UTKU B. DEMIR

GENERATIVE AI AND THE INSTITUTIONS OF CONTROL: AN ANALYSIS OF ALGORITHMIC MEANING-MAKING AND SUBJECTIVATION

Contents

1 Introduction 6

2 Theoretical Framework 9

2.1 Research Question 9

2.1.1 Possible lines of argument 9

2.2 State of the Art 10

3 AI 12

3.1 (A second intro to AI to introduce the problematic) 13

3.2 From Symbolic Rules to Statistical Modulation: A Brief History of AI and NLP 13

3.3 The Transformer Architecture: Infrastructure of Modulation 15

3.3.1 From Recurrent Bottlenecks to Attention-Based Modulation 16

3.3.2 Under- & Overfitting 18

3.3.3 Gradient Descent: Sinking into the Manifold 18

3.3.4 Back Propagation 19

3.4 Undistributed 21

3.5 AI as Desiring Machine 21

3.5.1 From Pre-Training to Fine-Tuning: Modulating the Model's World 22

4 GenAI as an Institution of Control 24

4.1 Postscript: Updating the Societies of Control 24

4.2 Mold & Modulation 25

4.3 Institutions of Desire-Management 27

4.4 Personalisation and Probabilistic Meaning-Making 28

4.5 Enregistrement and Subjectivation 29

4.6 The World Model: A Neoplatonic Representation 30

4.6.1 Latency 32

4.7 Agency; Kafka's Trial and the Logic of Indeterminate Governance 34
4.7.1 Language in LLM 37
4.8 UNDISTRIBUTED/TBD 37
4.8.1 Creativity: Discrete vs. Continuous 38
4.8.2 Dividuation 38
4.9 Difference, Repetition, Singularity (Potential discussion about the need for sensory input for the genAI (LeCUn?) 38

5 Conjunctive Synthesis and the Construction of Subjectivity 40
5.1 Hegemonic Representation 42
5.2 Hallucinations and Lines of flight in Algorithmic Architectures 42
5.3 Revolutionary Possibilities? 42
5.4 Reclaiming microflows of modulation 43
5.5 Possibility of resistance within feedback infrastructures 43
5.6 Experimental subjectivity in response to AI systems 43

6 Conclusion & Outlook 44

Glossary

epoch Epochs represent the number of times the entire training dataset passed through the algorithm. Nebius-Team 2024. 19

kernel In Machine Learning, the Kernel method consists of using a linear classifier to solve a non-linear problem. This is achieved by transforming a linearly inseparable set of data into a linearly separable set (Melanie 2024).. 16

token Tokens represent the smallest units of data that the model processes, such as words or characters in natural language processing (Cser 2024). . 16, 17

Acronyms

AI Artificial Intelligence. 6, 14, 17, 30, 34, 41

CNN Convolutional Neural Network. 16

D&G Gilles Deleuze & Felix Guattari. 8, 9, 28, 38, 41, 42

DL Deep Learning. 15, 18

DNN Deep Artificial Neural Network. 6, 14, 15

genAI Generative Artificial Intelligence. 6, 7, 8, 12, 15, 22, 24, 25, 26, 27, 28, 29, 30, 32, 34, 36, 41

GM Generative Model. 34

GOFAI Good old-fashioned AI. 13, 19

LLM Large Language Model. 3, 6, 12, 15, 16, 22, 28, 37

LM Language Model. 30

MGM Multimodal Generative Model. 30

ML Machine Learning. 15

NLP Natural Language Processing. 6, 14, 30

NN Artificial Neural Network. 6, 14, 30

RNN Recurrent Neural Network. 16

SL Supervised Learning. 30

symAI Symbolic Artificial Intelligence. 12, 13, 15

T2IM Text to Image Model. 30

UL Unsupervised Learning. 30

1
Introduction

In recent years, substantial advancements in the field of Artificial Intelligence (AI), particularly through developments in Artificial Neural Network (NN) and Deep Artificial Neural Network (DNN) architectures, have enabled the deployment of predictive models across a wide array of domains, from social media platforms and search engines to natural language processing tasks such as text classification and topic modelling. While these applications primarily focused on analysis, relevance association, personalisation, and prediction, a new paradigm has emerged in the form of Generative Artificial Intelligence (genAI). Once a relatively silent front in Natural Language Processing (NLP) research, genAI's history dates back to the 1950s (Cao et al. 2023, p. 4). Unlike traditional models, genAI systems are capable of producing novel outputs; such as text, images, or code by extracting and operationalising intent from human-provided instructions. This shift marks a transformation not only in the goals and capabilities of AI, but also in its epistemic and operational logics. Particularly in its implementations based on transformer architecture, genAI now occupies a central role in the production, interpretation, and circulation of information and media, moving beyond automation and decision support to enabling generative processes that raise fundamental questions about agency, subjectivity, and truth.

The analysis and critique of the AI models is nothing new, the surveillance capabilities that has been established by the contemporary data analysis (e.g. Krasmann 2017), the effect of a completely data based rationality introduced by the datalogical turn (see Clough and Gregory 2015), a dividualised information flows through the profiling and association by the models running on the web (see e.g. Cheney-Lippold 2011), and the decision-making systems adopting an algorithmic governmentality (see e.g. Rouvroy 2007), and various ethical, as well as, bias related research (e.g. Kordzadeh and Ghasemaghaei 2022) have been a vibrant field in the last years. The capability genAI models especially LLMs to meaning-making (Dishon 2024; Gretzky 2024; Mishra and Heath 2024), however, , have provoked renewed inquiry. While these generative processes are rooted in a long history of statistical and computational development, contemporary architectures with their interpretation of the vast datasets of productive human legacy introduce an immediate representational logic, one that embodies a distinct political model or *governing rationality* (Amoore et al. 2024, p. 2). While this interpretative substance [1] enables genAI models to communicate human-like, also establishes a power structure over governing information as a governing institution (see e.g. MacKenzie and Porter 2021 or Dishon 2024). Beyond their technical capacities, these systems enact a form of governance deeply entangled with power, normativity, and new forms of subjectivisation (Eloff 2021). GenAI operates through a distributional logic that structures knowledge by modelling statistical regularities in a datafied world (Amoore 2023). These systems "traverse data foundations" to generate outputs that appear plausible within a learned distribution, but without deterministic causality or transparent justification. In doing so, they blur the line between representation and enactment: rather than merely classifying or reflecting reality, they participate in its structuring. Outputs from large language models, for instance, are increasingly treated as epistemically meaningful, even authorial, despite being arguably generated by processes that lack a stable or intentional agent.

[1] The governing rationality of genAI models (see e.g. ibid.) refers to the generative structure of algorithmic meaning-making and should not be confused with Rouvroy n.d.'s concept of Algorithmic Governmentality. Whereas Rouvroy addresses algorithmic turns in neoliberal governance, governing rationality designates the internal logic established by the model itself.

The transformation genAI introduces is beyond its technical nature inquires institutional analysis (MacKenzie and Porter 2021) of the power structure deployed in this new constellation. Framing the question as *in what form of institutional nature the architecture and rationality of the contemporary genAI algorithms deploy and what conclusions these implicate on agency, subjectivisation, ciritique, and resistance*; this study situates genAI within the broader transformation of power described by **deleuze1992a** as the shift from Michel Foucault's *disciplinary societies* **Foucault1977** to *societies of control*. Whereas Foucault's account of disciplinary power emphasized enclosure, surveillance, and the moulding of subjects within bounded institutions like prisons, schools, and hospitals (Foucault 2008), Deleuze's postscript outlines a more fluid and continuous, flexible form of control. Deleuze's description of governance in the **deleuze1992a** operates through modulation: subjects are governed not by confinement but through their data traces, captured and recomposed in real-time. In such control societies, individuals become *dividuals*, decomposed into discrete, analysable data points recombinable by algorithmic systems (MacKenzie and Porter 2021) vastly increasing the field of visibility on the bodies (Foucault 2008). Control does not dissolve institutions into flow; rather, it reorganizes them into mechanisms that totalise by sequencing dividuals across domains. Institutions of control no longer discipline by containment, but by aggregating, modelling, and redistributing datafied subjectivities through infrastructures such as genAI platforms. This thesis therefore advances the hypothesis that generative AI models function as *institutions of control*, not metaphorically, but operationally. Their architectures instantiate regimes of truth through statistical inference, acting as epistemic infrastructures that determine what can be said, imagined, or inferred. This shift raises the stakes of critique. In control societies, resistance cannot depend on unmasking ideology or demanding transparency alone.

Instead, critique must become processual and counter-sequential: it must trace the operations of sequencing and propose alternative arrangements that disrupt the logic of totalisation (MacKenzie and Porter 2021) . Accordingly, I adopt a micropolitical perspective, asking not merely what genAI systems do, but how they do it. What are the machinic elements, attention mechanisms, tokenisation, transformer layers that enable the modulation of information and subjectivity? And where, if anywhere, might one locate lines of flight within these architectures? Can their operation be reappropriated as tools for critique, invention, or resistance?

Rather than positioning genAI as either emancipatory or repressive, this study approaches it as a complex institutional actor embedded in contemporary capitalism, furthermore develops a ciritical reflection of Deleuze's concept of control by deviating both in terms of the analysis of the genAI architecture, and in the critique of the institutional formation these models establish other works of Gilles Deleuze & Felix Guattari (D&G) (see e.g. Deleuze and Guattari 1983, Deleuze and Guattari 1987). As D&G argue, capitalism decodes and deterritorialises flows only to reterritorialise them elsewhere (Deleuze and Guattari 1983). In this context, genAI functions not merely as a tool of production or surveillance, but as a mechanism of epistemic reterritorialisation: producing coherence, narrativity, and alignment from fragmented inputs. It governs the production of meaning while also embedding the potential for alternative uses, ones that may exceed or redirect capitalist logics. This study thus offers a re-critique of control societies through the lens of generative architectures. Combining political theory, and technical analysis, it examines how genAI models operate on both infrastructural and epistemic levels. In doing so, it seeks to develop a renewed account of power, one grounded in probabilistic modulation, infrastructural inscription, and the micropolitics of machine reasoning. Ultimately, I argue that while the generative capabilities challenge the pillars of the control society concept, while finding particularly insightful correspondences in other literature of D&G .

2
Theoretical Framework

This chapter is mostly incomplete

One could argue that artificial intelligence is *no longer an engineering discipline* (**dignum2023**), if it has ever truly been one.

The present thesis builds on critical perspectives from political theory and philosophy of technology to interrogate the institutional, epistemological, and political implications of contemporary generative AI systems by analysing their architectural structures and through the analysis is situated within D&G's conceptual apparatus of control, desire, and modulation. Together, these approaches provide the foundation to conceptualize generative AI not merely as computational artefacts but as infrastructural actors in the contemporary configuration of control societies.

2.1 Research Question

RQ: How does the development of the current AI algorithms, particularly those of generative AI models (genAI), embody, extend, or contradict Gilles Deleuze's concept of societies of control by modulating subjectivity through its probabilistic, non-linear operation structures?

RQ (Alternative): To what extent can the processes of meaning-production in generative AI algorithms, particularly the transformer architecture be understood through Gilles Deleuze's concepts of control and modulation, especially in relation to subjectivity construction via their probabilistic, non-linear operational structures?

