Preparatory Space: Roland Barthes and Large Language Models

Sunil Manghani

In beginning this lecture, I should first declare some context: I have prepared these notes as part of composing a lecture course on the 'preparation of the artwork'.[1] The project is inspired by Roland Barthes' late lectures (held at the Collège de France), notably *The Preparation of the Novel*. I'm mindful of Barthes' opening remarks to his course (which also draws back to his 'Inaugural Lecture'[2]), whereby he notes his sincere belief 'that at the origin of teaching […] we must always locate a fantasy'.[3] Put more directly, I take this to mean an unwavering need to 'think aloud', to be ambitious; to be at liberty to consider new ideas – even perhaps to be wrong, albeit productively.

Equally, however, beyond my interest in form, Barthes' final lecture course provides the underlying thematic: *preparation*. I intend to open upon the prospect of a 'preparatory space', whereby language and its effects can be thought of as *finite*, and so calculable – *already prepared*. It will mean reading somewhat against the grain of Barthes' own writings, but nonetheless, I remain sympathetic to his overarching notion of the Text. A notable point of reference is Barthes' enigmatic essay, 'From Work to Text' (→ it is worth noting the essay's appeal to science and mathematics; the reference to Einsteinian science, 'which compels us to include within the object studied the *relativity of reference points*').[4] I am not going to cite a great deal from this essay. It is less a direct source than a 'situation': a situation Barthes remarks upon in the opening line: 'A change has lately occurred, or is occurring, in our idea of language'.[5]

Today, I want to speak similarly of a 'change', which again relates to language, not least the very 'preparation' of language. I am speaking of our approach to writing in the context of Artificial Intelligence, and specifically Large Language Models (e.g., OpenAI's GPT series, Google's PaLM and Gemini, xAI's Grok, Meta's LLaMA, Anthropic's Claude models, and the Mistral AI's open source models etc.). An alternative title for this lecture might have been: *What would Barthes have thought of Large Language Models?*

Barthes Studies, 10 (2024), 164–98.
ISSN: 2058–3680.

In/Computable

'From Work to Text' is a well-known and somewhat programmatic text. I have in mind a different, vicarious reading. It is my contention that an earlier colloquium text (i.e., a *spoken* text), 'Semiology and Urbanism', acts as an informative *preparatory* text.[6] But first, to set the scene: At the close of 'From Work to Text', Barthes declares the Text (written with a capital 'T') to be a 'social space which leaves no language safe […] the theory of the Text can coincide only with a practice of writing'[7]. In the opening of the essay, as I've already noted, he makes the analogy of a shift from a 'Newtonian physics' of the sign to an Einsteinian, relational one. Rather than see this a clarion call of post-structuralism, as if sweeping aside the claims of structuralism, I will read it (if contrarily) as the spreading of structuralism's wings with the falling of dusk. What was not available at the time Barthes was writing was the kind of computational power and techniques that we are now beginning to see, and indeed use in ubiquitous ways. Barthes' opening proposition of the Text is held within the limitations of the period → 'The text must not be understood as a computable object. It would be futile to attempt a material separation of works from texts.'[8]

According to Barthes, we should not state the work as one thing and the Text as another; rather 'the work is a fragment of substance, it occupies a portion of the spaces of books (for example, in a library). The Text is a methodological field'.[9] The Work contains the effects of the Text, and Text makes its way through countless Works. Yet, *what if* the Text were in fact computable? What if the Text were not just an operation (of calculation) – not just a 'methodological field' – but were itself subject to calculation? The remark that the Text 'must not' be understood a computable object can be read with a certain romanticism, contra science, that *something* ought to remain outstanding. Yet, read another way, Barthes' concern is narrower: he rightly does not want to conflate, nor place in a simple binary, Work and Text. In keeping with an Einsteinian physics, the Work is to the Text as the particle is to the wave[10] → so, to suggest they are *not* separate entities, rather they are different incidences, different utterances of the same. The problem for Barthes, at the time, was a simple brute reality: an inability to make the kinds of calculations we can make today. Yet, nonetheless, philosophically speaking, structuralism might at least *begin* to conceive of a different order of scale. (cf. Lévi-Strauss on the structural analysis of myth of the prospect of three-

dimensional models, which would require 'a spacious workshop … [and] IBM equipment, etc.'[11])

I intend to hold a little longer to the 'early', structuralist Barthes[12] – or at least attend to the interstice, the period in which Barthes was working between a structuralist and post-structuralist frame. 'From Work to Text' was published in 1971, while 'Semiology and Urbanism' was presented in 1967. Over this relatively short period the shift towards the (post-structural) incomputable starts to take shape. Subsequently, it is fair to say most contemporary commentary on Barthes focuses on the 'late', *incomputable* Barthes → the numerous pieces on *Camera Lucida*; the late lecture courses; and more generally the *writerly*. (In other directions the in/computable is held up as a political struggle for the human subject[13]).

Foucault's *The Order of Things* sets out the longer view, whereby, for example, Saussurean semiology is understood to have re-discovered the break from the 'resemblance' of the Renaissance to the rational, linguistic, cataloguing of the Classical era.[14] Foucault's relationship to structuralism (vis-à-vis hermeneutics) is complex and nuanced,[15] but, nonetheless, his remark that '[s]tructuralism is not a new method' but 'the awakened and troubled consciousness of modern thought' is apt.[16] The rise of new technologies, such as AI, machine learning and computer vision, and their underpinning mathematics, are not mere new inventions, but represent a culmination of a long history of categorising, classifying, and calculating. The explorations and many false starts in AI[17] continually lead us to question just how we think, which arguably brings us back to the project of structuralism as not merely the *analysis* of culture, but the exploration of the *structures of thought*. My contention, then, is that turning away from structuralism (in favour of post-structuralism) is not sufficient. We may choose to look away, but it does not take away the deeper resonance of what is structural.

Notes on Structuralism

For an edited collection, 'Notes on Structuralism',[18] I was fortunate to include a previously unpublished interview: a dialogue between Roland Barthes and Paolo Fabbri, which took place on 18 December 1965 in Florence, Italy.[19] Barthes offers an engaging account of his structuralist approach to narrative, as was published only months later in the well-known essay, 'Introduction to the Structural Analysis of Narrative' (originally published in French in 1966, as part of a special issue of *Communications*).[20] Note again: in the context of the earlier structuralist

phase, we have here the slippage between a spoken and written text (in this case between an interview and a published essay on narrative). I refer to the interview and essay as two distinct 'texts'.

The mid-1960s marked the apotheosis of Barthes' structuralist phase. *The Fashion System,*[21] described as a laborious study, or even cruelly as '*the* most boring book ever written about fashion',[22] had been completed in 1963 (although not published until 1967). Barthes sets out different structures from within the 'general system' of fashion. In 'Introduction to the Structural Analysis of Narratives', he takes on the more ambitious subject matter of the 'narratives of the world', which in the opening line he declares as 'numberless' (I shall return to this remark).[23] The *method* remains the same: to identify distinct 'levels of meaning', which, analogous to the specific levels of units of meaning identified by structural linguistics (i.e. phonetic, phonological, grammatical, contextual), aim to determine *invariance* across all narratives, all signifying systems.[24]

In Jonathan Culler's contribution to 'Notes on Structuralism', he carefully examines how the *interview* with Barthes diverges from the *essay*, and also with regards to Barthes' 'most detailed analysis of narrative, *S/Z*, in 1970'.[25] The latter text can be said to mark the more formal shift to post-structuralism, along with *The Pleasure of the Text*[26] (originally published in 1973); as well as the two well-known essays 'The Death of the Author', [27] published in English in 1967 (in *Aspen*), and the aforementioned 'From Work to Text', published in 1971.

A noticeable difference in the interview, in contrast to the published essay, is Barthes' repeated allusion to the *anthropological* aim of analysis. Given the occasion of the talk, with an august group of scholars in attendance, he appears to adopt this term as means to uphold a collective, scientific field of enquiry. As Culler suggests, in the interview 'Barthes explicitly presents himself as working to advance a collective analytical project and speaks with anticipation of progress he hopes will follow.'[28] The idea of a collective project, the fact it is based in empirical work, and which is concerned with the 'logic of human actions', is a particular feature of the interview, which falls away, partially, with the essay, and certainly in his subsequent writings.

