni-halle.de (Y. Ecker).

¹ ChatGPT 4, prompt „what can you do for regional planning and economic policy making in one sentence“ 07/10/2024.

2010).

In this paper, we are interested in the specific layer of machinic selectivity inherent in the probabilistic construction of spatial and economic “imaginaries” (Jessop 2010) – a concept that we prefer to the term of knowledge, since the latter implies human thinking and reasoning. We are particularly interested in how spatial imaginaries are incorporated, refracted and reproduced by LLMs, how economic and social problems of regions are presented and how this is translated into advice and policies. Furthermore, we are also very much interested in the way these systems represent the discipline of geography, what approaches and theories they refer to and how those are described. We want to get a sense of what data and knowledge are inscribed in these models, how weightings and alignments have been made and what implications this has for the relationship between these language models and the world. In order to approach this empirically and methodologically, we will analyze LLM-generated outputs in this paper. The approach remains necessarily experimental and partly speculative, which is not least due to the probabilistic and unstable logic of these models themselves.

The empirical object of investigation of this paper is generated output that deals with questions of regional development and economic challenges of peripheralized, economically disadvantaged and stigmatized regions. Its focus are three European regions that are currently particularly affected by the transition to a climate-neutral economy and are supported through the European Union’s Just Transition Mechanism (JTM) as part of the European Green Deal. The Midlands in Ireland, the Lausitz in East Germany and Asturias in Spain are regions with a long history of extractivist industries especially with regards to coal or peat and due to their decline are experiencing significant economic and political challenges. Choosing these regions in transition thus promises an investigation into the LLM’s representation of a politically controversial and challenged topic that is subject of extensive political interventions and addresses issues of justice and uneven geographies.

This text consists of three parts. In a first step, we will summarize some important strands of the discussions on a critical perspective on AI and LLMs. In particular, we are interested in discussions in geography and methodological debates on how to deal with the output of these systems. We will then outline our methodological implementation and research design. Therein, we introduce the idea of imaginaries as a sensitizing concept and propose a combination of qualitative content analysis with quantitative elements as an explorative method for analyzing LLM output. The main section presents the results of our analysis. In the conclusion, we contextualize our findings against the backdrop of the larger political economy of LLMs and explore the refraction of disciplinary knowledge as an issue of critique.

2. Researching large language models

Even if discussions about AI and society have a longer history that goes back at least to the early days of digital computers (Pasquinelli 2023; Tenen 2024), the current interest has largely been triggered and shaped by the publication and widespread use of general-purpose LLMs such as GPT-4 or TTI generators such as Stable Diffusion since the early 2020 s. In the public and academic discourse, these systems are often described with the term AI. Since AI is primarily a marketing term that attributes the technology involved a sense of disruptive agency and mixes different applications and tools as well as different understandings of “artificial” and “intelligence” and implies far-reaching assumptions about the future of humanity in general, there are good reasons to use the term AI cautiously. Instead of engaging in the ahistorical discussion about the disruptive potential of AI and its existential risks (Bender 2023) it is more fruitful to fall back on more modest terms and to understand these as socio-technical systems that are part of diverse assemblages of human and non-human actors, economic, social and political relations (Crawford 2021). With regards to the tools of interest here, instead of addressing them as artificial intelligence it is more

helpful to speak of machine learning and LLMs and to avoid metaphors and terms that suggest that these systems and their outputs should be described with terms such as “knowledge”, “understanding”, “cognition” or “consciousness”, but also “hallucinating” or “lying”.

LLMs became prominent within the discourse around AI in the late 2010 s and are computational systems designed to generate “human-like” text based on methods of machine learning within the field of natural language processing. These models are socio-technical systems that generate plausible outputs based on large data sets, computing power and human work of classification. When used, they generate a “probable” response to a specific prompt. LLMs are generally regarded as a significant advance for tasks of automated text generation, summarization, translation, and sentiment analysis. They are linguistic probability models that contain statistical relationships between words, or more precisely between tokens and they generate outputs that correlate with given prompts. As Bender et al. (2023) write in their highly influential paper “On the Dangers of Stochastic Parrots”, a LLM is a “system for haphazardly stitching together sequences of linguistic forms [...] according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot”. With regards to the interpretation and handling of the texts generated by these LLMs it is important to keep in mind that these models are models of language. “The program does not have a model of the world; it has only a model of language. Everything it knows comprises words (technically, tokens) expressed as vectors and manipulated mathematically through matrices, generating correlations associated with probabilities that are then output as words” (Hayles 2023, 258). The fact that these systems are only systems of signifiers and have no relation to the signified, means that the accusation that a model is “hallucinating” or making something up are unfounded. As Hicks et al. put it with reference to Harry Frankfurt (Frankfurt 2009), these models are “bullshit machines”: “Because these programs cannot themselves be concerned with truth, and because they are designed to produce text that looks truth-apt without any actual concern for truth, it seems appropriate to call their outputs bullshit” (Hicks, Humphries, and Slater 2024). True or false as well as questions of meaning and consistency are no useful categories for a model that generates plausible, probable or “mean” (Steyerl 2023) texts in relation to prompts (Munn, Magee, and Arora 2023).

Thus, what characterizes the systems we are interested in here and what distinguishes them from historically different understandings of AI is their reliance on data to train the models. These models are not based on an “understanding” but on statistical distributions within existing data. As Kate Crawford writes, “This is not magic; it is statistical analysis at scale” (Crawford 2021, 215). LLMs rely on huge amounts of data for training and their power and success is thus an expression of an extensive datafication of human-generated text. While the term “data suggests something more mundane and humble: numbers, spreadsheets, or graphs, for example” than “AI” and while data “seldom comes up in the public-facing discourse of the most influential players in AI” (Bode and Goodlad 2023), a significant part of the critical discussion around AI is concerned with this data and the social assumptions, cultural and language biases and inequalities inscribed in it (Roberge and Castelle 2021; Bode and Goodlad 2023; Raley and Rhee 2023; Mackenzie 2017).² This discussion is often linked to the wider discussions about the harms of datafication and big data, especially with regards to questions of gender and race (Noble 2018; Benjamin 2019; O’Neil 2017; D’Ignazio and Klein 2020; Broussard 2023) but also with regards to questions of ownership,

² This becomes particularly clear in relation to the classification of images, as these models rely on a relationship first being established between image and text. In one early critical engagement with the “anatomy of a training set”, Crawford and Paglen look at the politics of classification and in one of the most used image datasets used for machine learning pointing out how racial and gender stereotypes, among other social imaginaries, are inscribed there (Crawford and Paglen 2021).

appropriation and “Data colonialism through accumulation by dispossession” (Thatcher, O’Sullivan, and Mahmoudi 2016).