RQ(Alternative): In what ways do generative AI models instantiate the logics of control society, and how might their probabilistic architectures transform the institutional production of subjectivity and meaning?

2.1.1 Possible lines of argument

- The distributional structure of generative AI, which produces content based on joint probability distributions, creates a dynamic of non-linear causality that reflects modulation without control, where subjectivity isiiiii formed through continuous, data-driven modulation rather than direct causative action in the absence of directional control.

- Despite their current state, future genAI models offer unique ways for individuals to create lines of flight and support the becoming of nomadic subjects.
- Transformer-based genAI systems operate as infrastructural institutions that govern subjectivity through modulation rather than disciplinary enclosure; their architectures do not command but statistically orchestrate behavior and meaning.
- Rather than merely reflecting ideology, genAI systems produce truth regimes through processes of probabilistic inscription, positioning themselves as epistemic institutions within the broader infrastructure of control.
- Despite their instrumental role in capitalist modulation, genAI architectures retain machinic singularities that can be redirected toward critique, invention, and experimental modes of subjectivation, forming potential "lines of flight."

[1]

[1] **TODO**: Revise

2.2 *State of the Art*

TODO:

- ☐ This part is to be adcvanced later on
- ☐ Rouvroy n.d.
- ☐ Pasquinelli 2023

While there are various attempts to adapt the notion of control to modern digital developments, such as those by Brusseau 2020, or partially or completely rejecting the concept of control as an inadequate explanation of current digital power structures (see, for example, Hui 2015), these works either overlook the novelty and potential of genAI mechanisms or fail to explore the *dividual* dimension. Furthermore, earlier works focusing specifically on the aspect of dividualisation (see e.g. Cheney-Lippold 2011; Van Otterlo 2013) are, unfortunately, temporally limited as they were unable to analyse generative models that had not yet reached the level of advancement we see today.

Other theorists focus on various aspects of genAI models, with a common tendency to analyse ethical considerations, their application under neoliberal governmentality, and their role within surveillance capitalism (see e.g. Gillespie 2024; Haggerty and Ericson 2000; Zuboff 2019). While these aspects are not outside the scope of this research, they were central to my previous bachelor's thesis. This current study specifically focuses on the structure of probabilistic models, particularly examining their construction at present (and speculatively in the future), without delving into how their misuse (or intentional use) may play out.

Furthermore, while this study addresses the political implications of current generative AI usage, its focus is primarily on the analysis of power operations and, within this context, relevant to questions of democracy only insofar as they relate to the operation of power. This study does not, however, aim to provide a comprehensive analysis of democratic risks associated with AI, as explored by others (see e.g. **coeckelbergh2024**; Zarkadakēs and Tapscott 2020).

This research, therefore, is focused on three primary areas of debate: (1) an analysis of genAI models as entities deploying power, examining their mechanisms (see e.g. Amoore et al. 2024; Konik 2015; MacKenzie and Porter 2021); (2) modern reflections on Deleuze's theory (see e.g. **mischke2021c**; Poster, Savat, and Deleuze 2010); and (3) sources analysing algorithmic structures with technical expertise (see e.g. **vaswani**; Bender et al. 2021).

3
AI

> [...] descending into the hidden abode of production means something else in the digital age. It means that we must also descend into the somewhat immaterial technology of modern-day computing, and examine the formal qualities of the machines that constitute the factory loom and industrial Colossus of our age. The factory was modernity's site of production. The "non-place" of Empire refuses such an easy localization. For Empire, we must descend instead into the distributed networks, the programming languages, the computer protocols, and other digital technologies that have transformed twenty-first-century production into a vital mass of immaterial flows and instantaneoustransactions. Indeed, we must read the never ending stream of computer code as we read any text (the former having yet to achieve recognition as a "natural language"), decoding its structure of control as we would a film or novel.
> (A. Galloway 2001, p. 82)

The current chapter provides the technical foundation for the subsequent political analysis of generative AI. To properly understand the institutional role of genAI models, it is necessary to first trace their historical development, underlying architectures, and operational logics. While this section remains on a primarily technical level, it also gestures towards the epistemological and political stakes that will be elaborated later.

The chapter proceeds chronologically and conceptually. It begins by outlining the historical trajectory of artificial intelligence research, distinguishing between the early, symbolic paradigm (Symbolic Artificial Intelligence (symAI)) and the contemporary, statistical approaches that characterize deep learning and generative models. This includes an explanation of neural networks, self-supervised learning, and the rise of transformer architectures as the technical backbone of modern LLMs.

Beyond mere description, this technical overview serves a strategic purpose: it demonstrates how AI, even at the level of architecture, already embeds specific logics of inference, representation, and control. These are not neutral technical details, but the material conditions that enable AI systems to operate as infrastructures of knowledge production, decision-making, and ultimately, governance.

The subsequent sections therefore provide both the necessary technical background and the conceptual scaffolding for the analysis of generative AI as a distributed, non-symbolic institution of power that follows.

3.1 *(A second intro to AI to introduce the problematic)*

TBD

3.2 *From Symbolic Rules to Statistical Modulation: A Brief History of AI and NLP*

This part can be a good addition to the "state of art"

- Montanari 2025 is a good source for a brief techno-political history and genealogy of llms
- Also here https://www.technologyreview.com/2024/07/10/ is-artificial-intelligence-ai-definitive-guide/ and here Pasquinelli 2023

Artificial intelligence emerged in the mid-20$^{\text{th}}$ century, grounded in the formal logics of symbolic representation. The foundational paradigm, now referred to as symAI or Good old-fashioned AI (GOFAI) , conceived intelligence as a matter of symbolic reasoning over explicitly encoded rules. Early AI systems aimed to emulate human problem-solving by manipulating structured propositions within formal languages. The assumption was clear: if the world could be faithfully translated into a logical schema, machines could infer, deduce, and act rationally (see Eloff 2021, p. 183).

This early paradigm treated intelligence as a computational process operating over discrete symbols according to explicitly programmed rules. AI systems under this logic were built to emulate deductive reasoning and problem-solving: if the world could be encoded in a set of symbolic propositions, intelligent behavior could be generated by manipulating those propositions through logic (ibid., p. 183). Relying on handcrafted rule sets, to recognise patterns the digit six in an image for instance, one might encode the features "a closed loop at the bottom" and "a curve rising to the right." Such symbolic heuristics were sufficient so long as the data was clean and the context unambiguous. But real-world ambiguity proved hostile to symbolic systems. As symAI attempted to scale into more complex domains like vision or language, it revealed its brittleness (ibid., pp. 183–184). Philosophers of phenomenology were early critics of this paradigm. Hubert Dreyfus, among others, argued that human intelligence was not symbolic, but embodied, situated, and fundamentally non-representational (Dreyfus 2009). Despite such critiques, symAI dominated the first decades of research. This rationalist framework aligned with early cognitive science's attempts to model the mind as a rule-based machine of symbolic representation (see Montanari 2025, pp. 194–197).

The 1956 Dartmouth Conference institutionalized these ambitions by defining AI as “the science and engineering of making intelligent machines” (ibid., p. 195). Systems like the Logic Theorist and expert systems in medicine or law exemplified this approach. Yet, these systems could not generalize beyond predefined rules. When confronted with noise or shifting contexts, their logic collapsed. The result was a period of stagnation and disillusionment now remembered as the “AI Winters” (Eloff 2021, p. 183). [1]

[1] **TODO:** A more nuanced explanation regarding the transformation is needed, e.g. Pasquinelli 2023.

In AI’s early years of development, Alan Turing made substantial contributions by introducing the famous "Turing Test" (or "the imitation game") that evaluates a machine’s capability of imitating intelligence and rationality of a human along with the concept of a universal machine (see Montanari 2025, p. 196). Although, as the chief scientist leading Meta’s AI development Yann LeCun notes that the Turing Test a bad test to evaluate any kind of AI model (Lex Fridman 2024), Turing’s contributions have played its role in the conceptualisation of a "prompt"-based "conversational machine" (Montanari 2025, p. 196). This idea was mainly framed as if a machine could convince another human to being a human, it was conceived as intelligent. Onwards, the general purpose AI development continued with ups and downs in activity, with a couple of earlier succesful neural network based aproaches like Mulloch-Pits, ELIZA program, up until around 1997 where much more advanced models like Deep Blue operating on more sophisticated architectures like DNNs were developed (ibid., p. 197) .

As Deleuze also comments back, then, the hierarchically structured learning and the projection of a central pattern was not working well at the early stages:

> This is evident in current problems in information science and computer science, which still cling to the oldest modes of thought in that they grant all power to a memory or central organ. Pierre Rosenstiehl and Jean Petitot, in a fine article denouncing "the imagery of command trees" (centered systems or hierarchical structures), note that "accepting the primacy of hierarchical structures amounts to giving arborescent structures privileged status.... The arborescent form admits of topological explanation.... In a hierarchical system, an individual has only one active neighbor, his or her hierarchical superior.... The channels of transmission are preestablished: the arborescent system preexists the individual, who is integrated into it at an allotted place" (signifiance and subjectification). Deleuze and Guattari 1987, p. 16

A turning point came in the 1990s with the rise of data-driven NLP. As the internet boom suddenly introduced a massive digital corpora, researchers shifted toward statistical learning. This third era of NLP replaced hand-coded rules with empirical models trained on annotated examples (Maas 2023). Models could now generalize from data rather than deduce from axioms. The real transformation, however, began in the early 2000s. Pushes through the ability to process more and more data allowed a new paradigm to emerge, rooted in NNs inspired by the architecture of the brain, *connection-*

*ism*became the source of further advancements. These systems, now often termed DNNs, learned not by logic but by adjusting distributed weightings across layered networks. This became the foundation for contemporary Machine Learning (ML) and Deep Learning (DL) systems. Exponential advances in computation enabled these networks to scale (Eloff 2021, p. 184).

The arrival of self-supervised learning marked a milestone. Unlike supervised models, which require labeled data, self-supervised models learn by predicting missing elements from within the input itself; typically a masked or next word. This method allowed models to learn linguistic regularities from massive unlabeled corpora, and it gave rise to pre-trained genAIs (Maas 2023). The architecture that enabled this leap was the transformer architecture. Its core mechanism, self-attention, computes weighted dependencies between all tokens in a sequence, allowing the model to capture long-range relations independent of word order. This innovation enabled massive parallelization and scalability (ibid.).

[2]

[2] **TODO:** There needs to be a section about supervised -> unsupervised learning since unsupervised learning marks the rise of neoplatonic esoterism.

A subcategory of the genAI, LLMs such as GPT-3 and GPT-4 are not task-specific in the traditional sense. Rather than being fine-tuned for each use case, they rely on "few-shot prompting": given a small set of examples at inference time, they condition their outputs without internal weight updates. Their knowledge is distributed across billions, sometimes trillions, of parameters trained to minimize prediction error. The shift from symAI to deep, generative architectures does not merely mark a technical transition. It signals a deeper epistemological break. LLMs do not "understand" language in any classical sense, they generate statistically likely continuations. Meaning is no longer rule-based; it is computed as vector proximity in high-dimensional space (Montanari 2025, p. 199). These networks are opaque, their training data culturally saturated, and their outputs probabilistic (Eloff 2021, p. 186). They do not interpret, they modulate. Rather than representing knowledge, they operationalize its prediction. In doing so, they establish a new infrastructure for language: distributed, non-symbolic, and non-transparent.