In the published essay, the references to Saussure are brief, while during the interview Barthes 'follows more explicitly the steps of Saussure'.[29]In particular, Barthes refers to the 'commutation test' common in linguistics for identifying distinct units of meaning (which involves methodically testing replacements of phonemes or words to see if they significantly alter meaning). The problem at the level of narrative is that

large components (including, for example, whole chapters in a novel) can often be replaced or removed without altering the fundamental basis of the story. It is likely for this reason, Culler argues, that Barthes moves more swiftly in the published essay from his reference to Saussure to discussing '*functionality* as the criterion for the identification of narrative units', and which he also establishes as operating at different levels.[30]

The principle of *commutation* is emphasised in *The Fashion System* (with the 'structure' of clothing, the cross-matching of signs and forms, allowing for highly varied and complex constructions, yet which nonetheless remain all part of a system). In the interview on narrative, in thinking about the apparently numberless stories of the world, Barthes is still holding onto the idea of commutation as a means of analysis, but which seems to get thwarted by the sheer scale of the calculation. Yet, the references to 'numberless narratives', 'millions of narratives' and the 'infinite number of narratives' in the published essay, while echoed in the interview, are somewhat caveated, just slightly: 'we encounter a problem', he suggests, 'in trying to impose some type of scientific order on a body of material which appears, *at first sight*, to be impossible to master'; and similarly he says '*we believe* that, faced with an infinite number of narratives, we can start from a working hypothesis much like that of linguistics'.[31] In speech, in the interview (again keeping in mind the audience he is speaking to), Barthes retains a greater sense of optimism for the collective (pragmatic) project underway:

> As things stand at present, what we can develop is only a *theory* […] a theory of narrative just means that we're trying to figure out a hypothetical descriptive model. And this really is a necessary task, because, faced with this infinite number of narratives, if we're going to start seeing clearly, and find our way through, we need a hypothetical descriptive model.[32]

Today, we might suggest Barthes' hopes for a 'hypothetical' model have been realised beyond expectations with respect to recent developments in Large Language Models.[33] It is worth saying, the commutation test has been fundamental to the development of Large Language Models in several ways. The test helps in understanding the syntactic flexibility and constraints within a language. During the pre-training phase, Large Language Models are exposed to vast corpora of text, allowing them to implicitly learn the patterns of substitution and commutation within the language. Subsequently, the ability to substitute words or phrases while retaining the sentence's meaning can help Large Language Models

appropriately handle semantic roles and relationships; crucial for tasks like question answering, text summarization, and translation, where grasping the nuances of meaning is essential.

By learning which elements can be commuted without altering the sentence's core meaning, Large Language Models can paraphrase or creatively vary sentence structure in tasks like text generation. Outputs are made more diverse and natural → generative. Of course, the ability to appropriately substitute divides opinion, some feeling that tools like ChatGPT are simply parroting and plagiarising,[34] with others, taking a Wittgensteinian view of language, argue large language models show signs of conceptual reasoning, whereby conceptual meanings, while not derived from direct references, emerge through *internal reasoning*, due to the way concepts in language 'relate to *each other*'.[35]

More than a 'descriptive' model, the advances in AI and deep learning yield powerful probabilistic models far beyond that imagined at the time Barthes was working. The ability of machines to quickly parse through massive datasets of natural language usage (including millions of stories) is quite staggering. Recent natural language processing models, for example, are trained on data in excess of 300 terabytes of text, which would take 3 million lifetimes for a human to read.[36] When put into perspective of the lone researcher, as with Barthes working through a few James Bond novels (while nonetheless prospecting for a whole *theory* of narrative), or even with Lévi-Strauss reading hundreds of myths, we begin to realise how high-performance computing provides far greater potential for *structural analysis* than was ever envisaged. Commutation is now mere child's play in comparison to the multi-dimensional mathematics that underpin the deep learning of virtual neural networks, yet there is arguably an underlying principle at stake whereby the arbitrary nature of the sign (of structural linguistics) is as much a requirement for semiology as it is for the methods of predictive text and computer vision.

A few remarks on 'space': Underlying all AI computational work are concepts of mathematical *space* → vector spaces, metric spaces, topological spaces. Data points (like words, images, or user preferences) are represented as vectors in a high-dimensional space. In practice, what this means is that language is 'tokenised' into a vast array of unique strings of numbers, which can be used to determine patterns and relational positions. Natural Language Processing, for example, converts words or phrases into word embeddings in a vector space, whereby the geometric relationships between the vectors capture semantic relationships between the words (a complex form of commutation). Metric spaces define the distance between

any two points, critical for many AI algorithms that rely on measuring the similarity or dissimilarity between data points. Topological spaces are more abstract and deal with properties preserved under continuous transformations. They are less commonly mentioned in basic AI applications but find their place in advanced research. Crucially, a Hilbert space enables the generalization of Euclidean geometry to *infinite* dimensions – such a 'space' enables kernel methods to take original input data and transforms it into a higher-dimensional space. (Imagine a room in which you had infinite places to store all your belongings and, furthermore, the act of storing your possessions enables you to secure patterns of meaning and proximities).

As a visual metaphor, we might consider higher-dimensional space akin to the Magic Eye (or autostereogram) pictures made popular in the 1990s (except rather than just three-dimensions they are many dimensions!). The Magic Eye pictures are *two*-dimensional images that create the optical illusion of *three* dimensions. On first look, the 3D scene is usually unrecognizable until it is viewed with the correct vergence (the optical illusion of an autostereogram is one of depth perception and involves stereopsis: depth perception arising from the different perspective each eye has of a three-dimensional scene, called binocular parallax). The 'Ahha!' moment when the pictures comes into view provides the analogy of what happens when data placed in higher dimensional space is brought into alignment as a *pattern* through machine learning. In essence, I want to suggest this computational 'space' is a *preparatory space*, which structuralist analysis was already hinting at, even if it remained out of view.

Signifying Space

With these preliminary remarks, let me now turn to Barthes' 'Semiology and Urbanism', which evocatively opens with the idea that all 'human space … has always been a signifying space'.[37] The text performs at least two things: it again marks an interstitial space, theoretically (flirting with both a science and a writing of the sign); but also the topic of the 'city' provides a *spatial metaphor* that adds to our understanding of the complexities of high-dimensional mathematics, critical to the development of Large Language Models → critical to an understanding of contemporary preparations of/with language.

Barthes declares the city to be a discourse; 'and this discourse is actually a language: the city speaks to its inhabitants, we speak our city, the city where we are, simply by inhabiting it, by traversing it, by looking at it'.[38]

Consider the words of curator Iwona Blazwick, in her introduction to the catalogue of *Century City*, the inaugural exhibition held at Tate Modern in 2001:

> Who hasn't felt a thrill run up their spine on looking out of a plane at night and seeing the electric geometries of the nocturnal city? Lines of red and white light pulse in and out of a spatial web, sodium orange and fluorescent blue at its industrialised edges, points of rainbow-coloured neo through the centre. The city seems to stretch across the dark land mass: inhaling traffic, resources, people; exhaling refuse, spectacle, ideology and change. The city is the medium for the modern.[39]

What the exhibition epitomised was the city as both a represented space and a space of representations, fostering all kinds of significations and cultural expression. Blazwick's opening line ('Who hasn't felt a thrill run up their spine') is emblematic of a poetic call and response of the city. It is no surprise she writes descriptively ('Lines of red and white light pulse in and out of a spatial web'). The city, as she puts it, is 'the medium for the modern'. Of course, at the time of this exhibition, at the turn of the millennium (in the early days of the World Wide Web), it was more the medium for the postmodern.

The internet was fast becoming awash with hypertexts, websites and blogs; millions upon millions. Mostly these sites were of little shared significance, the mere musings of amateur writers, but gradually, as a mass accumulation of the Text, further fuelled with the emergence of social media, the online 'space' became critical to our everyday lives and experiences. In a Baudrillardian sense, we were more real than reality – *through* our mediated 'texts'. The internet allowed for a whole new combinatory, performative space, in the mode of Raymond Queneau's *Cent mille milliards de poèmes* [40] (to which I will return). → 'writing' not just with words, but with pictures, music and videos (to embed a YouTube video was to become no different from including a text quotation); we cut 'n' paste with freedom, we 'feed' in content from elsewhere and generally weave and re-weave a web of hypertext. When Barthes wrote 'the metaphor of the Text is that of the *network*', it was just that, a metaphor.[41] But, the poststructuralist gesture is now a lived (virtual) reality.