Bearing this process of making LLMs in mind is relevant as most of these models are black boxes – both because models and datasets can be proprietary and closed, and because the process of “learning” of these systems is no longer comprehensible to human observation (Hayles 2022). Critical analysis is thus especially interested in biases within the training data and how these affect the models. Which texts were used to train the models, which social positions are represented and dominant in them and which are not? It is also relevant how these models were adjusted, weighted and moderated – mainly by human labor. From a geographical perspective, the dominance of English in digital text corpora is significant, as is the marginalization of languages that are quantitatively less prominent in digital text corpora. This leads to concerns about reproducing the Anglophone hegemony in the representation of geographical knowledge and academic discourse. It also follows from the reliance on historical training data that these language models are created on the basis of the published language use of past worlds and thus use historical conditions to generate probable statements. One way to approach these models is to study their outputs – to read the texts they create and subject them to the scrutiny of textual interpretation.

3. Methodological Considerations

Given that texts generated by LLMs challenge our idea of understanding, meaning and significance, but also of authorship and truth, analyzing imaginaries within outputs generated by LLMs requires a nuanced review of fundamental methodological principles. One key concern is whether our various existing methods of text analysis are suitable to “understand mathematically correlated language production” (Hayles 2023, 258) or whether this material requires new methods and methodologies. Although it can be argued that the author of a text and the modes of its production do not play any role for its interpretation – after all the author is dead and one can argue that we do not have to know much about the printing press in order to discuss books –, we follow Hayles’ argument, that the fact that these outputs were created by machines and not by humans plays an important role when dealing with them (Hayles 2023). At the same time, this methodological discussion is still very much in its infancy. Especially when it is conducted beyond computational linguistics and takes place in and with regards to social sciences such as geography.

As a contribution to this discourse, we apply the concept of “imaginaries” (Jessop 2010; Sum and Jessop 2013) to investigate the representation of spatial and disciplinary knowledge within LLM-generated content. Following approaches from cultural political economy (Jessop 2010), we understand “imaginaries” as specific integrations of *genres* (e.g. policy planning documents, press statements), *discourses* (e.g. regarding the “social market economy” or “sustainability”) and *styles* (start-up lingo etc.). We argue that the LLMs’ statistical engagement with linguistic forms leads them to incorporate, de- and recompose such elements and imaginaries inherent in their training data, i. e. semiotic systems constituting different social fields, framing discursive selectiveness and orienting subjective interpretations. It is not so much that these are logically consistent, closed and empirically verifiable semiotic systems. Nor, are these models capable of generating meaning. In relation to LLMs, it is often precisely the moments of failure that attract our attention. Nevertheless, we think that the concept of imaginaries can be used to describe these outputs and their apparent plausibility. Therefore, we think that it makes sense to trace how their training leads them *resemble* and *refract* imaginaries dominant in their training data. By reproducing probabilistic consistencies and semiotic operations – e.g. regarding the definition of “the economic” or framing of “transition” –, these models then reflect “sedimented meaning” (Sum and Jessop 2013, 149) and selective interpretations back at their users. Thereby, LLMs introduce another machinic layer to the “organized field of human geography”, adding to existing inequalities with respect to (in)visibility and hegemony

within disciplinary knowledge production (Jazeel 2016). Hence, while it can be considered “off” to analyze LLM output regarding its truthfulness, we can and should be concerned with the form of the additional machinic layer of “discursive selectivity” (Jessop 2010: 344) these machines have apprehended and reproduce with respect to social fields fundamental to human thinking about society and social transformation.

Based on our interest in the representation of spatial imaginaries and disciplinary knowledge in dealing with regional economic challenges and policy responses, we chose to work with the popular LLM-service ChatGPT. For this, we generated a corpus consisting of “conversations” about three selected regions and approaches from economic geography addressing their challenges. We use the term “conversation” cautiously here, considering it more neutral and less anthropomorphizing than interview or talk. These “conversations” – i.e. a series of prompts and respective outputs – were made using the default settings of ChatGPT 4o.

In the absence of established methodological standards for working with these models and for dealing with their outputs in a systematic manner, the generation of these outputs was still largely experimental and exploratory in nature:

- In addition to the regions we were interested in, we generated conversations about regions that are discussed as economically vibrant (e.g. the respective capital cities) as well as a series of questions without a specific regional reference (regarding the definition of economic geography). This was to provide a sort of baseline for these results and to gain an understanding of the context sensitivity of the models.
- We also used other versions of the model, including the older ChatGPT 3.5 version and the Co-Pilot version, which gives references and sources for its outputs (most frequently corresponding Wikipedia entries and top results from the search engine Bing). Beyond using the English version of the model tests were run with the Spanish and German version for the Spanish resp. German region.
- We also experimented with the specific prompts. We varied terms, asked for different genres, and changed the chronology of the queries as well as the mode of inquiry (e.g. having more “dialogical” “conversations” or only one more extensive prompts). This was done to get a feel for the influence of each on the results. Given the absence of any extralinguistic reference in these models, it seemed crucial to us to investigate the extent to which specific word choice affects the surrounding words and connected tokens.
- Since these models not only generate probable statements based on their training data, but also have a certain amount of randomness built in – which in some interfaces can be adjusted via the “temperature” setting – all queries were recorded repeatedly generating a corpus with 55 “conversations” in total.³ This repetition allows for a more robust speculation about variability and randomness of the outputs (e.g. “How typical is the use of the term “decline” with reference to a certain region?” and “How frequently is David Harvey referred to as a key economic geography author when assessing an economic policy?”).

After some experimentation and a preliminary evaluation of the results, we chose three standardized questions in relatively reduced language without further and more dialogical inquiries or more inputs about roles and expectations. We asked to “describe the economic geography of the region”, to “analyze how the EU Just Transition Mechanism aims to support economic development in the region” and to

³ These include 10 conversations per region (Asturias, Lausitz, Midlands), 5 conversations per corresponding capital (Madrid, Berlin, Dublin) and 10 isolated conversations about central concepts and authors of economic geography. When cited short forms of these conversations are used, e.g. ast002 would refer to conversation number 2 on Asturias.

“critically assess the application of the Just Transition Mechanism in the region using concepts and authors from economic geography”.