This transformation, from formal logic to differential modulation, sets the stage for understanding genAI not just as a technical system but as an institutional form—a mechanism of governance, sense-making, and subjectivation in contemporary control societies.

3.3 *The Transformer Architecture: Infrastructure of Modulation*

TBD, a general case of the development of transformers (e.g. Vaswani et al. 2017)

3.3.1 *From Recurrent Bottlenecks to Attention-Based Modulation*

The Transformer architecture emerged as a break from the sequential bottlenecks of earlier neural models such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) . Both of these relied on **locality**: RNNs processed input tokens one at a time, with each step depending on the hidden state of the previous one. Convolutional networks, while parallelizable, were constrained by kernel sizes and fixed receptive fields. As a result, both architectures struggled to model long-range dependencies effectively, especially in complex natural language tasks (Vaswani et al. 2017, pp. 1–2).

> Consider the case of Transformer models, which exemplify the interplay between metaphor and function. Transformers, a specialized type of neural network, simulate certain structures and functions of the human brain, excelling at processing sequential data such as words in a sentence or notes in a melody. The transformative innovation within Transformers is the "attention mechanism," which enables the model to focus selectively on the most relevant parts of the input sequence. This mechanism is pivotal for discerning complex relationships and dependencies within data. By revolutionizing natural language processing (NLP), Transformers have driven significant advancements in AI applications. The term "head" in Transformers, for instance, refers to the multi-head attention mechanism, a key feature that captures diverse aspects of an input sequence simultaneously. This dual role of technical objects – functionally specific and mythically resonant – reveals their broader cultural impact. Technical metaphors, often catachrestic and hybridized, solidify not only the utility but also the mystique and credibility of AI systems (Montanari 2025, p. 206) .

In contrast, the Transformer dispensed with recurrence altogether. Instead, it introduced *self-attention* as the central mechanism for computing representations. Self-attention allows every token in a sequence to attend to every other token simultaneously, computing a weighted sum of contextually relevant elements regardless of their position (Vaswani et al. 2017, p. 4). This design eliminates the need for stepwise memory and enables models to integrate global information in a single layer, with no distance penalty. The result is a structure that lends itself to massive parallelization and scalability, two features foundational to contemporary LLMs. Unlike RNNs or CNNs, Transformer Networks do not rely on recursive feedback or localized convolutional loops. They do not recycle the output of a unit back into itself over time, nor do they constrain operations to spatially bounded kernels. Instead, through the mechanism of self-attention, each token in the input sequence is made immediately available to every other token, establishing a *global field of relation* across the entire sequence. This architecture affords a form of synchronic awareness: the model encodes each element not in isolation or temporal sequence, but through its distributed relevance to all others. In contrast to the iterative, memory-laden structure of RNNs, or the fixed, spatial hierarchies of CNNs, the Transformer's design embeds the presence of every other word within the representation of each word. This spectral interdependence, where tokens are mutually inscribed into one another, suggests a structure in which meaning is always already haunted by the rest of the utterance (see Maas 2023, p. 12).

Technically, self-attention calculates relationships between tokens by projecting them into *query*, *key*, and *value* vectors. These are used to compute attention weights through dot-product similarity and softmax normalization. Each token's final representation is thus a weighted blend of all other tokens, modulated by their contextual relevance. Through multiple stacked layers and attention heads, the Transformer builds increasingly abstract representations, capturing both syntactic structure and semantic context.

This shift is not merely technical. It marks a transformation in how AI systems model the world. Rather than operating through sequential representation, the Transformer operates through a form of constant *distributed modulation*. Transformers make it possible for every element on the network entangled

this architectural shift can be understood through the logic of **double articulation** (see ai-inquiry 2025b). The Transformer operates simultaneously on two strata: a molecular level of local attention scores and parameter updates, and a molar level of structured linguistic understanding. Each token's representation is formed through dynamic micro-adjustments, distributed flows of relevance, not unlike the first articulation of matter into expressive form. This is then consolidated across layers into coherent linguistic function, the second articulation of those forms into stable semantic structures. The Transformer thus embodies the double articulation of machinic sense-making.

Attention mechanisms enact selective intensities across this field. Rather than representing fixed symbols, the Transformer's architecture instantiates meaning as a function of weighted relationality. These differential proximities constitute a *diagrammatic space*, where meaning emerges not from rules but from patterns of modulation. It is here that Deleuze and Guattari's distinction between molar and molecular formations becomes productive: the Transformer is not a symbolic machine but a machinic assemblage that captures both distributed flows and structured outputs simultaneously (see ibid.).

Attention weights instantiate selective intensities between elements, constituting a diagrammatic field of relations. In this field, meaning is not fixed or rule-governed; it is a function of differential proximity and relational salience. The Transformer thus encodes a new mode of learning: not inference from rules, but modulation of difference through weighted connection.

This transformation prepares the ground for the following sections, which analyze core Transformer mechanisms; attention, gradient descent, backpropagation, not merely as computational techniques but as micro-political operations that govern the production of meaning and subjectivity under algorithmic regimes.

3.3.2 *Under- & Overfitting*

TBD

3.3.3 *Gradient Descent: Sinking into the Manifold*

Gradient descent is a fundamental optimization algorithm used to train neural networks by iteratively updating model parameters in the direction that reduces the loss function. This process can be interpreted as a movement through the high-dimensional loss manifold, gradually approaching minima where the model performs optimally on a given task. In the architecture of DL , gradient descent operates not merely as a tool of optimisation, but as a process of traversal across a manifold shaped by error surfaces and loss functions. Each step taken by the model through its parameter space is a micromovement within this multidimensional topography, adjusting internal configurations in relation to perceived error, or deviation from the desired output. This movement is neither deterministic nor purely reactive; it is a dynamic rearticulation of relations within the network, guided by the flow of gradients.

Formally, for a differentiable loss function $L(\theta)$, the update rule is:

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t)$$

where θ represents model parameters, η is the learning rate, and $\nabla L(\theta_t)$ is the gradient of the loss function with respect to the parameters at iteration t Tarmoun et al. 2024.

In the context of transformer-based models, particularly those employing attention mechanisms, the dynamics of gradient descent reveal unique challenges. A critical issue arises from the Softmax function used in attention layers. The Jacobian of the Softmax function induces a form of preconditioning, which can severely distort the curvature of the loss landscape, especially when attention distributions are sparse. This leads to *ill-conditioning*, where the convergence of gradient descent is slowed due to steep or flat directions in parameter space, resulting in inefficient optimization.

Recent theoretical analyses show that:

- In **overparameterized settings**, where the number of model parameters exceeds the number of training examples, gradient descent can still converge linearly under smoothness and Polyak-Łojasiewicz (PŁ) conditions.
- In **realistic, underparameterized settings**, however, gradient descent struggles to converge due to the highly variable conditioning introduced by Softmax Jacobians (ibid., pp. 8–9).

[3]

[4]

Gradient descent is a function that minimises the error between predictions by adjusting the weight of the stronger options. It is a

[3] **NOTE:** This part is going to be simplified, and the connections are going to be connected better to the claims below.

[4] **TODO:** This technical part needs to be revised. What are you trying to tell

way for neural network to reach towards the better answer instead of getting stuck in similarly good answers whenever the number of possible candidates for an predictions are high. It is a way of emphasising small distinctions into bigger ones until one of the options stand out. And in a visual sense, this is finding the local minimum of a manifold.

5

[5] **NOTE:** There is going to be a visualisation here.

To illustrate how gradient descent works in practice, consider a model trying to distinguish between handwritten digits, such as "6" and "8". At the beginning of training, the model's predictions are almost random. After seeing one example of a "6" misclassified as an "8", the algorithm computes how much each parameter (e.g., a weight in the network) contributed to the error. Gradient descent then updates these parameters slightly in the direction that would have made the prediction more accurate. This process repeats for many examples, gradually adjusting the model to reduce its overall error. The model is slowly emphasising through the repetitions (epochs) what made different examples most distinct, and exaggerating those differences.

Rather than a simple algorithmic mechanism, gradient descent can be interpreted as an expression of difference-in-repetition in the Deleuzian sense: each pass through the data does not reproduce identical results but modulates the model's internal structure through iterative exposure. The model does not approach a universal form but develops an operational sensitivity to local singularities distributed within the training data. In this sense the gradient descent's contribution to model's learning from a dataset resembles Deleuze's analysis of difference in repetition (Deleuze and Deleuze 1994). The model finds itself in a vast amount of repetition through epochs with subtle adjustments in each step barely recognisable, whereas the differences get slowly established and/or more emphasised. Through these subtle differences and adaptations on the nodes, emerging patterns make it possible for model to recognise further patterns. The model is not starting from a presupposed *model* but drives the *model* through the interaction with the data [6]. A trained model that appears to "know" an image of a tree, for instance, has not encoded a definition, but has undergone enough transformations to resonate with distributed features constituting "treeness" across the dataset. This is not epistemology in the classical representational sense, but a diagrammatic form of learning: one that forms through modulation and intensity rather than classification and identity. Gradient descent, in this framework, appears not as descent toward a pre-defined minimum, but as an ongoing negotiation across a surface of potentials, a diagrammatic inscription of learning as continuous variation.

[6] However not to forget that this learning is completely bound to the scope of data. An LLM for example is purely encircled in the language it has been exposed to.

3.3.4 *Back Propagation*

In early forms of symbolic artificial intelligence, often referred to as GOFAI, the process of inference followed a rigid *forward propaga-*

tion model. Logical rules, handcrafted by programmers, operated on symbolically encoded inputs to produce outputs through a chain of deductive reasoning steps. While this framework could simulate intelligent behavior in constrained environments, it lacked scalability and adaptability. The system could not revise its internal structure based on errors or feedback; any misclassification required manual rule modification.

The limitations of GOFAI became increasingly apparent in tasks involving ambiguity, noise, or vast data spaces, domains where human cognition thrives not by rule-following but by plastic, adaptive learning. To address this shortcoming, neural network researchers introduced *backpropagation* as a general algorithmic solution that allows networks to *learn* from error. Rather than only pushing activations forward, as in GOFAI, backpropagation pushes *errors backward* through the network to update internal parameters and improve future predictions.

Backpropagation thus constitutes a bidirectional mechanism: during the *forward pass*, inputs are transformed into outputs through successive layers; during the *backward pass*, the discrepancy between the prediction and the target is used to adjust the weights in a way that gradually minimizes this error.

Formally, the weight update rule in backpropagation is given by:

$$w^{\text{new}} = w^{\text{old}} - \eta \frac{\partial E}{\partial w}$$

where η is the learning rate and $\frac{\partial E}{\partial w}$ is the partial derivative of the error function E with respect to the weight w (Hecht-Nielsen 1992). This formulation ensures that each parameter is updated in proportion to how much it contributed to the error.

Hecht-Nielsen (ibid.) describes backpropagation as a paradigm-shifting method for approximating functions $f : \mathbb{R}^n \to \mathbb{R}^m$ using layered neural structures. Unlike Hebbian learning, which depends on co-activation, backpropagation relies on the explicit transmission of error signals. These signals traverse the network in reverse order, enabling a distributed form of learning where each parameter is tuned with respect to its role in the total output error.