A notable reference in 'Semiology and Urbanism' is Victor Hugo's *Notre-Dame de Paris*.[42] Particularly the chapter, 'This Will Kill That' – a didactic aside by the author, which follows the exchange between the Archdeacon of Notre Dame and his interlocutor, the Tourangeau, in

which they debate the passing of architectural (material) knowledge to the realm of books (to virtual knowledge forms).[43] Hugo expounds upon the changing modes of human thought and communication, marking, in this case, a transition from architecture to printing (from writing in stone, to writing on paper); from the grand Gothic cathedrals (which the Archdeacon symbolizes) to the rise of print and the dissemination of ideas through books, signalled by Tourangeau's curiosity and questioning. While presented as a binary logic (architecture/printing), Hugo's overall conception is more sophisticated → a palimpsestic account of knowledge: 'Architecture began like a writing system […] they produced books'.[44] Thus, Hugo reminds us of a deeper philosophical contention regarding the nature of expression and the materiality of thought itself, resonating with Derrida's critique of logocentrism, the critique that meaning precedes its articulation in language.[45] The idea to hold onto here is that 'writing' (be it architectural; urban; print-based) inherently transcends context. We are to posit a fluid continuum of thought, constantly being rewritten, wherein we conceive of one *big* language model (the Text).

Understood this way, the Text is plural, not just in that it allows for several meanings, but, as Barthes suggests, 'it fulfills the very plurality of meaning: an *irreducible* (and not just acceptable) plurality. The Text is not coexistence of meaning, but passage, traversal; hence, it depends not on an interpretation, however liberal, but on an explosion, on dissemination'.[46] The traversal of the Text comes into its own with the advent of contemporary AI Large Language models, which are fundamentally based upon statistical and probabilistic computations, focusing on identifying patterns. The challenge of speech recognition serves as a good example. It was only with the adoption of statistical methods, necessitating vast quantities of speech and text data for training, that any substantial advancements were made. Ultimately speech has been stripped down to *mere* data and analyzed without any linguistic insights, rendering the actual content of speech irrelevant.[47] Yet, in turn, this process enables meaningful exchange between humans and machines, and underlines significant advances in generative text. It is not only that recent advances *represent* the Text (as dissemination, as data). We might equally say, it is the *desire of the Text* (our compulsion to make meaning; to produce our 'data') that gives rise to AI in the first place. AI = Text, and vice versa.

Iwona Blazwick's remark, 'Who hasn't felt a thrill run up their spine on looking out…', captures, then, not just the wonder of the city, but what turned out to be a *desire* to re-present the world around us –

again and again (accelerated with the advent of the camera on the mobile phone). As Nicholas Mirzoeff reminds us: the last European 'who was thought to have read all available printed books was the sixteenth-century reformer Erasmus'.[48] Today, the explosion of text and image production is fast outpacing whole centuries' worth of 'data'; '[e]very two minutes, Americans alone take more photographs than were made in the entire nineteenth century'.[49] One wonders what the likes of Walter Benjamin, Henri Lefebvre and Michel de Certeau would have made of today's *flâneurs*; no longer a minority of detached observers but more like a majority, capturing and circulating the world around them through the lens of the camera phones and the networks of social media.[50] Fitting with Benjamin's analysis the potential is realized for 'the public' to be the 'examiner' – for all of us to become the resident critics and experts. Of course, as we know from the 'Artwork' essay, it comes with the admonition that the examiner is generally 'an absent-minded one'.[51] Yet, it is not necessarily *what* these new 'writings' say in themselves that is interesting or revealing. Instead, held up as a surface expression or 'Mass Ornament', the profusion of the (digital) Text provides a 'collective unconscious'.[52] And, it is this reservoir, and importantly its massive scale, that provides the basis for Large Language Models.

Of course, is also the *longue durée* of Text that is pertinent. Returning to the city as 'signifying space', Barthes looks as far back as the cartography of Greek antiquity → 'a veritable discourse'. A map of the world by Herodotus, for example, 'is constructed like a language, like a sentence, like a poem, on oppositions: hot countries and cold countries, known and unknown countries'. Similarly, the notion of *Isonomy* (the Greek concept of equality) is described as 'a truly structural conception', which in turn he equates with the design of the city in the same period: 'the conception of the city was exclusively a signifying one, for the utilitarian conception of an urban distribution based on functions and usages, which incontestably prevails in our day'.[53]

Coming forward in time, Barthes also attributes the idea of signifying space to Lévi-Strauss, in *Tristes Tropiques*, wherein he describes the occasion in which he rises at dawn to go out and examine the village in which he is staying. He sketches the arrangement of the inhabitants' huts, which circle around a central and large 'men's house', prohibited to women (Figure 1). He offers this structural metaphor: 'Seen from a treetop or a roof, a Bororo village looks like a cartwheel, the rim being the family huts, the spokes the paths and the men's house the hub'. The arrangement, he argues, was common across the region and 'is so important a factor in

[…] social and religious life that […] the surest way to convert the Bororo was to make them abandon their village in favour of one with the houses set out in parallel rows.[54]

At a formal level, the cartwheel diagram is suggestive of a vector space. In this case, it defines particular relations between men and women in the society. And more than that, this vector space is a form of *cultural data*, which was exploited by missionaries:

> Once [the inhabitant] has been deprived of their bearings and were without the plan which acted as confirmation of their native lore, the Indians soon lost any feeling for tradition; it was as if their social and religious systems […] were too complex to exist without the pattern which was embodied in the plan of the village and of which their awareness was constantly being refreshed by their everyday activities.[55]

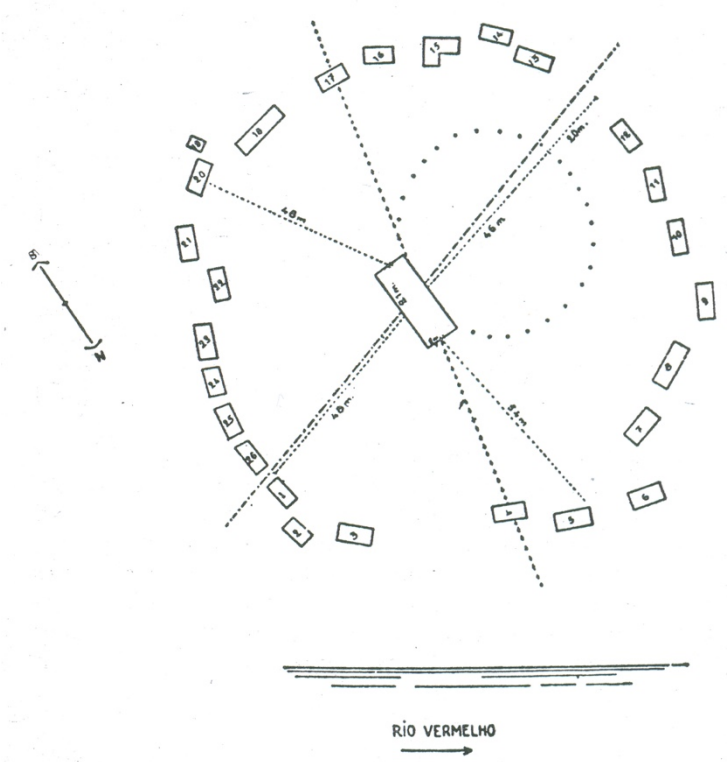

Figure 1: Plan of Kejara Village, from Lévi-Strauss' *Tristes Tropiques.*

We might reverse-engineer Lévi-Strauss' words, to suggest how contemporary data methods can take the complexity (and fecundity) of our everyday activities to find otherwise undeclared patterns and

'traditions'. Even a small multi-dimensional space (Figure 2) provides the freedom to group semantically similar items and keep dissimilar items far apart. Position (distance and direction) in the vector space encode semantics in the form of embeddings (e.g., to show geometrical relationships that capture semantic relations such as the relation between a country and its capital).