The rationale behind the selection of these prompts and modes of conversation was to encompass a range of dimensions for analysis. We use different linguistic operators to generate different modes of response (describe, analyze, assess). Furthermore, by addressing questions of spatial difference we inquire about the context sensitivity of these models and might even get an understanding of its construction of space and place. Lastly, the choice of regions and discipline allow us to operate in a field in which we have a certain degree of knowledge in order to be able to assess the outputs. This is in no way intended to devalue other possible questions and forms of conversation. On the contrary, we are concerned with diversity of outputs, which further reinforces the generally speculative character of these models. By using this material, we are not suggesting that many future policy papers will come straight out of LLMs, instead we are creating a simplified and standardized experimental setting to query the structure of the economic imaginaries that would come to bear even when more complex queries and parameters are used.⁴

The same openness and experimental approach that characterized the creation of our corpus also informs the methodological analysis of the outputs of these models. We broadly follow Hayles who proposes four strategies for a critical inquiry of LLM outputs. “1. Speculation about the source texts that inspired a particular response; 2. Analysis of a response’s stylistics in view of the input; 3. Analysis of how the response’s language refracts human language use; 4. Analysis of a response’s ideological implications in view of ideological biases in the source texts.” (Hayles 2022) However, a number of methodological questions arose in operationalizing these strategies. On the one hand, it seems plausible to assess stylistic and thematic particularities using quantitative methods of corpus linguistics. These are statistical texts, so it would be fitting to use statistical methods to describe them. This is especially true for the redundant outputs we generated. At the same time, qualitative methods approach texts at the level of their use by human readers and our interest lies precisely in this reading experience. While we would argue that it is not particularly convincing to apply hermeneutic approaches to reconstructing subjective *meaning* when working with machine-generated texts, we find cases for applying content analysis and a combination of quantitative and qualitative analytical elements more convincing (Dengel et al. 2023). Hence, we implemented a content analysis relying on deductively-generated categories and combined them with a quantitative focus on cooccurrences and word frequencies. We focused for example on categories such as “formatting”, “sentiment” and “ascription of agency” at the level of *style* and coded for “scope of transition”, “mentioned subject positions / actors” or “economic geographical concept” at the level of *discursive elements*. After an analysis of individual conversations, we integrated and systematized these findings across the conversations’ repetitions. It is only from this aggregation that we traced the recurring structure of the integrations of styles and discursive elements – i.e. imaginaries – reproduced in the LLM’s policy briefs and speculated about features of the source texts.

4. Findings

4.1. Spatial Imaginaries of Regions in Transition

In a first step, we will look at the construction of the three regions in transition in our material. We are interested in the way the respective regions are described, what are common motifs and what sets them

⁴ One might for example tweak the LLM-service to avoid bullet-point lists, to use web search results and consider only local geographers when analyzing a certain geographical region. However, the ensuing text would still underly the same probabilistic logics (for more detail, see chapter 4).

apart. Although all of the regions are classified by the European Union as Just Transition Fund Territories, they exhibit considerable variation in terms of their economic structure, historical context, and the nature of national policy regimes. The Lausitz region, for instance, was a prominent center for lignite mining during the German Democratic Republic (GDR) era and has since undergone a profound structural transformation. This process has entailed not only rapid deindustrialization but also substantial financial support from the federal government. The autonomous community of Asturias in northwestern Spain was a center for 19th century industrialization, largely due to its substantial coal and steel production. However, the regional transition process has been found to proceed more smoothly in the past (Bolet, Green, and Gonz ales-Eguino 2023). Besides, Asturias is the territory receiving the largest share of money through the EU’s Just Transition Fund in Spain.⁵ The Irish Midlands on the other hand is a relatively rural region where peat-based energy production rose to significant regional economic importance in the 20th century and where the pace of the transition has been discussed to undermine trust in the process (Banerjee and Schuitema 2022). Despite these differences, our focus on stylistic and discursive elements revealed a shared positivist spatial imaginary in the LLM output leveling most of the regional specificity.

All texts on our three sample regions exhibit a dominating formal structure. In general, the stylistic variability between the repeated conversations is relatively low. The typical formal structure consists of a short regional introduction, followed by a sector-based description combined with specific highlights and a closing reflection on challenges and opportunities. Almost all texts start with three sections bearing titles such as “Natural Resources and Primary Sector”, “Secondary Sector” and “Tertiary Sector”. They then usually include specific highlights as a fourth or fifth section (e.g. “Urban and Rural Dynamics”, “Energy Sector” or “Transportation and Infrastructure”). Within each of these sections, the text is ordered in lists of bullet points or numbering lists. For the capital cities, this formal structure is less rigid, mostly differentiating *within* segments of the tertiary and quaternary sector (“Technology Sector”, “Creative Industries”, “Financial Services”). Despite this at first glance systematic and analytical style, a closer look reveals inconsistencies, e.g. regarding the aggregation of activities into sectors (e.g. ast006).

The overall tone for the peripheral regions ranges from neutral to positive. For the corresponding urban centers, it is overwhelmingly positive. The economic geographies of the urban centers are typically associated – “shaped” or “characterized” – by their “strategic location” and “diverse”, “robust” or “dynamic” economies. They are described as economic “hubs” or “powerhouses” in over two thirds of the conversation and are often put into a national and European context. In contrast, the regions in transition are most frequently described passively as being “shaped” by their “past”, “industrial legacy”, “(natural) resources” and are typically not put into a supranational context:

“The interplay of natural resources, industrial diversification, and cultural heritage shapes its economic geography, presenting both challenges and opportunities for future growth.” (lau009)

“Madrid, the capital of Spain, is an economic powerhouse within the country and a significant hub in Europe.” (mad006)

Serious socio-economic problems such as “population decline” (lau002), “unemployment” (ast005) or “industrial pollution” (ast008) are addressed but are often integrated into internally balanced sections such as “Challenges” (ast006) or “Economic Challenges and Opportunities” (ast009). This neutral to positive tone is reflected in an almost in-existent ascription of agency. We can for example read of “investments in technology, research, and development [that] aim to foster innovation and attract new industries” (ast001), but virtually no actors or

⁵ https://www.transicionjusta.gob.es/es-es/Paginas/La_Transicion_Justa/Fondo-de-transicion-justa-UE2.aspx (14/08/2024).

agents are mentioned. Within the initial descriptions, “the region” or “Lausitz” are practically the only grammatical subjects exhibiting properties (e.g. touristic popularity, a robust fishing industry) and intentions. Where active voice is employed it is linked to these forms of collective subjects, such as in “Asturias is investing in renewable energy sources” (ast001).