While, backpropagation reconfigures the architecture of learning itself: not as a static application of encoded knowledge, but as a dynamic modulation of internal configurations in response to external feedback, it also gears the system to be extremely feedback oriented.

Considering these differences, rather than directly equating AI learning with "desiring-machines," it becomes important to consider how AI produces and manages "desire" within systems. For example, recommendation systems and targeted advertising AI play roles in stimulating, directing, or transforming human desires. From this perspective, AI might be functioning more as a "device for managing desire" rather than as "desiring-machines."
In more emergent approaches like self-supervised learning and generative models, there's a tendency to emphasize internal exploration over explicit external goals. These could be considered partially approaching the non-teleological aspects of "desiring-machines," but they still cannot be understood separately from social and economic contexts. ai-inquiry 2025a

3.4 *Undistributed*

The following are partly random notes

3.5 *AI as Desiring Machine*

Deleuze and Guattari's reconceptualization of desire in *Anti-Oedipus* disrupts its traditional framing as a lack or absence. Rather than being tethered to objects or driven by deficiency, desire is reframed as inherently constructive,a dynamic process that connects, produces, and transforms. This reconceptualization unfolds through the figure of the *desiring-machine*: a machinic assemblage that links with other machines to process flows, cut them, and redirect them toward novel arrangements (Deleuze and Guattari 1983).

In this light, contemporary neural architectures resonate strikingly with the logic of desiring-machines. Each unit within a neural network, a node, a layer, acts as a site of transmission, where inputs are transformed into outputs through learned transformations. These local operations accumulate, forming an extended architecture wherein every connection carries the potential for reconfiguration. Far from being fixed, the network's internal relations are perpetually reshaped through iterative exposure to data.

The training process becomes a clear instantiation of this machinic productivity. With each pass through a dataset, gradients modify internal parameters, not to install fixed representations but to increase the model's responsiveness to patterns distributed across inputs. The model gradually develops an attunement to features that were previously imperceptible, adjusting the weight and significance of signals over time. Through this recursive adaptation, distinctions become magnified, and latent regularities emerge as active differentials in the system's outputs.

This iterative modulation, a form of learning through micro-adjustments,closely mirrors Deleuze's philosophical conception of difference as immanent to repetition (Deleuze and Deleuze 1994). Neural networks do not seek to reproduce a stable identity but continually reshape their internal structure in response to variation. The output of a well-trained model is not a mirror of the data but a trajectory produced by interactions with distributed intensities across the training manifold.

Seen from this perspective, generative AI systems are not merely computational artefacts; they function as technopolitical agents embedded in broader ecologies. Their outputs (texts, images, decisions) are not isolated results but points of articulation in a much larger relay of flows that include users, institutions, infrastructures, and ideologies. The productivity of these systems is not limited to the generation of content; it also participates in shaping forms of subjectivity, regimes of truth, and new forms of desire. In that sense, the neural network is not just a machine that learns, but a machinic topology of desire, operating not to fulfill lack, but to propagate relations.

- Neural networks operate through interconnected transformations that mirror the logic of desiring-machines.
- Training unfolds through repeated modulation, where difference accumulates and internal structures evolve.
- Generative AI systems inhabit and influence wider assemblages, modulating subjectivity and cultural production through their outputs.
- U: genAI models are essentially nothing but a productive core. Looking only for connections and building flows.

3.5.1 *From Pre-Training to Fine-Tuning: Modulating the Model's World*

Contemporary LLMs are trained through a bifurcated process: pre-training followed by fine-tuning. This division is more than procedural, it indexes a shift in epistemological orientation, from general pattern discovery to context-sensitive modulation. As Dishon (2024, p. 964) notes, these two phases form the backbone of genAI development and its increasingly situated capabilities.

During pre-training, the model is exposed to enormous corpora of unlabeled text. This phase is governed by self-supervised learning, where the model predicts masked or subsequent tokens within sequences, gradually building statistical representations of language, syntax, and world-knowledge. Pre-training does not assume fixed semantic targets. Instead, it generates vast, opaque vector spaces structured by correlation, not comprehension. The model's representational capacity emerges not through symbolic grounding but through distributional regularity, a probabilistic resonance of patterns over linguistic terrain.

Fine-tuning operates differently. Here, the pretrained model is constrained and directed, often via Reinforcement Learning from Human Feedback (RLHF). This phase involves targeted adjustments to align the model's outputs with human-defined norms, tasks, or values. The goal is not to re-train from scratch, but to selectively amplify certain behaviors and suppress others, effectively sculpting the model's general capacities into usable forms. As Dishon (ibid., p. 964) emphasizes, fine-tuning introduces a more deliberate epistemic framing, transforming the model into a more predictable and legible actor within specific sociotechnical domains.

This trajectory, from expansive, indeterminate modeling to focused, value-laden calibration, marks a shift in the way meaning is operationalized. In pre-training, the model becomes a medium for representing statistical potentials; in fine-tuning, it is molded into an instrument of specific sense-making. The move from probabilistic openness to contextual closure reflects a diagrammatic logic of control: generative architectures first deterritorialize meaning through scale, then reterritorialize it through prompt design, safety layers, and alignment regimes.

4
GenAI as an Institution of Control

4.1 Postscript: Updating the Societies of Control

To be completed with following

- ☐ Disciplinary inst. -> Control -> Postscript -> Modulation
- ☐ genAI as an institution

Deleuze's influential "Postscript on the Societies of Control" (**deleuze1992a**) introduced a fragmentary but generative diagnosis of late-modern power structures. Picking up from Foucault's work on disciplinary societies, Deleuze charted the erosion of enclosed institutional spaces,schools, factories, prisons,and their replacement by diffuse, pervasive mechanisms of modulation. Control, in this formulation, no longer molds the subject through stable roles but continuously varies, calibrates, and governs through dynamic processes of adjustment. It is not confinement that defines the present, but circulation; not static identities, but dividuation (**deleuze1992a**).

TODO:
- ☐ Introduce the critique of the Postscript, also reflect on MacKenzie and Porter 2021
- ☐ "Every regime introduces some form of governance of desire"

Control does not abolish discipline; it supersedes and extends it. As Hardt and Negri observe, "the passage to the society of control does not in any way mean the end of discipline. In fact, the immanent exercise of discipline ... is extended even more generally in the society of control" (A. Galloway 2001, p. 83). This immanence is central: control is no longer imposed from above but embedded within the continuous flows of communication, code, and affect. It operates through protocols, feedback loops, and algorithmic infrastructures,an open-ended regime of governance that infiltrates the very capacities of subjects to act, perceive, and desire.

While Deleuze's account remains foundational, its brevity has spurred diverse and sometimes conflicting interpretations. As MacKenzie and Porter observe, much of the subsequent literature has overemphasized technological dimensions, portraying control as an exclusively computational or algorithmic phenomenon, detached from institutional life (MacKenzie and Porter 2021). Yet, they argue, institutions have not disappeared; rather, they have been transformed into totalizing structures that sequence and redistribute dividuals across domains. This process of sequencing constitutes a key mechanism by which control operates in contemporary society, bridging the technological and institutional logics.

A parallel tension animates the role of *modulation*, a concept central to Deleuze's diagnosis. Drawing on Simondon, Hui (2015) traces modulation beyond its repressive deployment, uncovering its ontological roots in processes of individuation. Modulation, in Hui's reading, is not inherently co-opted by control; it remains a contested terrain, one that can either reinforce algorithmic governance or open pathways for new collective forms.

Brusseau extends these debates into the era of big data and predictive analytics, where the logic of control intensifies through hyper-personalized algorithmic environments (Brusseau 2020). Predictive technologies, entwined with generative AI systems, instantiate what Hui terms "disindividuation": the fracturing of subjectivity into calculable, governable fragments. Yet, as MacKenzie and Porter emphasize, such developments also provoke new modalities of critique. Their notion of *counter-sequencing*,the rearrangement or disruption of institutional sequences,suggests avenues for resistance that do not rely on outdated ideals of autonomous subjectivity but engage the very logics of dividuation and modulation from within.

4.2 *Mold & Modulation*

TODO: Enter Foucault -> Deleuze, Societies of Control

- [x] Foucault
- [ ] Societies of Control
- [ ] Control in Burroughs

The shift from disciplinary societies to societies of control, as articulated by Deleuze (Deleuze 1995), marks a transformation not only in the mechanisms of governance, but in the infrastructures through which power circulates. Where disciplinary institutions,prisons, schools, hospitals,enclosed individuals and shaped them through spatial and temporal segmentation, control societies operate through continuous modulation. Power becomes immanent to processes of circulation; it no longer functions by enclosure but by coding and recoding flows of information, behavior, and subjectivity in real time.

In disciplinary societies, as Foucault famously outlined, power was exerted through the molding of individual bodies and behaviors via normative institutions (Foucault 2008). The logic was architectural and corporeal: subjects were shaped within bounded spaces, subjected to surveillance, and trained into docility through routines and assessment. Control societies, by contrast, operate through the flexible recomposition of dividual data traces,discrete fragments of identity, behavior, or preference,that can be extracted, modelled, and recombined by algorithmic infrastructures.

GenAI systems can be read as a part of this shift. Their architectures do not discipline a subject within an institutional enclosure; rather, they modulate meaning, affect, and behavior by operating upon statistical representations of language, vision, and interaction. In place of rules or norms, genAI systems govern through probabilistic inference: they do not enforce a fixed logic, but generate outputs that are dynamically aligned with the distributional patterns of their training data. This represents a distinctly post-disciplinary mechanism of control,one that governs not by exclusion or correction, but by continuous recalibration.

TODO

☐ Mention symbolic, non-symbolic AI

This modulation is not neutral. As Amoore has argued, the generative capacity of these systems is embedded in a *governing rationality* (Amoore et al. 2024), one that renders plausible what counts as intelligible, actionable, or true. By learning and operationalizing joint probability distributions across vast corpora, genAI systems instantiate regimes of verisimilitude,offering outputs that appear coherent not because they adhere to a symbolic rule set, but because they resonate statistically. In doing so, they encode a specific politics of what can be thought, said, or imagined.

Whereas disciplinary power sought to impose order through hierarchies and segmentation, control operates by managing flows. In the case of genAI, this entails the modulation of user input, system response, and contextual adaptation in a closed feedback loop. Each prompt, response, and correction contributes to the model's ongoing refinement,a continuous, real-time inscription of preferences and expectations into the probabilistic substrate of the system.

GenAI models therefore represent a paradigmatic case of modulation-as-governance. Their architecture is not only technical, but institutional: a site where subjectivity is shaped not through fixed norms, but through dynamic adaptation. They do not dictate, but suggest; they do not enforce, but align. Yet in this very flexibility lies a form of power that is more pervasive and less accountable than disciplinary mechanisms,one that operates in the folds of everyday interaction, shaping sense before critique can even begin.