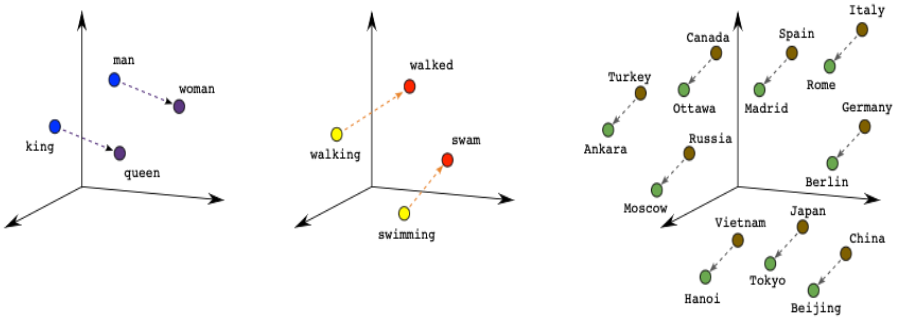

Figure 2: Visualizations of embeddings illustrate 'geometrical' (spatial) relationships between the words for a country and its capital. E.g., the 'distance' from words 'Canada' and 'Ottawa' is about the same as the distance from 'Turkey' and 'Ankara'. I.e., as cities of the respective countries, there is greater propensity to find these words being used together (or, put another way, the writing about Turkey and Ottawa in a single sentence represents a lower probability). Source: Google for Developers.[56]

Today, we extend the idea of data methods leveraging cultural patterns and traditions on a much grander scale. In *The Age of Surveillance Capitalism*,[57] Shoshana Zuboff highlights how tech giants like Google amass vast quantities of data on individuals' behaviours, preferences, and interactions. This data is not just a record of activity; it's a rich seam of insight into cultural norms, values, and shifts. With respect to 'signifying space', it is pertinent that a pivotal moment in Google's history is the observation of spikes in searches that correlated with a TV game show as it aired across the different time zones of America. This suggested a form of 'time travel', an ability to predict behaviour, which in turn led to a way of monetising web search. Zuboff refers to this as 'behavioural surplus' → the timing and context of searches can be analysed, commodified, and used for purposes beyond the initial intent. This is the dark side of a 'preparatory space'. For now, however, allow me to remain within the confines of a formal critique → to secure an understanding of what

structuralism might still offer in relating to our *situation* in the 'change in language'; a change in how we might consider the 'preparation' of language, of *what we (already) say*.

An Imaginary Science

Let us now read on a little further with 'Semiology and Urbanism' (albeit against the grain). Barthes suggests an urban semiotics 'would consist in dissociating the urban text into units, then in distributing these units into formal classes, and thirdly, in finding the rules of combination and of transformation for these units and for these models'.[58] Yet, he pulls back from delivering any such system of analysis. As Diana Knight suggests, 'the lecture is in fact an apology for Barthes' own sort of semiology'.[59] Drawing out Barthes' own words, she explains:

> the 'scientific' approach ('investigations or functional studies of the city') is gradually displaced in favour of a plea for 'a certain ingenuity', an accumulation of personal readings of the city, 'of which, unfortunately, till now, only writers have given us some examples' […] No one, of course, could be fooled by Barthes' 'unfortunately'. By the end of a short lecture he has moved from an opening statement of his modest credentials as speaker on the topic – both lover (*amateur*) of signs and lover of the city – to an unambiguous statement of his own creative investment in writing the city.[60]

Barthes acknowledges that the 'usable data' of the social sciences had up until that time not been adequately incorporated. He writes:

> if we have difficulty inserting into a model the urban data supplied us by psychology, sociology, geography, demography, this is precisely because we lack a final technique, that of symbols. Consequently we need a new scientific energy in order to transform such data, to shift from metaphor to the description of signification, and it is here that semiology (in the broadest sense of the word) may by a still unpredictable development afford us some assistance.[61]

The appeal to semiotics – or this 'final technique' of symbols – as a foundational approach and as part of some 'new scientific energy' would seem initially to locate with a structuralist domain (even to relate to contemporary Large Language Models). But, as noted by Knight, Barthes

thwarts such an account (→ it is in preparation but put in abeyance). He nonchalantly announces: 'It is not my intention to evoke here the procedures for discovering an urban semiology.'[62] (NB. The paper is titled 'Semiology *and* Urbanism', not 'Urban Semiology'.) Thus, while Barthes takes steps towards outlining an urban semiotics ('It is likely that such procedures would consist in dissociating the urban text into units'[63]), in the end, in keeping with an emergent post-structuralist perspective, he puts his faith in a textual notion (and experience) of the city, which becomes a very different object of/for analysis. The most important thing, he suggests, 'is not so much to multiply investigations or functional studies of the city as to multiply the readings of the city'.[64] → We might begin to suggest connections with recent developments in generative AI, which by default multiply readings/renderings.

NB. Seductive nature of Barthes' writing → high degree of ideas 'formulated through an explicit vocabulary of utopia'.[65] As Barthes explains: 'I have a utopian imagination and very often when I write, even if I am not referring to a utopia, if, for example, I'm analysing particular notions in a critical way, I always do this through the inner image of a utopia: a social utopia or an affective utopia.'[66] Unsurprisingly, in 'Semiology and Urbanism', methodology remains an open (utopian?) question:

> I have not approached the problem of methodology. Why? Because, if we seek to undertake a semiology of the city, the best approach […] will be a certain ingenuity on the reader's part. It will require many of us to attempt to decipher the city […] beginning, if necessary, with a personal report.[67]

However, in the context of today's Big Data (whereby individuals on mass are generating calculations, predictions, texts, locations, transactions, likes, dislikes), we can begin to see how all of these 'personal reports' become something more, beyond what Barthes might have ever imagined. In what follows, I will trace Barthes' three main observations on semiology and the city. While proposed in such a way as to lead us toward a post-structuralist position, I hold to the structuralist account (which I demonstrate through extension to a reading of AI). Typically, utopic imagination = the *writerly*; scientific utopia = structuralist → what if the 'utopia' of language = the previously *imagined* science of the sign, now realised in AI.

Let us now proceed through the three key terms designated in 'Semiology and Urbanism': symbol, signifier and the erotic.

(1) Symbol: Barthes begins by considering the shifting status of the symbol, which can no longer be considered as something fixed, or as being of a determined lexicon. '[I]t would be absurd', he writes, 'to attempt to elaborate a lexicon of the significations of the city'.[68] He draws attention, for example, to the notion of an 'empty signifier' – a site of non-meaning around which signification can accumulate, yet without any metaphysical foundation. He refers to the centre of Tokyo, with its forbidden, secret territory of the Imperial Palace, which, with its 'deep moat and hidden by verdure, is experienced as an empty center'.[69] Beneath an old map of Tokyo, illustrated in *Empire of Signs*, Barthes writes elliptically: 'The City is an ideogram; the Text continues.'[70] He wishes to mark a contrast with the 'very movement of Western metaphysics, for which every center is the site of truth, the center of our cities is always full.'[71]

In *formalist* terms (→ setting aside criticism of ethnocentrism), Tokyo provides Barthes with respite from the constraints of signification. He is literally unable to read off the language, which is entirely foreign to him (visually and aurally). Taking a 'holiday' from the strictures of systems of signification, he is able to imagine *other* possibilities, and like the typical holiday-maker he is imbued with a good deal of optimism. He does not wish to write *of* Japan, but rather to make something out of being there: 'The author has never, in any sense, photographed Japan. Rather, he has done the opposite: Japan has starred him with any number of "flashes"; or, better still, Japan has afforded him a situation of writing.'[72]

A further example gleaned of a 'touristic gaze' can be found in Lévi-Strauss' *Tristes Tropiques*. In amongst his various attempts to explain and divide up units of meanings, there are occasions that appear to go beyond *systems* of meaning:

> Every time I emerged from my hotel in Calcutta, which was besieged by cows and had vultures perched on its windowsills, I became the central figure in a ballet which would have seemed funny to me, had it not been so pathetic. The various accomplished performers made their entries in turn: a shoeblack flung himself at my feet; a small boy rushed up to me; whining 'One anna, papa, one anna!' a cripple displayed his stumps, having bared himself to give a better view; a pander – 'British girls, very nice…'; a clarinet-seller; a New Market porter begged me to buy everything…[73]

Emerging out of his hotel, Lévi-Strauss is more tourist than social scientist; the 'world' is infinitely stranger; more fluid and rhythmic (a 'ballet'). Of course, the scene is as much revealing of Lévi-Strauss (as subject) as it is of

the city and its inhabitants. But, for present purposes, what we can take from these detachments of meaning, or engulfing decipherments, is a searching question about the ordering of meaning. We might think of Large Language Models as acting like tourists, unable to 'read', yet capable of 'witnessing' order in a cacophony of words (although advantage is gained from the fact language is sequential).