Aside from these peculiarities with respect to form, the content of these descriptions deserve attention. We explore this by focusing on the way in which the concept of economic geography is associated with the respective regions and how the ongoing transition is described.

Regarding the first aspect, it remains obscure in what way this term is tokenized within the LLM, whether for example the relation between the two substrings – “economic” and “geography” – is weakened through the processing. When first experimenting with the models and testing our queries, we for example asked for descriptions of the “economic structure” instead of the “economic geography” and arrived at remarkably similar results, hinting towards a stronger influence of the “economic” over the following output. Within our dataset there are some hints at this uneasy integration. Half of the descriptions link the description of the primary sector to “natural resources” or “geographical setting”, pointing towards an association of “geography” with “physical geography”. In most cases, specific locational references to important places, hubs, ports and transport corridors were included. In one case, the “geographical” seems to have been disaggregated from the “economic”, being confined to an independent section “geographical factors” at the end of the sector-based description and its SWOT-analysis-like conclusion (ast007). The descriptions are thus spatial in the sense that they feature geographical location and use it to explain economic assets and sectoral importance, resembling a weak form of geo-determinism (e.g. of agriculture or mining). Temporality and history are often integrated into these descriptions (e.g. a tradition of mining), but not used as explanans for ongoing processes, for example to relate the Lausitz’s economic structure with past politics of regional specialization in the socialist GDR.

Even though we did not ask about these aspects in the initial questions, all conversations situate the three regions in transition towards “renewable energies” or “climate neutrality” and include concerns regarding “economic diversification”. In none of the cases, however, this process is linked to an explanation of why it is happening (e.g. linking it to climate change, EU integration or needs to confront global competition) – or why it is pursued by “the regions”. While the later inquiry regarding the Just Transition Mechanism (JTM) does produce texts that put these processes into the context of the “European Green Deal” and its aim “to support regions and communities that are most affected by the transition towards a climate-neutral economy” (ast003), no rationales are given for the necessity of a transition as such. In line with this reservation towards the explanation of ongoing processes, few actors and subjectivities are mentioned. These include “the workforce” – frequently associated with “reskilling” –, “affected communities/stakeholders”, “SMEs/firms/private capital” and “different levels of government (EU, national, regional)”. These are mentioned as factors that have to be addressed, not as actors having interests and shaping decision-making. Typical categories for discerning differential vulnerability to economic processes (e.g. class, gender, race, ethnicity) are fully absent. The JTM is analyzed as a “comprehensive” (lau007) or “holistic” (ast010) policy framework addressing economic, social and environmental aspects by mitigating negative impacts and promoting positive change. Even though we did not explicitly ask for this, the LLM frequently provided us with an affirmative assessment regarding the policy mechanism in its responses’ conclusions.

Overall, the resulting spatial imaginaries remind us of a refracted positivist regional science understanding of economic geography and regional development. Certain aspects of the specific configuration of stylistic and discursive elements are conditioned by our queries (e.g. using operators such as “describe” or “analyze” instead of “explain”). However, others – the tonality, the understanding of “economic sectors”

and the (dis)aggregation of “economic + geography” – point towards what we would call the LLM’s *machinic selectivity*. In what follows, we will turn towards the construction of disciplinary knowledge about economic geography to explore the mechanisms behind this probability-based text generation in greater detail.

4.2. Constructing Economic Geography

To further explore the de- and re-composition of spatial imaginaries in our dataset, we first turn to our separate conversation about the definition of economic geography the model generates and then to the understanding reflected in the assessments within the specific regional and political context.

When queried in isolation (“What is economic geography? Explain using key paradigms and authors from economic geography”), the LLM robustly includes three key perspectives. In nine out of ten chat conversations, “location theory” connected to authors such as Thünen or Weber is explicitly mentioned. “Globalization” and “Marxist” or “Political Economy” approaches are also included in almost all responses. “New economic geography”, “evolutionary economic geography”, “path dependence” and “institutional approaches” are mentioned in more than half of the queries. This is also reflected in the key authors associated with the discipline – all responses mention both Massey and Harvey, more than half mention Krugman, Dicken and Boschma. Beyond this probability-based produced “core” of the discipline, the less frequent positions among the on average six concepts included in the responses are taken by concepts such as “Global Production Networks”, “Network Theories” or “Spatial Economics” – and authors such as Christaller, Florida or Sassen.

However, asking the LLM to assess economic processes and regional policies through concepts and authors from economic geography does not imply that the above-queried imaginary of economic geography is applied. Rather economic geography is then relationally re-imagined in and through the given context. Throughout our 30 conversations on the selected regions in transition, we could identify both constants and variable components of the model’s answers. The most frequent components are displayed in Fig. 1, the number of their occurrences are mentioned in brackets. All responses draw on concepts of “path dependency” to assess the JTM. 23 out of 30 reference “regional innovation systems” or “learning regions”. More than two third of the responses include varying concepts used to address social aspects (e.g. “spatial justice”, “spatial justice” or “equity”). Two third reference “resilience” and still more than one third of the responses makes direct reference to “cluster theory”. These three components represent the core of the on average six perspectives used to assess this policy mechanism for regional development, other references are less frequent (e.g. “Growth Pole Theory”, “Global Production Networks” or “Cumulative Causation”). The concepts are relatively robustly associated with one or two corresponding authors. (Philip) Cooke⁶ is mentioned in 20 out of 30 responses – almost exclusively collocated with “regional innovation system” – making Cooke the most cited economic geographer in the dataset. The second and third most likely “authors from economic geography” in our context then include (Ron) Martin, (Peter) Sunley and (Paul) David who are most frequently associated with “path dependence”. Still more than one third of the responses reference (David) Harvey, (Doreen) Massey and (Michael) Porter. Despite these consistencies the LLM seems to have picked up with respect to dominant concepts and authors – e.g. to link Harvey and Massey to spatial justice or Cooke to regional innovation systems etc. – it is worth noting that the model does slip up occasionally. In one case, we for example encountered “David S. Massey” (mid010). Just as with author Sunley who never occurs independently from Martin, there seems to be a machinic

⁶ Technically, the LLM output usually only uses the last names in conjunction with a concept. Hence, including the first name is already an interpretation.

Critically assess the application of the Just Transition Mechanism in the region using concepts and authors from economic geography.