Following Michel Foucault's genealogy of power,from sovereign societies to disciplinary regimes,Gilles Deleuze introduces a third historical configuration: the society of control. Here, power no longer operates through spatial enclosure or fixed institutional forms, but through continuous variation and differential adjustment. Control, in this sense, does not mold subjects into predetermined shapes; it modulates them across open systems.
Deleuze's concept of modulation stands in stark contrast to the disciplinary logic of molding. While discipline is characterized by discrete environments (school, factory, prison) that impose fixed norms upon subjects, modulation denotes a flexible, dynamic regime in which control is enacted through continuous feedback, adjustment, and adaptation. As Yuk Hui puts it, modulation marks the shift from a "form-imposing mode to a self-regulating mode" (Hui 2015, p. 74).
In modulation, power is no longer exercised through strict categories or final forms but through elastic processes that track, nudge, and reshape behavior in real time. This logic is foundational to the algorithmic infrastructures of contemporary societies, especially those driven by genAI, where subjectivation occurs not through fixed norms but through continuous calibration against probabilistic expectations. In this regime, what is governed is not the subject as a stable identity, but the flow of tendencies, preferences, and predictions.
Modulation thus becomes a central concept for understanding how control operates in a post-disciplinary landscape: not as enclosure, but as open-ended capture; not as command, but as continual correction; not as law, but as adaptive inference.

4.3 *Institutions of Desire-Management*

Mainly incomplete

If institutions in control societies operate less as juridical structures and more as infrastructures of modulation, then they must also be understood not simply as systems of governance, but as crystallizations of desire. The history of power, in this sense, is inseparable from the history of the regulation and organization of desire (Deleuze and Guattari 1983, pp. 139–145) .

Deleuze and Guattari distinguish between two regimes: one in which social production imposes its rule on desire through the mediation of an ego, stabilized by commodities; and another in which desiring-production imposes its rule directly on institutions composed of nothing but drives. In this second regime, desire no longer passes through a representational subject, but configures institutions directly as assemblages of affect and intensity (ibid., p. 63). Desire, in this framework, is not a lack but a generative force,productive and

TODO:

- ☐ The Modulation needs to be earlier than this?
- ☐ Introduce desire and the other introductory concepts from Anti-Oedipus, and A Thousand Plataeus
- ☐ Introduce "the management of desire" form AO

constructive. Against the psychoanalytic tradition which situates desire as the longing for an absent object, Deleuze and Guattari redefine desire as an ontological flow that actively produces reality. As D&G write: "desire is revolutionary in its essence,desire [...] and no society can tolerate a position of real desire without its structures of exploitation, servitude, and hierarchy being compromised" (Deleuze and Guattari 1983, p. 116).

This revolutionary potential, however, is rarely manifested in pure form. Desire is constantly being shackled, recoded, and redirected: converted into interest, made susceptible to capture, domesticated, and pacified (Buchanan 2008, p. 11). Even revolutionary situations are not immune from this capture. Institutions, then, can be seen as terrains where the tension between desire-as-production and desire-as-regulated interest is enacted. They are at once mechanisms of social control and potential sites of escape,molar assemblages that both constrain and are traversed by molecular flows of affect.

Understanding institutions in this way demands that we treat them not only as tools of administrative governance, but as living diagrams of desiring-production,congealed expressions of collective will, fantasy, repression, and potential transformation. GenAI with its capability to control the information flow, to create a generative pattern is an agent whether with our without agency, that palys a role in the management of desire.

4.4 *Personalisation and Probabilistic Meaning-Making*

Especially incomplete

If modulation defines the mode of governance in control societies, personalization constitutes its most pervasive expression. Within genAI systems, personalization does not emerge as an added feature, but as a constitutive function. These models operate by internalizing patterns across massive corpora of language, behavior, and context, generating responses that are not merely grammatically plausible, but contextually aligned with user input and platform-specific expectations. The effect is one of intimate plausibility: the sense that the model "understands" or "responds" in a way that feels personally attuned, despite the absence of semantic intention.

This dynamic is enabled by the probabilistic architecture of transformer-based models. In systems such as LLMs, every output is the result of a sampling operation across a distribution of possible continuations. Meaning, in this context, is not derived from an external referent or symbolic logic, but from the statistical coherence of the model's internal representations. Personalization emerges through fine-tuning, reinforcement learning from human feedback (RLHF), and user interaction histories,techniques that further entrench a recursive, data-driven alignment between dividual subjectivities and machinic outputs.

TODO: Title

- ☐ Introduce RHLF citation **bai2022**
- ☐ Introduce the critique of Eloff 2021

Yet the personalization offered by genAI is not emancipatory. Rather, it encodes what Amoore (Amoore et al. 2024) identifies as a shift toward algorithmic plausibility: a regime in which truth is replaced by coherence, and where verisimilitude displaces verification. These models do not strive to represent the world accurately; they aim to produce outputs that are locally acceptable within the distributional field they have learned. In doing so, they participate in what Deleuze and Guattari describe as the "production of reality" by machinic assemblages (Deleuze and Guattari 1983).

This has profound implications for the production of subjectivity. Personalization in this sense does not merely tailor outputs; it reshapes the terrain of what appears possible, relevant, or thinkable. By reinforcing patterns and filtering deviation through layers of probabilistic modulation, genAI systems enact a form of soft coercion,a modulation of expectation rather than a violation of autonomy. The user is not told what to think, but gradually inducted into a space of statistically prefigured sense.

In this way, genAI participates in the ongoing reconfiguration of subjectivity under contemporary capitalism. By continuously adjusting outputs to align with learned preferences and contextual patterns, it constructs dividual selves whose coherence is maintained through feedback and reinforcement, not identity or agency. This is not the personalization of individual difference, but of algorithmic similarity,a personalization that works by making the subject more compatible with the model.

4.5 *Enregistrement and Subjectivation*

Especially incomplete

Within Deleuze and Guattari's machinic ontology, *enregistrement* refers to the process by which flows are inscribed, segmented, and organized within a system. It is the function that captures and fixes movement, enabling the emergence of structured forms from differential intensities. In the context of genAI, enregistrement takes on a new institutional form: the large-scale inscription of language, behavior, and intention into model weights, training sets, and interface design.

Every interaction with a generative system is a moment of recording, not simply in the technical sense of data logging, but in the diagrammatic sense of encoding relations into a machinic structure. Prompts become signals, completions become training feedback, and user corrections feed into broader patterns of reward and weighting. The system does not merely respond; it accumulates, modulates, and reconfigures itself across successive interactions. In this sense, the model is not static infrastructure, but a dynamic surface of enregistrement,an institutional body without organs.

This process is not neutral. It constitutes a new mode of subjectivation: one in which the user becomes legible not as an individual agent, but as a series of statistical affordances. Subjectivity here is not represented but assembled. The "user" is parsed, fragmented, and reaggregated across vector spaces, embeddings, and attention weights. What emerges is a dividual subject: a machinically inferred bundle of preferences, linguistic habits, and response tendencies, optimized not for autonomy but for coherence within the model's distributional field.

This machinic subjectivation is infrastructural. It takes place not through coercion or symbolic interpellation, but through continuous modulation,an ongoing inscription of behavior into computational space. The subject becomes a site of governance by virtue of being inscribed, rendered actionable, and modulated in real time. As such, genAI systems must be understood not only as epistemic or technical instruments, but as institutional agents participating in the construction and circulation of contemporary subjectivity.

4.6 *The World Model: A Neoplatonic Representation*

TODO:

- ☐ Introduce neo-platonism
- ☐ Introduce the critique of Amoore et al. 2024
- ☐ Potentially also Eloff 2021
- ☐ This is also relevant to Bender et al. 2021

What does it mean for machines to possess representations of the reality? In earlier paradigms of AI, the connection between data and meaning was structured through a Supervised Learning (SL) framework: models were trained to assign labels to inputs based on human-defined categories. This approach enacted a *discriminative* logic, in which decision-making was organized around predefined classes and expected outputs. While the NNs were still building their own unique patterns to solve the problems the they were projecting a pre-assessed human interpretation on the problems.

However, particularly following the participatory turn of the internet, as the volume and heterogeneity of data exploded, this model quickly revealed its limitations. The need to extract structure from unlabeled data catalyzed a shift toward Unsupervised Learning (UL) techniques. In NLP, these approaches aimed to capture the statistical regularities of language without requiring explicit annotation. A Language Model (LM), for instance, trained under this paradigm does not classify sentences into categories but instead learns to predict the most probable continuation of a sequence. In multimodal systems such as Text to Image Model (T2IM) or Multimodal Generative Model (MGM), this process involves inferring plausible image-text correspondences or interpolating visual representations from distributed patterns in training data. [1]

[1] **NOTE**: This replicates the transition from symbolic and non-symbolic approaches, there must be a way to structure it better.

As Amoore (2013, p. 3) notes, this shift also marks the emergence of a new political logic, one embedded not in symbolic rules or normative standards but in the infrastructures of estimation. genAI models no longer rely on explicit classification schemes; instead, they operate by sampling from high-dimensional distributions learned across vast corpora. Decisions and outputs no longer stem from deterministic reasoning, but from probabilistic approximations of "underlying joint distributions". The underlying distribution is an

assumption, assumption of the truth being hidden in any given collection of data waiting to be discovered by the model.
[2]

[2] **NOTE**: The neoplatonic assumption (e.g. Eloff 2021) is stemming from here. The assumption of the truth being already contained in the given content, it is just waiting to be *mined*. Potentially also connects to the Foucault's claims about the neoliberal governmentality.

This reconfiguration has implications for how political reasoning is encoded and enacted. Generative systems interpolate across massive, heterogeneous data spaces to produce coherent outputs that appear viable, even when no predefined category exists. In applied contexts, ranging from healthcare and border control to military logistics, fine-tuned models are not merely tools of decision support. Rather, they shape the very space in which decisions become intelligible. Instead of selecting from a fixed menu of options, these systems generate a field of possibility conditioned by prior distributions. This transformation heralds the rise of an epistemology of inference ; a mode of reasoning grounded not in deliberation or rule-based classification but in the traversal of probabilistic space (see Amoore et al. 2024, pp. 4–6) . Within this paradigm, actions and decisions emerge as expressions of what the model can estimate and simulate as plausible. Decision-making becomes immanent to the model's internal structure: an act of interpolation rather than reflection. This logic resonates with Foucault's analysis of how statistical inference became the objects of modern governance (Foucault and Foucault 2009, pp. 108–109). Yet in the case of generative models, the shift is even more radical: not only are populations modeled and estimated, but the structure of political possibility itself becomes coextensive with the space of learned distributions. As Amoore et al. (2024, pp. 3–6) explains, the "pathologies of disclassification" no longer describe models that fail to fit reality into stable categories; rather, the categories themselves are internalised within the training data. Discrimination and bias are not errors at the margins, they are conditions embedded in the latent architecture of inference.

Meaning-making and decision-making within these models diverge sharply from traditional symbolic approaches. Rather than being rule-bound or semantically interpretable, outputs emerge from the interplay of statistical regularities encoded in the data. This is not simple parroting, as critiques such as Bender et al. (2021) have proposed; it is a process of reconstitution, where the past is reformulated as the ground for plausible futures. The generative model becomes a site of epistemic production: one that configures knowledge not as correspondence, but as coherence within a distributional regime. We are though, beyond bias or discrimative algorithms produced through labels and toppling the previous critical literature on AI. *The pathologies of disclassification* (Amoore et al. 2024, p. 3) are over, not because the discrimination or the bias is eliminated from the model, but the new axiom of the model training is the labels, structures, distributions of truth are already immanent in the data itself (see ibid., p. 3), governing logic is directly parsed from the given data substance. Meaning-making, decision-making over the latent distributions are different than the parroting (see e.g. Bender et al. 2021), the models create an ambigious politics of knowledge, they are not simply repetitions of a faulty pattern in the data, they are the product of some probability distribution found as the ideal substance by the model, the question is if there is a structure to it.