As Mustafa Suleyman explains, AI models draw upon very large samples of text to 'abstract representation of the information contained within', so breaking language into 'tokens' of meaning, which include syntactical elements as well as words and letters. As previously noted, this process amounts to the creation of many strings of numbers, acting as coordinates or identifiers. 'The challenge lies in designing an algorithm that "knows where to look" for signals in a given sentence.'[74] This capability, referred to as 'attention', is attributed as the major breakthrough in 'transformer' models,[75] → the 'GPT' in ChatGPT refers to generative pre-trained *transformer*. Not so different to Lévi-Strauss in attending to the situation as he emerges from his hotel, we can understand:

> When a large language model ingests a sentence, it constructs what can be thought of as an 'attention map'. It first organizes commonly occurring groups of letters or punctuation into 'tokens' […] making it easier for the model to process the information. It's worth noting that humans do this with words of course, but the model doesn't use our vocabulary. Instead, it creates a new vocabulary of common tokens that helps it spot patterns across billions and billions of documents. In the attention map, every token bears some relationship to every token before it, and for a given input sentence the strength of this relationship describes something about the importance of that token in the sentence. In effect, the [large language model] learns which words to pay attention to.[76]

The AI model's ability to look dispassionately (starred with any number of flashes), combined with oversight of masses of training text, gives rise to *a new formulation of writing*. The empty signifiers, the seeming chaos that awaits the viewer stepping out onto the street from the hotel does in fact have an order (as indeed Lévi-Strauss' eloquent writing attests).

NB. The attention mechanism (and backpropagation → a fundamental algorithm in neural network training, capable of correcting, or *going back on* a calculation to avoid unnecessary compute time) takes the AI's 'preparation' of language beyond what is referred to as 'brute force' techniques (i.e., a method of exhaustive search through all possible

solutions, which is simply not feasible with the combinatory possibilities of words). → Barthes' initial interest in the principle of commutation lifts to a whole other order of dimensions.

(2) Signifiers: Barthes' second observation, which builds on the first, is that 'symbolism must be defined essentially as the world of signifiers, of correlations […] never imprisoned in a full signification'. For a semiology of the city, 'we must intensify, more meticulously, the signifying division'.[77] He refers to a 'descriptive technique' as an analysis of the city in close and loving detail. Barthes appeals to his 'experience as an amateur of cities', arguing the need to look beyond the function of certain spaces within the city (whether markets or places of entertainment etc) and rather find a means to chart not only these sites, but equally one's movement through them:

> [H]ere we rediscover Victor Hugo's old intuition: the city is a writing; the man who moves about in the city, i.e., the city's user (which is what we all are, users of the city), is a sort of reader […] When we move about in a city, we are all in the situation of the reader of Raymond Queneau's *100,000 Million Poems*, where we can find a different poem by changing a single verse; unknown to us, we are something like that avant-garde reader when we are in a city.[78]

A member of the Oulipo collective, Queneau published *Cent mille milliards de poèmes* in 1961. It is comprised of a set of ten basic sonnets placed in a form referred to as 'combinatory literature': each of the sonnets have the same rhyme scheme and employ the same rhyme sounds. As a result, any line from a sonnet can be combined with any from the other nine, giving 10^{14} (= 100,000,000,000,000) different poems.[79] Working twenty-four hours a day, it would take some 140,000,000 years to read them all.[80] Barthes' point is that the city offers the same sort of fixed, *countable* structures, but which in their endless combinations allow for all number of different stories or viewpoints.

In a rather cleverly written article, 'The Potential Literature Horseshoe', Tom Savage emulates the combinatory method to produce a meta-critique that itself offers 2.66×10^{19} versions – each version of the text offering the same comprehensible, consistent argument (NB. the quotations that follow will invariably not be the same when the article is next view!).[81] He includes reference to Berge's diagram of *Cent mille milliards de poèmes* (Figure 3), suggesting '[v]erses act equivalently to

neural network layers, and phrases correspond to discrete nodes', only, 'what is missing is the mathematical transformations from layer to layer, token embeddings'. [82] In considering 'the thematic similarities of complexity, composition, and distillation (or lack thereof) of language' of both members of Oulipo and machine learning researchers, Savage poses a provocative account of the *form* of language and how we get to 'see' the underlying structure. Arguably, Large Language Models attend to language in as (if not more) complex ways as did the Oulipo group. A key difference, however, is that we foreground that complexity when we 'read' works of combinatory literature.

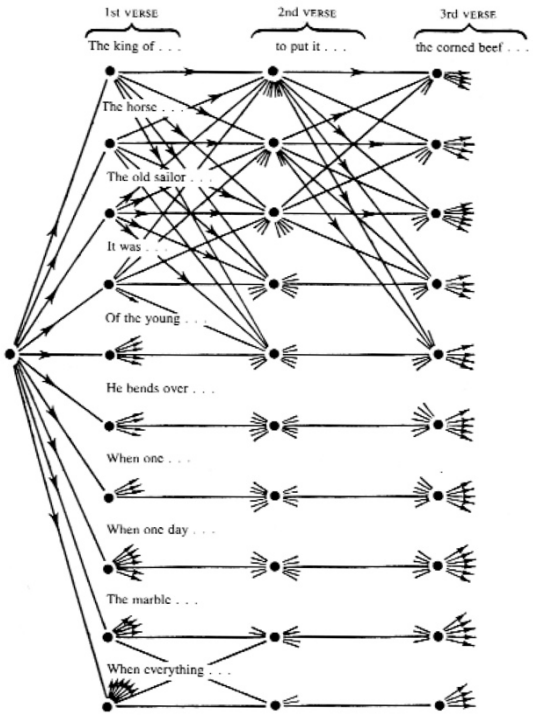

Figure 3: 'Principle of the graph of the *Cent Mille Milliards de poèmes* (not all of the arcs and vertices have been drawn)': Diagram from Claude Berge's 'For a Potential Analysis of Combinatory Literature'.[83]

By contrast, the output of Large Language Models *dissolves* to normative text. Placing a 'sense of order' and 'rule complexity' on a graph, Savage demonstrates this as a horseshoe effect (Figure 4): 'The sheer quantity of rules within [a Large Language Model] reflects a *removal* of linguistic constraints, order is returned from chaos, and the resulting

underlying form is obscured.'[84] The Large Language Model's 'space' of preparation is more radical that a human handling of language (tokens devolve language beyond linguistic units of meaning): this complex preparation is not revealed, yet it is thoroughly worked upon. A more fundamental position is that all possible combinations have already been observed (we revert to an understanding of the finite; → despite the limits of human comprehension giving way to *seeming* infinity. I shall return to this, with reference to 'The Library of Babel'; another innovative digital example of combinatory writing).

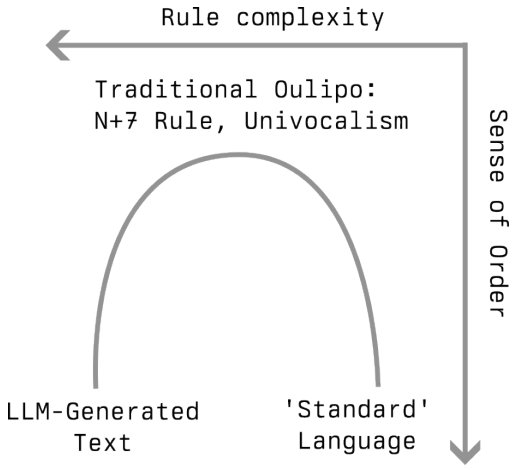

Figure 4: Tom Savage, 'The Potential Literature Horseshoe', 2024 (https://sav.phd/posts/oulipo).