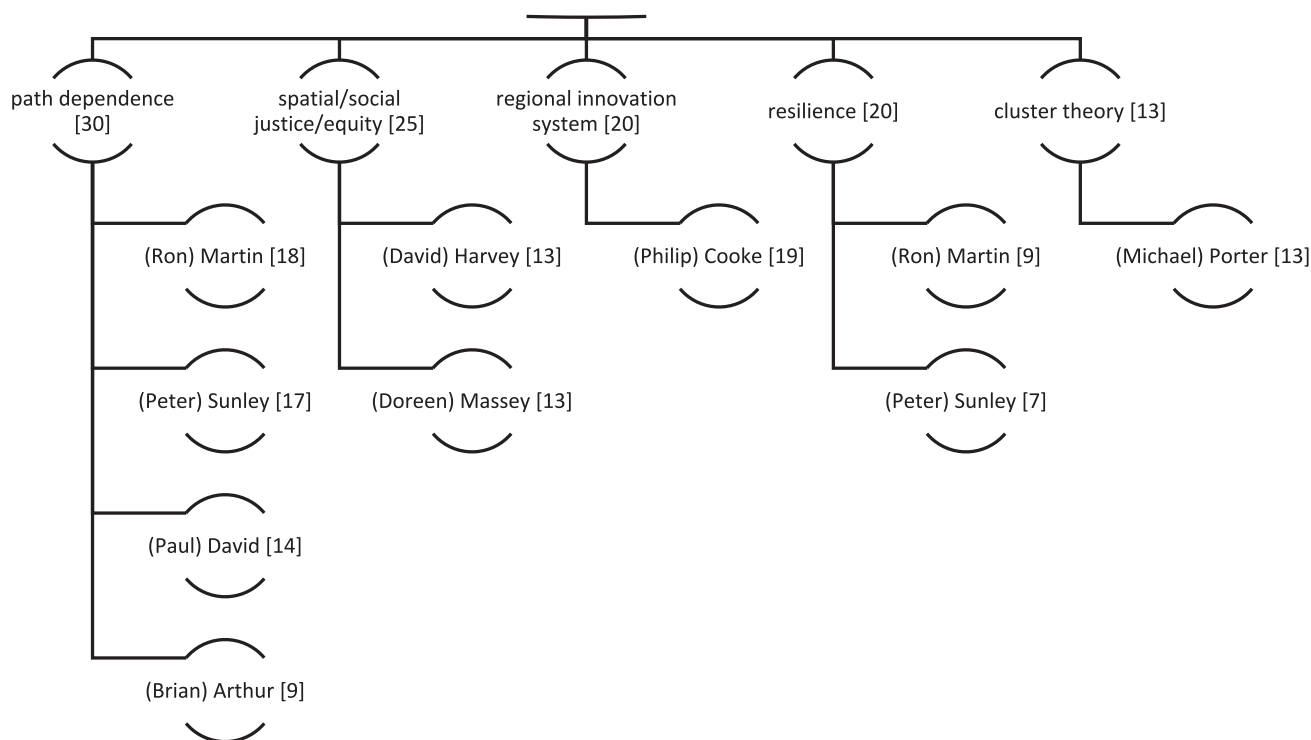


Fig. 1. Overview of concepts referenced over 10 times and corresponding authors mentioned over 5 times.

aggregation of related tokens rendering them so close to each other – i.e. probably consistent – that they become almost replaceable or even merge.

Considering these results also allows us to speculate about the way our context consisting in the respective region (Asturias, Lausitz, Midlands) and policy mechanism (JTM) affected the content generation. All analyses of the application of the JTM in the region included concerns for workers and communities such as, “To ensure a just transition, the mechanism includes measures to support social inclusion and mitigate adverse social impacts.” (lau006) “Just” [164 mentions]⁷ and “social” [167 mentions] are among the 20 most frequently used words in our dataset which likely influenced the following assessment and the authors and concepts deemed likely to relate to these themes. Aside from words related to the “economy/economic”, “region(al)” is the most frequent word in our corpus [534 mentions]. Therefore, we believe it is plausible to assume that this prevalence in the text conditioned the surprisingly robust reference to “Regional Innovation Systems” and corresponding authors. The highly robust reference to “path dependence” is the most interesting of these cases for us as formulations such as “path” or “path dependent” did not come up in the answers prior to the question regarding concepts and authors. We assume that this representation is associated with the frequent exploration of the theme “Transition(ing)” (fourth most frequent word, 454 mentions). The discussion of the concept “path dependence” (response 3) echoes the typical SWOT-styled juxtaposition of opportunities and challenges to be found in the model’s regional descriptions (response 1):

“Overall, Lausitz is navigating a significant economic transition, balancing its industrial legacy with new opportunities in sustainable development and innovation.” (lau003, from answer to question 1)
 “Lausitz’s history of coal mining creates a path-dependent economic structure that is challenging to shift. The JTM aims to overcome this by funding alternative industries and retraining programs. However, the entrenched skills, infrastructure, and social identity linked to coal mining present significant barriers.” (lau008, from answer to question 3)

It is thus not simply a frequently used word, but the structure of the description and the motif of challenge that seems to render reference to a specific concept probable.

Overall, while the focus on the region and transition processes notably affected the choice of authors and concepts compared to the control question, the *specific* region seems to play an almost negligible role. Even though we could identify some variation in respect to the five most frequently mentioned components of the model’s responses (see Fig. 2), the choice of corresponding authors and concepts was overall surprisingly homogeneous. The assessment for the Midland region is for example less often connected with the concepts “regional innovation system” and “cluster theory” which might be related to its more rural descriptions in the preceding questions. However, the different national contexts do not lead to referencing scholars not mentioned in the other ones (e.g. German or Spanish geographers).

Looking not only at the choice of authors and concepts, but also their representation can help to illuminate how the LLM (dis)aggregates text in its construction of imaginaries. In our conversations, the concepts and authors are referred to in a highly superficial way, often contributing one positive (appreciation/strength) and one negative aspect (concern/weakness) to the assessment of the application of the JTM in the given region. Harvey’s concern for inequalities is for example sometimes illustrated using the concept of the “spatial fix”:

⁷ The word frequencies in this paragraph all refer to the conversations prior to asking the third question regarding authors and concepts.