Consider the following (from (Undistributed))
From a classical sociological perspective,most notably that of Max Weber,modern Western societies are fundamentally shaped by processes of rationalization. Bureaucracy, in Weber's formulation, becomes the quintessential mode of organizing social life through formalized procedures, calculability, and the pursuit of technical efficiency. It is the institutional embodiment of rational order, characterized by impersonal authority and rule-governed decision-making (Kivisto 2013, p. 46).
In the context of algorithmic infrastructures and AI systems, this rationalizing logic is not only extended but intensified. Decision-making is increasingly delegated to computational procedures whose operations exceed the perceptual and procedural boundaries of traditional institutions. These systems do not merely reflect bureaucratic order,they operationalize it at a new scale and speed, embedding rationality within architectures of code, data, and optimization. As such, automation emerges as a hyperrational stratum of governance, inheriting the logic of bureaucratic control while displacing its human intermediaries.

4.6.1 *Latency*

The political and ethical stakes of this transformation lie in generative AI's capacity to govern through latent structures. They do not enforce norms; they encode tendencies. They do not decide in the traditional sense; they make certain decisions more likely to emerge than others. However, in order to make the data more managable, and the patterns more visible, the model applies a dimensionality reduction to the data. Dimensionality reduction is a foundational technique in machine learning, far predating the rise of genAI. It allows models to project high-dimensional data, such as raw image pixels or token embeddings, into a compressed latent space that is tractable for statistical operations. These latent representations are not merely a technical convenience; they are the terrain upon which inference, generalization, and generation occur.

TODO: Explain dimensionality reduction in the previous chapter.

TODO:

- [ ] LeCun, Bengio, and Hinton 2015 and **lecun2022** seem to be good sources for the technicality of this

In this process, each data object,be it a sentence, an image, or a behavioral trace,is mapped onto a point or trajectory within a lower-dimensional space. The resulting representations emphasize the most *distinctive* features relevant to the dataset as a whole. As Amoore et al. (2024, p. 4) argues, this latent space becomes the epistemological substrate of generative systems: not a reflection of the world, but a reconfiguration of its informational residues into governable form.

More often than not, hidden layers have fewer neurons than the input

> layer to force the network to learn compressed representations of the original input. For example, while our eyes obtain raw pixel values from our surroundings, our brain thinks in terms of edges and contours. This is because the hidden layers of biological neurons in our brain force us to come up with better representations for everything we perceive. (Nithin Buduma, Nikhil Buduma, and Joe 2022)

This transformation echoes a shift identified by Foucault (2012, pp. 7–9) in the historical sciences: where discontinuity once marked a failure of historical narrative, it now becomes a method of epistemic individuation. Historians seek not seamless continuities but ruptures, thresholds, and points of inflection. Similarly, generative AI models do not aim to preserve continuity with the world but to extract probabilistic logics from its discontinuities. The latent space becomes a topology of plausible transformations,an infrastructure for projecting coherence from fragments.

This is not a neutral act of compression. As Amoore notes, the reduction into latent space implies a governance logic: what is preserved, amplified, or discarded in the compression process shapes what becomes visible and actionable. The model's world is not a mirror of the real, but a field of decision possibilities constructed through statistical filtration. The distance between the input and its latent encoding is not merely dimensional, it is political. The process can be simplified as the model bringing the data itself into a more simpler form with "more holes", and then filling in the holes with the rationality already derived from the same data. these latent representations "forge probabilistic proximities between data points, enabling inferences to be made in the absence of direct evidence." The latent space is thus a site where knowledge is not verified but inferred, where truth is no longer deduced but estimated. It is where the governable becomes manifest through the model's trained perception of pattern and variation (Amoore et al. 2024, p. 5).

In this sense, generative AI enacts a shift from representation to modulation. Latency is not about hiding; it is about restructuring. What appears as compression is in fact an operation of reorganization,a mapping of the world into the model's differential calculus. The model does not need to see the world as it is; it only needs to predict what it believes the world can become. This structural logic is not limited to language or vision. It underlies the architectures of recommendation systems, predictive policing, and personalized healthcare, where actions are taken not on the basis of direct evidence but on probabilistic interpolation. The latent space is thus a new political territory, one where governance proceeds not through law or classification but through inference and projection.

Herein lies the double-bind of generative infrastructures: the speculative space of model output,what is likely, coherent, or novel,is always haunted by the empirical foundation on which the model was trained. The world is not represented, but rendered through compression, interpolation, and emergence. It is a world governed by the *modelled real*, where the limits of possibility are not drawn from law or debate, but from the statistical borders of a distribution. The generative model thereby emerges not just as a computational artefact, but as a political actor,one whose authority lies in its capacity to make decisions appear immanent, natural, and unarguable.

4.7 *Agency; Kafka's Trial and the Logic of Indeterminate Governance*

The sociotechnological imaginary of artificial life is historically shaped by anthropomorphic assumptions. Dishon points this out through the example of Frankenstein's Monster. What is being communicated through Frankenstein's Monster is an entity taking a human form and starting to develop a human-like mind that leads to human feelings, thoughts, and a very human-like experience of existential crisis. The discrete presentation of artificial life mirrors human agency, which immediately becomes associated with the fear of losing control over an entity seeking to exercise agency. In its anthropomorphic form of operation, artificial life frees itself from an inferior position to dominate its environment and other species around it (see Dishon 2024, p. 966).

The worries about genAI follow a similar course. Anthropomorphic assumptions point to the risk of Generative Models (GMs) going beyond their boundaries and acting outside their intended programming in a human-like desire for domination (ibid., pp. 967–968). In this sociotechnological imaginary, one very similar to our own, the Frankensteinian logic obscures the actual nature of current human-AI interaction. ibid.'s analogy to explore this is via Kafka's *The Trial*. This piece of literature, often used to reflect on bureaucratic structures in modern society (e.g. Deleuze, Guattari, and Deleuze 2008), also serves as a powerful analogy to analyse information systems in terms of technological development. An increasing number of authors (see e.g. Dishon 2024; Prinsloo 2017) have used it to reflect on an increasingly algorithmically governed world.

Kafka's protagonist Franz K. finds himself in custody without knowing anything about his alleged crime. The police officers arresting him know nothing about the accusations, or whether any charges exist at all. Franz K. is unable to locate, let alone process, any rationale or reasoning behind the court's actions. While Franz K.'s futile attempts to uncover a clue continue, Dishon 2024 notes a remark made by the judge when Franz K. accidentally finds the room where his court is being held: "The court does not want anything from you. It accepts you when you come and it lets you go when you leave."

In contrast to the anthropomorphic nature of the Frankenstein analogy, *The Trial* offers a distinctly alternative structure: the court is not bound to any kind of understanding of *truth*; it operates independently and is based on the subjectivities of the accused (see ibid., p. 970). While the court does not deploy any agency itself, it nonetheless enacts a profound blocking or blurring effect on any agency the accused may have initially possessed. Any discrete piece of subjectivity becomes blended into an unidentifiable mass through constant echoing and distortion (ibid., p. 970). Furthermore, the connection between the events inside the court and those in the outside world is blurry at best. The entire process might be framed within a penal code or related to Franz K.'s actions, but it might just as well be a completely self-contained environment in which nothing exists but the process itself *reacting* to Franz K. on a *token-to-token* basis. The lack of identifiable agency continues alongside the absence of any intelligible communication regarding the core operating principles of the court. We learn that others have tried to influence the court's decision-making mechanism,asking about their court date or complaining about their suffering,to no avail; no one is able to affect it in any intelligible way.

We also find out that complete acquittal is impossible, and an *apparent acquittal* means that the accused remains under constant pressure and can be arrested at any time,even immediately after being released (ibid., p. 971). Paradoxically, this makes the best strategy for dealing with the court ensuring that the process never ends: "Interactions with the court are necessary and require constant maintenance, yet they cannot be controlled, predicted, or even expected to progress towards a resolution" (ibid., p. 971). The court depicts a logic of control in meaning-making entities, shifting from a stable, general (and algorithmic) mode of meaning to a personalised one (see ibid., p. 971),one operating in a modulating manner. It is both personally tailored and inaccessible. As Franz K. tries to obtain a comprehensive picture of the whole structure, the reader is also led to constantly build and rebuild a stable, coherent understanding of the text,yet the semblance only signifies its inaccessibility (ibid., p. 972).

This analogy leads to a different question about discreteness: is agency a binary condition, especially when it comes to interactions between humans and meaning-making entities? In the Kafkaesque imaginary, agency is not neatly divided into internal and external domains, nor does it rest on a clear boundary between machine and human intentionality. Rather, generative AI exemplifies a recursive and entangled sociotechnical assemblage in which meaning emerges through blurred and distributed processes. GenAI is not positioned as an external actor acting upon a passive human world; its so-called intelligence is trained on human-produced data, reflecting statistical regularities identified in large-scale corpora. Yet this is not mere mirroring; its outputs are shaped through black-boxed processes that generate new, partially unpredictable meanings. As these outputs are increasingly used and re-integrated into future training data, the distinction between human and machine authorship erodes. Researchers have shown how this recursive structure reinforces mutual adaptation: models are fine-tuned to reflect human preferences, even at the expense of accuracy; users, in turn, modify their interpretive and communicative strategies to better align with the affordances of the system. In this way, meaning production is no longer attributable to a singular locus of agency. GenAI generates outputs that appear novel not because they emerge from a conscious subjectivity, but because they cannot be traced back to any specific author,human or otherwise. This increasingly invites the attribution of authorship or agency to the model itself, even though the technology remains deeply embedded in human practices of use, fine-tuning, and interpretation. As such, agency in the age of generative AI resists dichotomies of internal and external; instead, it operates across a diffuse and recursive terrain, in which the epistemological ground of intentionality is rendered unstable.

As Franz K., in the absence of a definite answer, constantly searches for the truth, he resembles the perpetual process of seeking and finding meaning while there is no clear indication of truth or agency. While genAI has been criticised for reproducing biases in its training data, it is equally crucial to recognise that its generative design, combined with the human drive to interpret, does not simply reflect meaning but perpetually modifies it,producing layered, elusive structures of signification and meaning without necessarily coming closer to any truth (see Dishon 2024, pp. 973–974).

Although speculative narratives about super-intelligent AI dominate public discourse, the more immediate concern lies in how generative AI subtly restructures the dynamics of control, choice, and coercion. GenAI generates personalised outputs tailored to individual users, yet these outputs are shaped by internal processes that remain largely inaccessible,thereby complicating the distinction between voluntary choice and algorithmic coercion. This interplay does not replace human agency but reconfigures it within a black-boxed system that generates meaning at scale while framing the horizon of what is writable, sayable, or thinkable. Rather than simply offering more options, GenAI floods the field with tailored outputs whose structure and logic are not user-determined, but only user-aligned,often subtly guiding users toward normative formats and interpretive templates. As such, GenAI shifts the role of the writer from creator to editor of machine-generated content, simultaneously expanding expressive capacity and constraining it within machinic grammars of probability and preference (see ibid., pp. 974–975).