(3) Erotic: Barthes makes reference to eroticism, in a broad ethical sense: 'it would be absurd to identify the eroticism of the city merely with the neighbourhood reserved for such pleasures […] I am using eroticism or *sociality* here without differentiation. The city, essentially and semantically, is the site of our *encounter with the other*'.[85] It is perhaps worth remembering, the Turing Test (a measure of a machine's ability to exhibit intelligent behaviour indistinguishable from that of a human) was based upon the 'imitation game': a Victorian parlour game, whereby participants had to guess the gender of another person. We might consider, here, a certain ludic quality to Barthes' interest in the erotic → in structuralist terms: *all* language is a constant process for simulating and dissimulating, of imitating or approximating as a means to *carry* meaning; to attend to the *other*.

Barthes suggests of a different 'economy' of meaning between city centre and the suburbs: 'the center-city is always experienced as the space in which certain subversive forces act and are encountered, forces of rupture, ludic forces'. He suggests the peripheries of Paris, for example, experience the centre 'semantically as the privileged site where the other is and where we ourselves are the other, as the site where one plays'.[86] His interest in 'play', 'encounter' and 'rupture' suggest of another significant new methodology: Deep Learning. Considered a 'breakthrough moment', deep learning was first properly exhibited in 2012 with a computer vision model, AlexNet, designed with a deceptively simple goal: to identify the primary object in an image. Deep learning systems use 'neural networks' (modelled on the neural network of the human brain) to 'learn' through an iterative process of free-form training. The underlying technique of backpropagation is key here, allowing a model to adjust its weightings:

> [W]hen an error is spotted, adjustments propagate back through the network to help correct it in the future. Keep doing this, modifying the weights again and again, and you gradually improve the performance of the neural network so that eventually it's able to go all the way from taking in single pixels to learning the existence of lines, edges, shapes, and then ultimately entire objects in scenes.[87]

The association with 'play' is pertinent, as the deep learning methodology was more fully realised through the development of Deepmind's AlphaGo – a model that famously beat the world champion Go Master Lee Sedol in 2016. Deepmind began by letting AI models learn (from scratch) rudimentary computer games. What is crucial about the deep learning method is the ability to learn 'intuitively', *without supervision* (i.e., without prior human labelling of data), and without needing to go through all possible configurations. Indeed, there is not enough compute in the world to 'learn' Go:

> It is exponentially more complex than chess. After just three pairs of moves in chess there are about 121 million possible configurations of the board. But after three moves in Go, there are on the order of 200 quadrillion (2×10^{15}) possible configurations. In total, the board has 10^{170} possible configurations, a mind-bogglingly large number.[88]

Beyond a *gridded* terrain of a game such as Go or Chess, the ability of computer vision to detect (in real time) the existence of lines, edges, shapes etc., suggests of radical 'eroticism'; open to ever greater degrees of learning.

Image diffusion models (such as DALL-E and Midjourney) generate massive virtual neural patterns (using high dimensional maths to store a huge array of probabilities). From these patterns 'decisions' can be made as to what is the most 'likely' appropriate rendering. In both cases, there is an attempt to render the *most appropriate* image. Rather than the Turing Test, we might understand diffusion models akin to the Game of Consequences, whereby you are filling in the very next piece of the picture (the surprise of the 'next' render of an image diffusion model can be just as disarming as when opening up the folds the paper after a round of Consequences!). Similarly, as with the Large Language Models, these self-supervised image models have no 'mind's eye' → they are not working to constraints of art historical categories or aesthetic terms such as figure and ground. They do not work with any such categories, but instead, due to the enormity of information parsed, they can locate and operate with wholly different 'units of meaning', which may coalesce of mere 'clusters' of pixels that the human eye may never notice or be able to see. The diffusion method first reduces data to noise before then building up an image (→ imagine tuning out of a radio station, then re-tuning to a coherent station). The technique provides the means for the *original* generation of imagery, thereby introducing a new *creative* property to computer vision. Despite the complexity of imagery, which does not adhere to 'grammar' in the sense we would say of language, the 'art' of this technique is its ability to form predication models of pixels in a similar manner to predicting words in language models; i.e., by training models on the smallest units of meaning (whether words or pixels). Furthermore, the integration of Large Language Models and image diffusion knits together both word and image making. In so doing, the models appear to solve the apparent 'problem' of the complexity of images, which human culture has obsessed over, over millennia, and which computers can now format quickly as *information*.

In each of the three observations (symbols, signifiers and the erotic), Barthes' emphasis is increasingly upon the *writerly*. Yet, he does not let go entirely of the utopia of a certain science of semiotics (→ contemporary AI methods suggest a 'certain' science is indeed possible). In drawing his talk to a close, he states: 'Starting from these readings, from this reconstitution of a language or of a code of the city, we might orient ourselves toward means of a scientific nature: investigation of units, syntax,

etc.'[89] NB. Despite having drawn our attention to Barthes' 'unfortunate' sleight of hand (that he never really intended a science of signs), Knight is as much complicit in the turn to the post-structural: 'one shudders at the thought of how the Eiffel Tower essay might have turned out if Barthes had applied this sort of approach'.[90] Indeed, Barthes' iconic essay is of course attuned to the importance of the movement of the signifier. The Eiffel Tower, he writes, is present to the entire world:

> First of all as a universal symbol of Paris, it is everywhere on the globe where Paris is to be stated as an image […] [T]here is no journey to France which isn't made, somehow, in the Tower's name, no schoolbook, poster, or film about France which fails to propose it as the major sign of a place and of a people.[91]

Figure 5: The image of the Eiffel Tower overlaying the Brandenburg Gate during restoration work. Berlin, 2003.

The signifier travels again and again: in the early 2000s, I spotted the Eiffel Tower in Berlin – presented on a hoarding during the restoration of the Brandenburg Gate: a marker of a myth of the various uniting crossroads

of Europe (Figure 5). As we know from Barthes: 'myth is a type of speech', defined not by 'the object of its message, but by the way in which it utters [its] message'.[92] It enables a double structure (→ e.g., the Eiffel Tower is at once historical and ahistorical), both *parole* and *langue* = myth residing somewhere between the two (Barthes: myth is depoliticised speech). Looking at the photograph I took over 20 years ago, of the Eiffel Tower wittily woven into the Brandenburg Gate (under the 'banner' of progress at the behest of Germany's T-Mobile), it could very well pass as any one of many images generated today in seconds by an AI generative model. The plasticity of myth (foretold by Barthes) is realised in mathematical space (an infinite (?) preparatory space for the signifier).

Argo

Barthes ends his article with an allusion to the galley *Argo*, 'of which each piece was no longer an original, yet still remained the ship *Argo*, i.e., a group of readily legible and identifiable significations'. [93] In his 'autobiography', Barthes gives further explanation to this allusion: 'by dint of combinations made within one and the same name, nothing is left of the origin: *Argo* is an object with no other cause than its name, with no other identity than its form'.[94] *Argo* is the city: it is always there, yet it is refashioned every time we speak of it, every time we enter it. Today, this refashioning of meaning is extended and made ever more evident with the millions of texts online, all writing and re-writing about the city and much more besides. *Argo* ≈ Large Language Models.