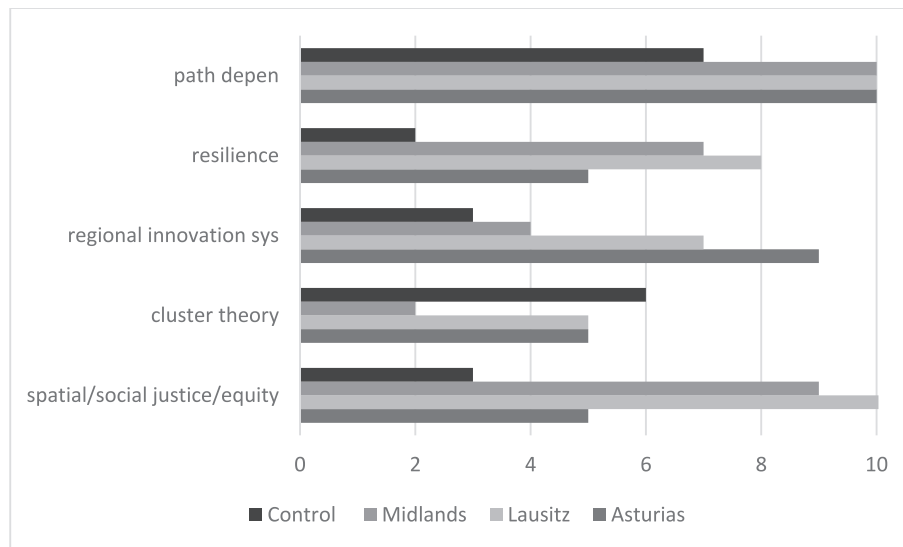


Fig. 2. Overview over concept mentions.

“Harvey’s concept of the spatial fix can be applied to understand how the JTM seeks to address the spatial inequalities created by the decline of the coal industry in Lausitz. By redistributing resources and investments to this lagging region, the JTM attempts to fix the spatial imbalances and stimulate regional economic development.” (lau001)

At first sight, this application of the concept of the “spatial fix” might seem plausible – at least semantically –, but theoretically speaking the concept is reduced to an intentional balancing effort associated with a collective actor (the JTM), rather than being analyzed structurally in the context of a crisis of over-accumulation of capital. The conversations are characterized by a lack of depth which sometimes irritated our sense of commensurability or common sense as *concepts* are stripped from their theoretical tradition and social context, reduced to *words* and plugged into a SWOT-styled feedback loop. The imaginary is one of ontological flatness and cybernetic feedback. While the LLM is able to group concepts into containers such as “New economic geography” or “Political Economy” in isolated queries, the additional tokens introduced through the regional and policy context seem to weaken these consistencies.

4.3. Representation & Machinic Selectivity

Which logics and mechanisms help to explain this selectivity? In what remains, we reflect on the LLM’s logo-centric construction of spatial imaginaries and speculate on its training data as sources of selectivity.

As we noted regarding both the regional descriptions and their discussion using authors and concepts from economic geography, the conversations for the three regions in transition were remarkably similar. All regional descriptions exhibit the same spatial imaginary (4.1): They are described in a neutral to mostly positive way based on a set of similar economic sectors, geographical factors and the packaging of the “challenge” of transition in SWOT-styled conclusions. We do in fact not know whether the model treats spatial references (names of regions, cities etc.) any different from other tokens and how this may give rise to a missing sense of place or historical geographical specificity. We could, however, observe that all three regions in transition were constructed through a similar positivist regional science imaginary of economic geography and stylistic formalism. Viewed this way, the regions do seem very similar, giving little reason for applying different geographical concepts or assessing the JTM differently. As a model of language (Hayles 2023, 258), the LLM by default operates in the purely

relational space spanned between its tokens, resulting in the kind of abstract comparative global perspective showcased by this section in one critical assessment:

– **Successful Examples:** Regions like the Ruhr Valley in Germany, which successfully transitioned from coal to a diversified economy, provide valuable lessons. These include the importance of strong local governance, substantial public investment, and community engagement.

– **Potential Pitfalls:** The decline of coal in the Appalachian region of the United States highlights the dangers of inadequate support for displaced workers and insufficient diversification efforts.” (mid008)

The generated response here draws highly abstract lessons for the challenges and opportunities the Midland region faces in its transition away from peat extraction. Spatial context matters in that the focus on the region (“regional innovation systems”) and questions of transition (“path dependence”) co-constitute the appropriate disciplinary knowledge. But it matters in an ontology of space flattened through the model’s logo-centric gaze.

While such analyses have to be very wary of how the own input conditions the model’s responses, repeated conversations have helped us to approximate logics inherent in its grasp of geographical knowledge. OpenAI does not disclose the corpus that goes into the training and how later alignment interferes, but we can assume that a broader representation of authors and concepts in that dataset translates into a higher probability of making it into the LLM’s “short lists”. With far more than 100,000 citations each, the sheer number of mentions in association with the words “economic geography” likely reserves authors such as Massey, Harvey or Krugman prime spots. From a geographical perspective, the recurring inclusion of Krugman is most curious as the integration of his “new economic geography” into the disciplinary knowledge of economic geography is sometimes contested (for an argument in favor of their inclusion, e.g. Barnes/Christophers 2018: 45). This semantic conflict over the label “new economic geography” seems to be lost on the model. Just like this frequent and therewith probable semantic tie “naturalizes” Krugman as a disciplinary geographer, the underlying logic likely also explains why less represented approaches from feminist economic geography, themes such as gentrification, ethnic economies or gender and authors exceeding the “Anglophone economic imaginary” (Sheppard 2024) are marginalized in the model’s representation of economic geography.

The way these LLMs are presented to us as black boxes means that we do not know which geographical knowledge went into the model. We

can for example only make an educated guess that the robust (and merged) reference to the authors “Martin/Sunley” when assessing the JTM could relate to their 2006 paper on path dependence in the highly-ranked *Journal of Economic Geography* (Martin and Sunley 2006). This reference might be constructed as a more probable response either because of its representation in the training data or a weighting mechanism introduced through alignment. More complex series of questions could surely improve or alter the model’s performance, pivoting it towards reproducing different imaginaries, for example by asking it to work only with suitable regional case studies when representing an economic geographical perspective, computing google scholar citation metrics for assessing the representativity of authors or by relying on web search results for its responses. However, the *machinic selectivity* we have explored would still apply.