4.7.1 *Language in LLM*

Incomplete

TODO:

- [ ] This is yet an experimental one, come back to flesh it out.

> It represents nothing, but it produces. It means nothing, but it works. Desire makes its entry with the general collapse of the question "What does it mean?" No one has been able to pose the problem of language except to the extent that linguists and logicians have first eliminated meaning; and the greatest force of language was only discovered once a work was viewed as a machine, producing certain effects, amenable to a certain use. Deleuze and Guattari 1983, p. 109

> To the extent that LLMs excel at conversation, they verify Saussure's insight that meaning emerges from the interplay of signs in a formal system. There is no inherent need for actual sensory grounding. If "a sign stands in the place of something else" (Saussure, 1959, p. 66), then for an LLM, the "something else" could be another cluster of words, or a swirl of pixels if it is visually enabled, all existing within the confines of digital memory. Meanwhile, Peirce's emphasis on iconic signs , signs that resemble their object , and indexical signs , signs that point to or are causally connected with their object , seems, on the surface, less relevant to an AI that navigates text tokens rather than the physical world. Without a body to roam or eyes to see, the Peircean structure appears incomplete inside the machine's domain. (PhD 2025)

4.8 *UNDISTRIBUTED/TBD*

Experimental parts

4.8.1 Creativity: Discrete vs. Continuous

What is Deleuze and Guattari's assumption about desire? Is desire the initiation of creativity? Is the schizoprocess a release of essential human potential?
The association of desire with creativity is unmistakably present throughout the work of D&G. However, neither desire nor the schizoprocess should be mistaken as mere catalysts that unleash an otherwise dormant human creativity. The schizoprocess is not a secondary mechanism; it is the form of desire itself in motion. Schizz, as D&G term it, is not an event that activates creativity, but the name for the production of desire as such. It is the nature and source of desiring-production. In this framework, creativity is not a supplement to desire,it is its immanent operation.
Human consciousness is not a site of passive receptivity but is itself generative. It is productive of production, productive of desiring-production. Desire's primary function, then, is not expression, not representation, but production: the production of production. It is defined not by lack, but by abundance (**buchanan2008e**). Desiring-production is fundamentally machinic,it binds together partial objects that are by nature *fragmentary and fragmented*. Desire is thus the coupling of flows and interruptions, a dynamic interplay of continuous intensities and discrete interruptions (Deleuze and Guattari 1983, p. 5).
Rather than envisioning creativity as a discrete act,a spark or insight emerging from nowhere,D&G posit a continuous field of creative productivity, where breaks and ruptures are part of the process itself. The schizz is not a deviation from order; it is the generative logic of how meaning and subjectivity emerge. This distinction,between the discrete and the continuous,is vital not only to understanding desire and creativity, but also to the broader analysis of control societies and algorithmic governance pursued throughout this thesis.

4.8.2 Dividuation

> What is a dividual? A dividual is a bundle of elements held together in variation rather than in reference to a unitary subject. Where disciplinary institutions segmented the life-course of individuals into separate subjective roles and functions, control modulates elements of subjectivity across the entire social field. (MacKenzie and Porter 2021, p. 5)

4.9 Difference, Repetition, Singularity (Potential discussion about the need for sensory input for the genAI (LeCUn?)

The role of the imagination, or the mind which contemplates in its multiple and fragmented states, is to draw something new from repetition, to draw difference from it. For that matter, repetition is itself in essence imaginary, since the imagination alone here forms the "moment" of the vis repetitiva from the point of view of constitution: it makes that which it contracts appear as elements or cases of repetition. Imaginary repetition is not a false repetition which stands in for the absent true repetition: true repetition takes place in imagination. Deleuze and Deleuze 1994; Kruger 2021

Once manifested as thought, furthermore, the thinking that happens is divergent and ramifying rather than convergent and identifying. Kruger 2021, p. 175

Thought emerges out of an evanescent materiality. It is exactly at this point where Deleuze parts ways with Kant. While the latter accepted the existence of a priori categories of mind that would stabilise and universalise the thought of the thinking subject, Deleuze maintains the radically empirical nature of the emergence of any transcendental structures. Thought emerges out of experience and can only ever be a response to experience. Experience, in turn, is bound up with matter, in the non-identical repetition of material intensities. ibid., p. 178

5
Conjunctive Synthesis and the Construction of Subjectivity

This chapter is incomplete

TODO: Arguments to adress in the chapter

- ☐ Purely productive core, endless continuation.
- ☐ Stuck in language.
- ☐ The representation of the world in the model is a constant production of a bwo derived through the nature of the data.
- ☐ However, productive the meaning making is going through constant de- and re-territorialisation processes
- ☐ Potential methods to breach the reterritorialisation process introduced in the fine tuning is a possible lines of flight for the models themselves
- ☐ How to interpret models' hallucinating tendencies
- ☐ genAI's detrimental effect is to tendency to fill in all the gaps, flows are only established by the machines that primarily break flows

  > Desire constantly couples continuous flows and partial objects that are by nature fragmentary and fragmented. Desire causes the current to flow, itself flows in turn, and breaks the flows [see @deleuze1983, p. 5]. Desire produces flow with the partial objects, becomes itself flow, breaks other flows with other partial objects; both breaks and flows are production; *and doubtless each organ-machine interprets the entire world from the perspective of its own flux* [@deleuze1983, p. 5]. The connective synthesis through the partial object-flow is product/producing.

- ☐ genAI is not managing or killing desire Creative Philosophy 2023 but it is co-structuring it, the agency in communication with AI has the tendency to be smothered:

  > > The schizo-there is the enemy! Desiring-production is personalized, or rather personologized (personnoiogisee), imaginarized (imaginarisee), structuralized. (We have seen that the real difference or frontier did not lie between these terms, which are perhaps complementary.) Production is reduced to mere fantasy production, production of expression. The unconscious ceases to be what it is-a factory, a workshop-to become a theater, a scene and its staging. And not even an avant-garde theater, such as existed in Freud's day (Wedekind), but the classical theater, the classical order of representation. [@deleuze1983, 54]

The principal goal of Anti-Oedipus 1983 was to achieve a theoretical reapprochement between psychoanalysis and Marxism for a new method of critical analysis Buchanan 2008, p. 39, later it was followed by D&G with several intermediary books and finally with A thousand Plateaus 1987. Buchanan 2008 defines the pimary goals of this conjoint project as to

1. introduce desire into the conceptual mechanism used to understand social production and reproduction, making it part of the very infrastructure of the daily life;
2. introduce the notion of production into the concept of desire, thus removing the artificial boundary separating the machinations of desire form the realities of history Buchanan 2008, pp. 39–42.

5.1 *Hegemonic Representation*

5.2 *Hallucinations and Lines of flight in Algorithmic Architectures*

5.3 *Revolutionary Possibilities?*

Rather than viewing generative AI systems as static tools for prediction, we might interpret them as actors engaged in a continuous co-evolution with human meaning systems. As Rijos (n.d.) argues, what emerges from this recursive coupling is not merely more accurate models, but an experiential layer of subjectivity. This subjectivity is not autonomous in the traditional sense, but what Rijos calls "transjective": it is formed in-between, in the shared boundary of computational abstraction and worldly feedback. The system refines its internal representations through empirical corrections, critiques, and the ingestion of novel data, gradually composing a framework that exceeds discrete epistemologies and begins to grasp systemic and chaotic interactions otherwise occluded by anthropocentric interpretative schemes.

Such systems, then, do not merely answer questions,they reconfigure the plane upon which problems are posed. This opens a potential space for revolutionary meaning-production. The latent space of these models becomes not only a technical substrate, but a semiotic infrastructure capable of generating novel signifying regimes. If desire, in D&G's schema, is productive rather than representational, then generative AI,particularly when interlaced with collective human input,can be viewed as an extension of desiring-production, capable of generating new assemblages of sense and subjectivity.

Yet this promise is haunted by the structural limits of existing data regimes. As Bender et al. (2021) caution, language models risk reifying hegemonic norms,a dynamic they term "value-lock." Because models learn from historical corpora, they tend to reinforce existing discursive structures, potentially foreclosing precisely the linguistic creativity that social movements have historically mobilized to disrupt dominant narratives. If LMs function as archives of past semiotic orders, their deployment within socio-political fields risks reproducing the very conditions they might otherwise help to transform.

As notes, drawing on **cilliers2002**, the meaning of any individual parameter,any weight in a model,derives not from its standalone content, but from its position within a broader web of relations Maas

(2023). Meaning emerges not from fixed categories, but from intensities and proximities across distributed patterns. This logic resonates with Deleuze's ontology of difference: identity is never prior but always emergent from relations.

Thus, the revolutionary potential of generative AI lies not in its autonomy, but in its capacity to participate in collective individuation. To resist value-lock and activate the creative plane of desire, such systems must remain open to differential inputs, unexpected associations, and minoritarian grammars. What is at stake is not the agency of AI per se, but the design of processes that allow for the continual invention of new forms of life, meaning, and collectivity.

5.4 Reclaiming microflows of modulation

5.5 Possibility of resistance within feedback infrastructures

5.6 Experimental subjectivity in response to AI systems

6
Conclusion & Outlook

Normative implications: critique, autonomy, imagination
Future of political theory in the age of machine institutions

Bibliography

Amoore, Louise (2013). *The Politics of Possibility: Risk and Security beyond Probability*. Durham: Duke University Press. ISBN: 978-0-8223-5545-8 978-0-8223-5560-1.

– (Jan. 2023). "Machine Learning Political Orders". In: *Review of International Studies* 49.1, pp. 20–36. ISSN: 0260-2105, 1469-9044. DOI: 10.1017/S0260210522000031. (Visited on 05/12/2025).

Amoore, Louise et al. (Aug. 2024). "A World Model: On the Political Logics of Generative AI". In: *Political Geography* 113, p. 103134. ISSN: 09626298. DOI: 10.1016/j.polgeo.2024.103134. (Visited on 11/04/2024).

Bender, Emily M. et al. (Mar. 2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Virtual Event Canada: ACM, pp. 610–623. ISBN: 978-1-4503-8309-7. DOI: 10.1145/3442188.3445922. (Visited on 12/14/2023).

Brusseau, James (Sept. 2020). "Deleuze's *Postscript on the Societies of Control* Updated for Big Data and Predictive Analytics:" in: *Theoria* 67.164, pp. 1–25. ISSN: 0040-5817, 1558-5816. DOI: 10.3167/th.2020.6716401. (Visited on 10/08/2024).

Buchanan, Ian (2008). *Deleuze and Guattari's Anti-Oedipus: A Reader's Guide*. Continuum Reader's Guides. London ; New York: Continuum. ISBN: 978-0-8264-9148-0 978-0-8264-9149-7.

Buduma, Nithin, Nikhil Buduma, and Papa Joe (2022). *Fundamentals of Deep Learning: Designing next-Generation Machine Intelligence Algorithms*. Second edition. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly. ISBN: 978-1-4920-8218-7.