Let us cast our minds back once more to Barthes' structural analysis of narrative, whereupon he remarks of the phenomena of stories as 'human material, a class of thing which humans produce'. At first, it appears we cannot impose any sort of order upon narratives: 'There are millions and millions of narratives', he states, 'developed over an indefinite period of time, the origins of which are unknown. […] Narrative is everywhere'.[95] To the human faculty, narratives (and melodies) might as well be infinite: the maths is too big to contain. Yet, this is not the same as saying narratives of the world *are* infinite. In *Words and Rules*, Steven Pinker provides us with some of the numbers. He evokes Jorge Luis Borges's story 'The Library of Babel' (as '[p]erhaps the most vivid description of the staggering power of a combinatorial system'). As the story goes, 'somewhere in the library is a book that contains the true history of the future (including the story of your death), a book of

prophecy that vindicates the acts of every man in the universe, and a book containing the clarification of the mysteries of humanity'. Of course, even after the human species is made extinct, the library (and its combinatorial possibilities) remains. Yet, *technically*, Pinker explains, 'Borges needn't have described the library as "infinite". At eighty characters a line, forty lines a page, and 410 pages a book, the number of books is around $10^{1,800,000}$, or 1 followed by 1.8 million zeroes. That is, to be sure, a very large number – there are only 10^{70} particles in the visible universe – but it is a finite number.'[96]

Pinker's calculations become all too real in an online 'artwork', *The Library of Babel*, by Jonathan Basile.[97] Organised into hexagon shaped rooms, Basile's library provides numbered 'locations' for any possible 3200-character combination of English letters, comma, space, and period. Each room has four walls of books containing five shelves with 32 volumes of 410 pages each. Each page is given a unique sequential page number in base 10. The text on each page is encoded into this number, which in turn (based on an algorithm) creates a seed to generate a unique large number. This number is converted into base 29 (representing each letter in the English alphabet, as well as the comma, the space and the period). Basile's algorithm ensures each combination, and the same page number, will create the same output every time, meaning what is on each page is already *predetermined*. In other words: every page already exists in principle, or rather in a virtual, mathematical space. It is already 'prepared' and only needs to be looked up. Furthermore, each page can be converted through the inverted algorithm and turned into the exact page number they are found on. It is an uncanny experience to be able to find the permanent location for any 3200-character text ever written or to be written.

True to Borges's tale, somewhere in the library is description of your own death, along with every poem, every joke, every confession. Anything that can or could be said is on this site. The quandary, then, is whether you are only ever able to look up what has already been prepared. → Suffice to say, this very thought, this very utterance is already located in the Library:

Figure 6: Output from Jonathan Basile's *The Library of Babel*, listing (as part of a 3200 character long entry): 'True to Borges's tale, somewhere in the library is description of your own death, along with every poem, every joke, every confession. Anything that can or could be said is on this site. The quandary, then, is whether you are only ever able to look up what has already been prepared'.

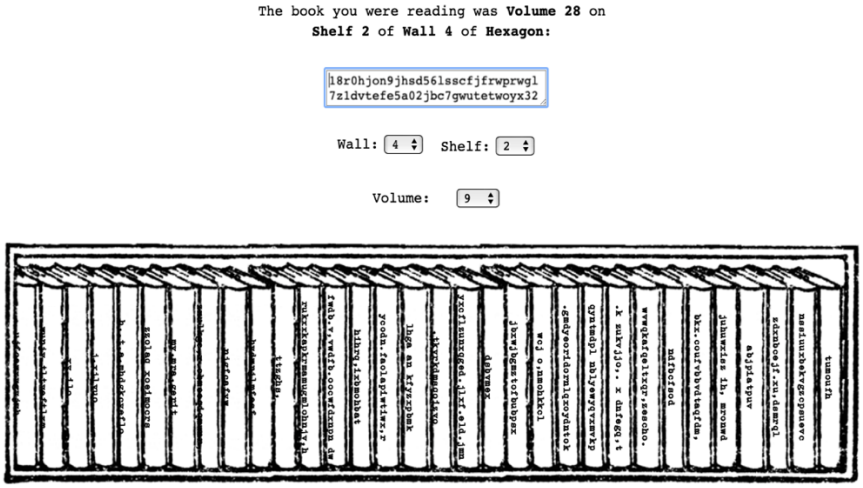

Figure 7: Jonathan Basile's *The Library of Babel,* showing contents (in Figure 6) held at Vol.28, Shelf 2, Wall 4, Room number '18r0hjon9jhsd56lsscfjfrwprwgl […] 2jucnca8csg0wp5eu2kz' [3200 characters long].

Of course, we want to believe there is a difference between the artificial generation of words and what a person actually says → we put this down to a matter of agency. Yet, is it just a matter of time? By Basile's 'calculus', all language is in preparation, in storage, ready to be used. While *seemingly* infinite, it is the (albeit massively) finite nature of language that draws us back to an understanding of the statistical turn in AI development and of preparatory space. At a human level, we can have no sense of the magnitude of words, sentences and narratives, yet for high performance computing (and potentially with quantum computing still to come) the sums are within range. And not just words, but the computation of the massive array of style in words, sounds and gestures, even the spaces between and around words. → Benjamin Bratton's discussion of internet addresses in his conception of the 'Stack' is suggestive of how the spatialising of digital information, including not only words but their relational positions can all be uniquely addressed. [98] The current configuration of Internet Protocol (v6) addresses are 128 bits long, theoretically allowing for 3.4×10^{38} combinations, meaning we could assign an address to every atom on the surface of Earth a hundred times over.

Whether human or computer generated, knowledge is *produced* by classification, with its gathering together (whether encyclopaedias, mythical songs, libraries, archives or in the high-dimensional clusters of information of virtual neural networks). What is common to *practices* of knowledge is the need to rely upon prior information and information beyond immediate grasp. Akin to the 'Text', Umberto Eco evokes the idea of the 'Total Encyclopedia', to suggest the notion of a 'totality of knowledge' that is always *productive* and in operation:

> I've been used to putting my faith in other people's knowledge. I confine my doubts to some specialized sector of knowledge, and for the rest I put my trust in the Encyclopaedia. By 'Encyclopedia' I mean the totality of knowledge, with which I'm only partly acquainted but to which I can refer because it is like an enormous library composed of all books and encyclopedias – all the papers and manuscript documents of all centuries, including the hieroglyphics of the ancient Egyptians and inscriptions in cuneiform.[99]

Eco places himself in the 'trust' of the Encyclopedia (with the first letter capitalised to denote its totality), *which is not the same as being fully trustworthy*. Nonetheless, the pursuit (and inhabiting) of a reservoir of knowledge is formed of an enduring desire (→ the 'bit' of information[100]).

In a remarkable article in *Science News-Letter*, from 1937, an entity called the 'world brain' (borrowed from H.G. Wells) is discussed, with 'Librarians, scientists and editors, and others who marshal and create the written record of civilisation' bringing together 'the intellectual resources of this planet into a unified system'.[101] The statement is predictive of the World Wide Web (suggesting at the time the 'solution' resides in the new technology of the time: microfilm): 'The nuclei of this world brain exist in the various great intellectual centers – the libraries, journals and indices of recorded knowledge – and the task considered is how to exchange and distribute more effectively the past, current and future accumulations in all fields of human endeavor.'[102] Inevitably, it is this 'brain' or network of information (= Text) that has made possible the massive training of data for AI.

Typically, given the media stories that prevail, it might be said AI technology marks a paradigm shift, or, to use Michel Foucault's term, a new episteme (i.e., a new way of thinking).[103] To watch the speed and adeptness of an AI application producing highly credible text (even a summary of Foucault's work!), can seem, on the surface, almost 'magical'. Yet, at root, its operation is drawn *through the network* of knowledge, the 'world brain'; through both a temporal and material set of connections (= traversal of the Text).

To the end, Barthes placed his 'trust' in *writing*, within Literature → echoes of Hugo's account of printing, which, were it a building, would be 'colossal'; 'Some statistician or other', Hugo writes, 'has calculated that if all the volumes printed since Gutenberg were piled one on top of another they would reach as far as the distance from the earth to the moon'.[104] However, rather than consider literature in its *entirety*, Barthes would suggest all accumulated knowledge can as much harbour within a single novel. 'In a novel like *Robinson Crusoe*', he notes, 'is a historical knowledge, a geographical, a social (colonial), a technological, a botanical, an anthropological knowledge (Robinson proceeds from Nature to culture).'[105] Were all else to be expelled, he argues, it is Literature that must be saved:

> for all knowledge, all sciences are present in the literary monument. […] Yet literature, in this truly encyclopaedic respect, displaces the various kinds of knowledge, does not fix or fetishize any of them; it gives them an indirect place, and this indirection is precious. On the one hand, it allows for the designation of possible areas of knowledge – unsuspected, unfulfilled. Literature works in the interstices of science. It is always behind or ahead of science […]

> The knowledge it marshals is, on the other hand, never complete or final. Literature does not say that it knows something, but that it knows *of* something.[106]

As Large Language Models have already demonstrated (with the techniques of deep learning, backpropagation and attention), it is feasible to work with smaller corpuses (albeit with rather more than just *Robinson Crusoe*) to nonetheless generate more generalised and fluent systems of meaning. Barthes had the right intuition (= we are always in the space of preparation), just the wrong conclusion: the turn to the signifier ≠ the infinite.