This brings us to questions of social practice and the production of subjectivity associated with the LLM interaction. While LLMs are large models of language the use and interaction generate subjects and subjectivities. On the one hand, we here find certain “subjects of enunciation” (Lazzarato 2008): the helpful assistant with its characteristic affective dispositions (Magee, Arora, and Munn 2023) and the kind of dispositions it calls forth in us as its reflexive users, e.g. endlessly asking us to further specify our questions and suggesting possible follow-ups. On the other hand, we find certain subjectivities which are typically objectified into “subjects of the statement” (Lazzarato 2008) in the model’s discourse (policy makers, firms, the workforce, the region etc.). It takes transparency regarding the model and conscious labor on the user’s side to actively mold these dialogical subjective dispositions and escape or deconstruct the most probable productions of subjectivity.⁸ We would argue that rather than turning this circumstance into a task for its users – “It can’t be that stupid, you must be prompting it wrong” as the saying goes –, our conclusion as digital geographers should be to point out potential mechanisms through which machinic selectivity could have adverse effects regarding the reproduction of hegemonic narratives and knowledge. The normalization of a specific imaginary of the economic and economic geography, the model’s canonization of disciplinary knowledge and its limited relational conception of space are three such effects we could observe in our conversations.

5. Conclusion

There is a growing body of literature on the political economy surrounding the making of LLMs which problematizes the *accumulation by dispossession* and precarious labor involved in the extraction of training data and manual refinement of these models’ outputs. While the debate to what extent the use as training data incurs *individual* copyright infringement is still ongoing, a less legalistic approach helps to understand that what is appropriated in the making of these models is a product of *collective* work formerly not-owned – the relationality within the totality of accumulated data (Guadamuz 2024; Chesterman 2024). What companies such as OpenAI thus market to their subscribers is a cloud computing service resting on the appropriation of the linguistic selectivity inherent in the commons they themselves helped to build.

⁸ Arguably, as we interact with the LLM it is precisely our confusion regarding these machinic subjectivities – misunderstanding ourselves as the “cause of the statements”, when we are really inserted into a machinic setting “folding” our utterances into probable frames of reference – that causes much of the confusion regarding the question of meaning and intelligence. What is mistaken for thinking is the labor of aligning our and the model’s dispositions: „At the end of the interview, you are the subject of the statement, an effect of the semiotics of the machine of communication, believing itself to be a subject of enunciation, feeling itself to be the absolute, individual cause and origin of statements, whereas in reality it is the result of a machinery, no more than the end point in the process. Your words are folded over statements and modes of expression which are imposed on you and expected of you.” (Lazzarato 2008).

Aside from this unpaid appropriation, the datafication involved in making LLMs is also built on the exploitation of (paid) “ghost work” (Gray and Suri 2019), for example for the manual validation of these models’ outputs. Reaping the benefits of an emerging “planetary labor market” for spatio-temporally flexibilized digital work (Graham and Ferrari 2022), this is often done by poorly paid gig-work in the Global South (Muldoon et al. 2023; Zinn 2023), while having a strong bias towards the norms and values of urban people in the Global North. In our contribution, we seek to expand these critiques for a concern regarding the ideological implications of LLM-based text production.

Based on the assumption that general-purpose LLMs are increasingly built into software stacks through logics of “infrastructural platformization” (van Dijck et al. 2019; van Dijck et al. 2018), we are likely to meet more LLM-generated representations of spatial imaginaries in the future. Therefore, we argue that we need concepts and methods on how to deal with them. We found a cultural political economy perspective on “imaginaries” (Jessop 2010; Sum and Jessop 2013) and its sensitivity for the integration of genres, stylistics and discursive elements useful to identify logics of structuration within LLM output. By focusing on three peripheralized regions in transition and the LLM’s analyses of regional economic policy, we explored an experimental approach of dealing with LLM output. We generated a corpus of 55 conversations which we analyzed using a combination of qualitative content analysis and quantitative methods. Our methodology allowed us to identify the refracted positivist regional science imaginary of economic geography and flattened spatial ontology the model generates in its default mode (4.1). Furthermore, we could observe a reductionist representation of economic geographic disciplinary knowledge combined with a superficial, cybernetic handling of concepts and theoretical traditions (4.2). In providing its answers, the model reproduced highly selective notions of the economic, geography and disciplinary knowledge, while struggling to fact or in the peculiarities of a given spatial contexts. While we could speculate plausibly about the machinic logics and potential impacts of the training dataset, important questions – e.g. regarding the operationalization of space or volume and weighing of academic publications within the training data – remain opaque (4.3).

We draw a threefold conclusion from our engagement with the LLM’s output. Firstly, it points towards the responsibility of those who by pillaging the commons without disclosing the academic corpora which went into the model have introduced an opaque additional layer of selectivity into the representation and proliferation of geographical knowledge. These include private companies, public institutions, the producers of data sets, further underlining that machinic-selectivity is socially produced – and outlining the normative criteria a commons-oriented development of LLMs has to address. In the absence of access to the training data, we secondly argue that scholars need concepts and methodologies for engaging with LLM output and for sensitizing for the logics of (mis)representation. In spite of the fact that we are not dealing with logically consistent, closed and empirically verifiable semiotic systems, we found the adaptation of the concept of imaginaries and a combination of quantitative and qualitative analytical approaches helpful for identifying recurring structures within the text. We hope that our speculative methodology contributes to the academic discourse of dealing with these kinds of texts. For future research, we would argue that working with much larger datasets and applying genuine corpus linguistics would be a promising methodological approach to further investigate the selective logics we have identified. Thirdly, a political argument follows from the first two: The hype and spread of LLM tools should worry us not only due to the environmental costs and precarious labor relations that go into their making (Crawford 2021; Graham and Ferrari 2022); rather we should also problematize the kinds of potential ideological implications a cultural political economy perspective sensitizes for. There are enough mechanisms of exclusion and marginalization at work in disciplinary knowledge production as it is, unless we gain a better understanding of underlying mechanisms, we would argue that there are good reasons for a politics of refusing additional stochastic

reinforcement.