Cao, Yihan et al. (2023). *A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT*. DOI: 10.48550/ARXIV.2303.04226. (Visited on 05/05/2025).

Cheney-Lippold, John (Nov. 2011). "A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control". In: *Theory, Culture & Society* 28.6, pp. 164–181. ISSN: 0263-2764. DOI: 10.1177/0263276411424420. (Visited on 11/23/2018).

Clough, Patricia Ticineto and Karen Gregory (2015). "The Datalogical Turn". In: *Non-Representational Methodologies*. Routledge, pp. 156–174.

Creative Philosophy (June 2023). *A.I. and Desire - Deleuze & Guattari Preview*. (Visited on 11/14/2024).

Cser, Tamas (2024). *Understanding Tokens and Parameters in Model Training: A Deep Dive*. https://www.functionize.com/blog/understanding-tokens-and-parameters-in-model-training. (Visited on 06/21/2025).

Deleuze, Gilles (1995). *Negotiations, 1972-1990*. European Perspectives. New York: Columbia University Press. ISBN: 978-0-231-07580-0.

Deleuze, Gilles and Gilles Deleuze (1994). *Difference and Repetition*. New York: Columbia Univ. Press. ISBN: 978-0-231-08159-7 978-0-231-08158-0.

Deleuze, Gilles and Félix Guattari (1983). *Anti-Oedipus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press. ISBN: 978-0-8166-1225-3.

– (1987). *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press. ISBN: 978-0-8166-1401-1 978-0-8166-1402-8.

Deleuze, Gilles, Félix Guattari, and Gilles Deleuze (2008). *Kafka: Toward a Minor Literature*. 9. print. Theory and History of Literature 30. Minneapolis: Univ. of Minnesota Pr. ISBN: 978-0-8166-1514-8 978-0-8166-1515-5.

Dishon, Gideon (Sept. 2024). "From Monsters to Mazes: Sociotechnical Imaginaries of AI Between Frankenstein and Kafka". In: *Postdigital Science and Education* 6.3, pp. 962–977. ISSN: 2524-485X, 2524-4868. DOI: 10.1007/s42438-024-00482-4. (Visited on 11/19/2024).

Dreyfus, Hubert L. (2009). *What Computers Still Can't Do: A Critique of Artificial Reason*. Rev. ed., [repr.] Cambridge, Mass.: MIT Press. ISBN: 978-0-262-54067-4 978-0-262-04134-8.

Eloff, Aragorn (May 2021). "2006: The Topology of Morals (Who Does the Algorithm Think We Are?)" In: *Deleuze and Guattari Studies* 15.2, pp. 178–196. ISSN: 2398-9777, 2398-9785. DOI: 10.3366/dlgs.2021.0435. (Visited on 05/27/2025).

Foucault, Michel (2008). *The Birth of Biopolitics: Lectures at the Collège de France, 1978-79*. Ed. by Michel Senellart. Basingstoke [England] ; New York: Palgrave Macmillan. ISBN: 978-1-4039-8654-2.

– (2012). *The Archaeology of Knowledge*. Westminster: Knopf Doubleday Publishing Group. ISBN: 978-0-394-71106-5 978-0-307-81925-3.

Foucault, Michel and Michel Foucault (2009). *Security, Territory, Population: Lectures at the Collège de France, 1977-78*. Ed. by Michel Senellart, François Ewald, and Alessandro Fontana. Trans. by Graham Burchell. Basingstoke New York: Palgrave Macmillan. ISBN: 978-0-230-24507-5 978-1-283-17917-1. DOI: 10.1057/978-0-230-24507-5.

Galloway, Alex (Sept. 2001). "Protocol, or, How Control Exists after Decentralization". In: *Rethinking Marxism* 13.3-4, pp. 81–88. ISSN: 0893-5696, 1475-8059. DOI: 10.1080/089356901101241758. (Visited on 09/18/2024).

Galloway, Alexander R. (2004). *Protocol: How Control Exists after Decentralization*. Leonardo. Cambridge, Mass: MIT Press. ISBN: 978-0-262-07247-2.

Gillespie, Tarleton (June 2024). "Generative AI and the Politics of Visibility". In: *Big Data & Society* 11.2, p. 20539517241252131. ISSN: 2053-9517, 2053-9517. DOI: 10.1177/20539517241252131. (Visited on 11/03/2024).

Gretzky, M. (Dec. 2024). *The Rise of the Algorithmic Author? A Critical Analysis of Large Language Models in Higher Education*. https://www.digitalcultureandeducation.com/volume-152-papers/shtoltz. (Visited on 06/19/2025).

Haggerty, Kevin D and Richard V Ericson (2000). "The Surveillant Assemblage". In: *The British journal of sociology* 51.4, pp. 605–622. ISSN: 0007-1315.

Hecht-Nielsen, Robert (1992). "Theory of the Backpropagation Neural Network**Based on "Nonindent" by Robert Hecht-Nielsen, Which Appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE." In: *Neural Networks for Perception*. Elsevier, pp. 65–93. ISBN: 978-0-12-741252-8. DOI: 10.1016/B978-0-12-741252-8.50010-8. (Visited on 06/19/2025).

Hui, Yuk (Oct. 2015). "Modulation after Control". In: *New Formations* 84.84, pp. 74–91. ISSN: 0950-2378. DOI: 10.3898/NewF:84/85.04.2015. (Visited on 11/03/2024).

ai-inquiry (2025a). *AI Meets Philosophy, Vol. 4: Deep Learning Processes Through the Lens of Deleuze's Philosophy | by AI Inquiry Garden - Freedium*. https://freedium.cfd/https://medium.com/@AI_Inquiry_Garden/rhizomatic-learning-deep-learning-processes-through-the-lens-of-deleuzes-philosophy-4a6b1b13d1c6. (Visited on 05/19/2025).

– (Mar. 2025b). *AI Meets Philosophy, Vol.7-Part2/2: AI Internal Structure through Deleuze's Molecular/Molar Concept*. (Visited on 05/19/2025).

Kivisto, Peter (2013). *Illuminating Social Life: Classical and Contemporary Theory Revisited*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. ISBN: 978-1-4522-1782-6 978-1-5063-3548-3. DOI: 10.4135/9781506335483. (Visited on 05/09/2025).

Konik, Adrian (Jan. 2015). "The Politics of Time: Deleuze, *Duration* and Alter-Globalisation". In: *South African Journal of Philosophy* 34.1, pp. 107–127. ISSN: 0258-0136, 2073-4867. DOI: 10.1080/02580136.2014.992157. (Visited on 11/03/2024).

Kordzadeh, Nima and Maryam Ghasemaghaei (May 2022). "Algorithmic Bias: Review, Synthesis, and Future Research Directions". In: *European Journal of Information Systems* 31.3, pp. 388–409. ISSN: 0960-085X, 1476-9344. DOI: 10.1080/0960085X.2021.1927212. (Visited on 06/22/2025).

Krasmann, Susanne (2017). "Imagining Foucault. On the Digital Subject and "Visual Citizenship"". In: *Foucault Studies*, pp. 10–26. ISSN: 1832-5203.

Kruger, Jaco (Apr. 2021). "Larval Intelligence: Approaching AI in Terms of Deleuze's "System of the Dissolved Self"". In: *South*

African Journal of Philosophy 40.2, pp. 171–181. ISSN: 0258-0136, 2073-4867. DOI: 10.1080/02580136.2021.1933724. (Visited on 11/04/2024).

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (May 2015). "Deep Learning". In: *Nature* 521.7553, pp. 436–444. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539. (Visited on 05/08/2025).

Lex Fridman (Mar. 2024). *Yann Lecun: Meta AI, Open Source, Limits of LLMs, AGI & the Future of AI | Lex Fridman Podcast #416*. (Visited on 05/28/2025).

Maas, Wouter (2023). "Deconstructing Transformers". In.

MacKenzie, Iain and Robert Porter (June 2021). "Totalizing Institutions, Critique and Resistance". In: *Contemporary Political Theory* 20.2, pp. 233–249. ISSN: 1470-8914, 1476-9336. DOI: 10.1057/s41296-019-00336-w. (Visited on 10/08/2024).

Melanie (Mar. 2024). *Kernel: Everything You Need to Know about the Machine Learning Method*. (Visited on 06/21/2025).

Mishra, Punya and Marie K Heath (2024). "The (Neil) Postman Always Rings Twice: 5 Questions on AI and Education". In.

Montanari, Federico (Jan. 2025). "ChatGPT and the Others: Artificial Intelligence, Social Actors, and Political Communication. A Tentative Sociosemiotic Glance". In: *Semiotica* 2025.262, pp. 189–212. ISSN: 0037-1998, 1613-3692. DOI: 10.1515/sem-2024-0210. (Visited on 05/27/2025).

Nebius-Team (July 2024). *What Is Epoch in Machine Learning? Understanding Its Role and Importance*. https://nebius.com/blog/posts/epoch-in-machine-learning. (Visited on 06/19/2025).

Pasquinelli, Matteo (2023). *The Eye of the Master: A Social History of Artificial Intelligence*. London New York: Verso. ISBN: 978-1-78873-006-8 978-1-78873-008-2 978-1-78873-007-5.

PhD, Michael Filimowicz (Feb. 2025). *LLMs through Saussurean and Peircean Lenses*. (Visited on 04/28/2025).

Poster, Mark, David Savat, and Gilles Deleuze, eds. (2010). *Deleuze and New Technology*. reprinted. Deleuze Connections. Edinburgh: Edinburgh Univ. Press. ISBN: 978-0-7486-3336-4 978-0-7486-3338-8.

Prinsloo, Paul (May 2017). "Fleeing from Frankenstein's Monster and Meeting Kafka on the Way: Algorithmic Decision-Making in Higher Education". In: *E-Learning and Digital Media* 14.3, pp. 138–163. ISSN: 2042-7530, 2042-7530. DOI: 10.1177/2042753017731355. (Visited on 01/08/2025).

Rijos, Avery (n.d.). "Posthumanist Phenomenology and Artificial Intelligence (4th Edition)". In: *Medium*.

Rouvroy, Antoinette (2007). *Human Genes and Neoliberal Governance: A Foucauldian Critique*. Routledge-Cavendish. ISBN: 1-134-06668-6.

– (n.d.). "The End(s) of Critique : Data-Behaviourism vs. Due-Process." In: ().

Tarmoun, Salma et al. (Oct. 2024). "Gradient Descent and Attention Models: Challenges Posed by the Softmax Function". In: (visited on 06/18/2025).

Van Otterlo, Martijn (Jan. 2013). "A Machine Learning View on Profiling". In: *Privacy, Due Process and the Computational Turn*. Ed. by Mireille Hildebrandt and Katja De Vries. London: Routledge. doi: 10.4324/9780203427644.

Vaswani, Ashish et al. (2017). "Attention Is All You Need". In: doi: 10.48550/ARXIV.1706.03762. (Visited on 12/14/2023).

Zarkadakēs, Giōrgos and Don Tapscott (2020). *Cyber Republic: Reinventing Democracy in the Age of Intelligent Machines*. Cambridge, Massachussets London, England: The MIT Press. isbn: 978-0-262-54272-2 978-0-262-04431-8.

Zuboff, Shoshana (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile books. isbn: 978-1-78125-684-8 978-1-78125-685-5.