CRediT authorship contribution statement

Boris Michel: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Yannick Ecker:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Banerjee, A., Schuitema, G., 2022. How just are just transition plans? Perceptions of decarbonisation and low-carbon energy transitions among peat workers in Ireland. *Energy Res. Soc. Sci.* 88, 102616. <https://doi.org/10.1016/j.erss.2022.102616>.
- Bender, Emily M. 2023. Talking about a 'schism' is ahistorical. <https://medium.com/@emilymenonbender/talking-about-a-schism-is-ahistorical-3c454a77220f>.
- Benjamin, R.h., 2019. *Race after technology: Abolitionist tools for the new Jim code*. polity, Cambridge.
- Bode, K., Goodlad, L.M.E., 2023. Data Worlds: An Introduction. *Critical AI* 1 (1–2). <https://doi.org/10.1215/2834703X-10734026>.
- Bolet, D., Green, F., Gonz ales-Eguino, M., 2023. How to Get Coal Country to Vote for Climate Policy: The Effect of a "Just Transition Agreement" on Spanish Election Results. *Am Polit Sci Rev* 1–16. <https://doi.org/10.1017/S0003055423001235>.
- Broussard, M., 2019. *Artificial unintelligence: How computers misunderstand the world*. First MIT Press, paperback edition. The MIT Press, Cambridge, Massachusetts, London, England.
- Broussard, M., 2023. *More than a glitch: Confronting race, gender, and ability bias in tech*. The MIT Press, Cambridge.
- Chesterman, S., 2024. Good models borrow, great models steal: intellectual property rights and generative AI. *Policy and Society*. <https://doi.org/10.1093/polsoc/puae006>.
- Crawford, K., 2021. *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press, New Haven.
- Crawford, K., Paglen, T., 2021. Excavating AI: the politics of images in machine learning training sets. *AI & Soc.* <https://doi.org/10.1007/s00146-021-01162-8>.
- Dengel, Andreas, Rupert Gehrlein, David Fernes, Sebastian G orlich, Jonas Maurer, Hai H. Pham, Gabriel Gro mann, and Niklas D. g. Eisermann. 2023. Qualitative Research Methods for Large Language Models: Conducting Semi-Structured Interviews with ChatGPT and BARD on Computer Science Education. *Informatics* 10 (4). doi: 10.3390/informatics10040078.
- D'Ignazio, C., Klein, L.F., 2020. *Data feminism*. The MIT Press, Cambridge, Massachusetts.
- Frankfurt, H.G., 2009. *On Bullshit*. Princeton University Press, Princeton <http://gbv.eblib.com/patron/FullRecord.aspx?p=457797>.
- Graham, Mark, and Fabian Ferrari, eds. 2022. *Digital work in the planetary market*. The MIT Press-International Development Research Centre series. Cambridge Massachusetts: The MIT Press. Accessed July 23, 2024.
- Gray, M.L., Suri, S., 2019. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Houghton Mifflin Harcourt, Boston, New York NY.
- Guadamuz, Andres. 2024. A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs. *GRUR International* (2). doi: 10.2139/ssrn.4371204.
- Hayles, N.K., 2022. Inside the Mind of an AI: Materiality and the Crisis of Representation. *Nlh* 54 (1), 635–666. <https://doi.org/10.1353/nlh.2022.a898324>.
- Hayles, N.K., 2023. Subversion of the Human Aura: A Crisis in Representation. *Am. Lit.* 95 (2), 255–279. <https://doi.org/10.1215/00029831-10575063>.
- Hicks, M.T., Humphries, J., Slater, J., 2024. ChatGPT is bullshit. *EthicsInf Technol* 26 (2). <https://doi.org/10.1007/s10676-024-09775-5>.
- Jessop, B., 2010. Cultural political economy and critical policy studies. *Critical Policy Studies* 3 (3–4), 336–356. <https://doi.org/10.1080/19460171003619741>.
- Lazzarato, Maurizio. 2008. The Machine. [https:// transversal.at/transversal/1106/lazzarato/en](https://transversal.at/transversal/1106/lazzarato/en).
- Lehued , S., 2024. An elemental ethics for artificial intelligence: water as resistance within AI's value chain. *AI & Soc.* <https://doi.org/10.1007/s00146-024-01922-2>.
- Mackenzie, A., 2017. *Machine learners: Archaeology of a data practice*. The MIT Press, Cambridge, Massachusetts.
- Magee, L., Arora, V., Munn, L., 2023. Structured like a language model: Analysing AI as an automated subject. *Big Data Soc.* 10 (2). <https://doi.org/10.1177/20539517231210273>.
- Martin, R., Sunley, P., 2006. Path dependence and regional economic evolution. *J. Econ. Geogr.* 6 (4), 395–437. <https://doi.org/10.1093/jeg/1bl012>.
- Morreale, F., Bahmanteymouri, E., Burmester, B., Chen, A., Thorp, M., 2023. The unwitting labourer: extracting humanness in AI training. *AI & Soc.* <https://doi.org/10.1007/s00146-023-01692-3>.
- Muldoon, James, Callum Cant, Mark Graham, and Funda Ustek Spilda, 2023. The poverty of ethical AI: impact sourcing and AI supply chains. *AI & Soc.* <https://doi.org/10.1007/s00146-023-01824-9>.
- Munn, L., Magee, L., Arora, V., 2023. Truth machines: synthesizing veracity in AI language models. *AI & Soc.* <https://doi.org/10.1007/s00146-023-01756-4>.
- Noble, S.U., 2018. *Algorithms of oppression: How search engines reinforce racism*. New York University Press, New York.
- O'Neil, C., 2017. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books, London.
- Pasquinelli, M., 2023. *The eye of the master: A social history of artificial intelligence*. verso, London, New York.
- Raley, R., Rhee, J., 2023. Critical AI: A Field in Formation. *Am. Lit.* 95 (2), 185–204. <https://doi.org/10.1215/00029831-10575021>.
- Roberge, Jonathan, and Michael Castelle, eds. 2021. *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*. 1st ed. 2021. Cham: Springer International Publishing. Accessed April 17, 2024.
- Sheppard, E., 2024. Exceeding the Anglophone economic geographical imaginary. *Progress in Economic Geography* 2 (1), 100010. <https://doi.org/10.1016/j.peg.2024.100010>.
- Steyerl, Hito. 2023. Mean Images. *New Left Review* (140/141): 82–97.
- Sum, N.-L., Jessop, B., 2013. *Towards a cultural political economy: Putting culture in its place in political economy*. Edward Elgar E-Book Archive. Edward Elgar Pub. Ltd., Cheltenham.
- Tenen, Dennis Y. 2024. *Literary theory for robots: How computers learned to write*. Norton shorts. New York, NY: W. W. Norton & Company Inc.
- Thatcher, J., O'Sullivan, D., Mahmoudi, D., 2016. Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environ. Plann. D: Soc. Space* 34 (6), 990–1006. <https://doi.org/10.1177/0263775816633195>.
- van Dijk, Jos , Thomas Poell, Martijn de Waal., 2018. *The platform society: Public values in a connective world*. Oxford University Press, New York.
- van Dijk, J., Nieborg, D., Poell, T., 2019. Reframing platform power. *Internet. Policy Review* 8 (2). <https://doi.org/10.14763/2019.2.1414>.
- Zinn, Z., 2023. Data Work and its Layers of (In)visibility. *Just Tech, Social Science Research Council*. <https://doi.org/10.35650/JT.3060.d.2023>.