Notes

[1] This paper presents the 'fantasy' of a lecture. It deliberately adopts the style (and typography) of Roland Barthes's late lecture courses, published as Roland Barthes, *The Neutral: Lecture Course at the Collège de France (1977–1978)*, trans. by Rosalind E. Krauss and Denis Hollier (New York: Columbia University Press, 2005); Roland Barthes, *The Preparation of the Novel: Lecture Courses and Seminars at the Collège de France, 1978–1979 and 1979–1980*, trans. by Kate Briggs (New York: Columbia University Press, 2011); and Roland Barthes, *How to Live Together: Novelistic Simulations of Some Everyday Spaces: Notes for a Lecture Course and Seminar at the Collège de France (1976–1977)*, trans. by Kate Briggs (New York: Columbia University Press, 2013). The specific thematic of 'preparation' (responding to *The Preparation of the Novel*) is important to my overall argument regarding the 'computation' of language (vis-à-vis recent developments with Large Language Models and massively scaled compute). The form of the lecture gives an additional 'risk' of subjectivity (a degree of the preparatory, of continency through speech). The form, then, in itself is a preparatory method → there is an equivocation at stake: we can take these *written* notes as merely preparatory to a lecture still to be delivered (borrowing the typographic use of the → , found in the publications of Barthes' lecture course, which can be read as a cue to a possible aside or further development of an argument should there be time when presenting). Yet, equally, the deliberate use of typography, subheadings and the overall structure of the text might readily suggest an edited transcript – whether of a real lecture given in the past, or only ever devised as a fiction. The argument made for a mathematical 'preparatory space' *always already* in existence of language is thus maintained in juxtaposition to a deliberate *play* of speech and writing. It is worth noting, Barthes was sceptical of 'thinking aloud', of speech equating to *thinking*, 'if by this we understand a genuine interaction where the subject, live and in real time, puts his initial assumptions at risk'. Annette Lavers, 'Barthes and the Interview: The Politics of Seduction', *Nottingham French Studies*, 42.1 (2003), 48–60 (p. 48). Barthes' preference was to write: 'to write means in a certain manner: *I think better*, more firmly; I think less of you, I think more for the "truth"'. Roland Barthes, *The Grain of the Voice: Interviews 1962–1980*, trans. by Linda Coverdale (New York: Hill and Wang, 1986), p. 6. My paper largely accepts this view, but nonetheless allows for an oscillation, which perhaps is more in keeping with Barthes' wish for a transferential and textual space, which he attributed to the seminar. Roland Barthes, 'To the Seminar', in *The Rustle of Language,* trans. by Richard Howard (Berkeley: University of California Press, 1989), pp. 332–42.

[2] In his inaugural lecture in 1977, Barthes refers to research ('an enormous, almost unjust, privilege') as the 'professor's sole activity' at the Collège de France: 'to speak – I shall even say to dream his research aloud – not to judge, to give preference, to promote, to submit to controlled scholarship'. Roland Barthes,

'Inaugural Lecture, Collège de France', in *A Roland Barthes Reader* (London: Vintage, 2000), pp. 457–78 (p. 458).
[3] Barthes, *The Preparation of the Novel*, p. 3.
[4] Roland Barthes, 'From Work to Text', in *The Rustle of Language,* trans. by Richard Howard (Berkeley: University of California Press, 1989), pp. 56–64 (p. 56).
[5] Barthes, 'From Work to Text', p. 56.
[6] Roland Barthes, 'Semiology and Urbanism', in *The Semiotic Challenge*, trans. by Richard Howard (Berkeley: University of California Press, 1994), pp. 191–201.
[7] Barthes, 'From Work to Text', p. 64.
[8] Barthes, 'From Work to Text', p. 57.
[9] Barthes, 'From Work to Text', p. 57.
[10] Cf. Karen Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning* (Durham: Duke University Press, 2007).
[11] Claude Lévi-Strauss, 'The Structural Study of Myth', in *Structural Anthropology Vol. 1*, trans. by Claire Jacobson and Brooke Grundfest Schoepf (London: Penguin Books, 1993), pp. 206–31 (p. 229).
[12] In 'Late Barthes', Jonathan Culler upholds the relevance of Barthes's early work, praising its 'ludic systematicity'. He acknowledges the allure of Barthes' later writings, although maintains the discerning reader can manage to engage with the early work, even as they draw inspiration from the complexities and meta-linguistic qualities of Barthes' later work. Jonathan Culler, 'Late Barthes', *The Conversant*, 2013. Available online: https://www.academia.edu/37664041/The_Renaissance_of_Roland_Barthes [accessed 9 March 2020]; see also: Jonathan Culler, 'Analyzing Narrative: Roland Barthes' Forgotten Interview', *Theory, Culture & Society*, 39.7–8 (2022), 175–80.
[13] See Bernard Stiegler, *Technics and Time, 1: The Fault of Epimetheus*, trans. by Richard Beardsworth and George Collins (Stanford: Stanford University Press, 1998); and Achille Mbembe, *Brutalism*, trans. by Steven Corcoran (Durham, NC: Duke University Press, 2004).
[14] Michel Foucault, *The Order of Things: An Archaeology of the Human Sciences* (London: Routledge, 2002).
[15] Hubert L. Dreyfus and Paul Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics*, 2nd edn (Chicago: University of Chicago Press, 1983).
[16] Foucault, *The Order of Things*, p. 226.
[17] Michael Wooldridge, *The Road to Conscious Machines: The Story of AI* (London: Penguin Books, 2020).
[18] Sunil Manghani, 'Notes on Structuralism' (edited section), *Theory, Culture & Society*, 39.7–8 (2022), 117–80.
[19] Paolo Fabbri, Monica Sassatelli, and Sunil Manghani, 'On Narrative: An Interview with Roland Barthes', *Theory, Culture & Society*, 39.7–8 (2022), 159–174.

[20] Roland Barthes, 'Introduction to the Structural Analysis of Narrative', in *Image-Music-Text*, trans. by Stephen Heath (London: Fontana, 1977), pp. 79–124.
[21] Roland Barthes, *The Fashion System*, trans. by Matthew Ward and Richard Howard (Berkeley: University of California Press, 1990).
[22] Brian Moeran, 'A Japanese Discourse of Fashion and Taste', *Fashion Theory*, 8.1 (2004), 35–62 (p. 36).
[23] Barthes, 'Introduction to the Structural Analysis of Narrative', p. 79.
[24] Émile Benveniste, *Problems in General Linguistics*, trans. by Mary Elizabeth Meek (Coral Gables, FL: University of Miami Press, 1971).
[25] Culler, 'Analyzing Narrative', p. 175.
[26] Roland Barthes, *The Pleasure of the Text*, trans. by Richard Miller (New York: Hill and Wang, 1975).
[27] Roland Barthes, 'The Death of the Author', in *Image-Music-Text*, trans. by Stephen Heath (London: Fontana, 1977), pp. 142–48.
[28] Culler, 'Analyzing Narrative', pp. 175–76.
[29] Culler, 'Analyzing Narrative', p. 176.
[30] Culler, 'Analyzing Narrative', p. 177.
[31] Fabbri et al., 'On Narrative: An Interview with Roland Barthes', p. 163 (emphasis added).
[32] Fabbri et al., 'On Narrative: An Interview with Roland Barthes', p. 164.
[33] It is worth noting that later, in *S/Z*, Barthes turns explicitly away from the structuralist endeavour. In the opening line of the book, he references Buddhist ascetic practices ('to see a whole landscape in a bean'), which he equates with narrative analysis, in attempting to 'see all the world's stories […] within a single structure'. A task, he suggests, that is exhausting and 'ultimately undesirable, for the text thereby loses its difference'. Roland Barthes, *S/Z*, trans. by Richard Miller (Oxford: Blackwell, 2000), p.3. Arguably, Large Language Models remove the 'exhaustion' of the task. Accordingly, we might ask: does massively scaled compute only further navigate 'difference', or does it suggest of a massive, yet nonetheless finite combinatory language? See also Culler's 'Analyzing Narrative' (pp. 178–79) in which he ponders what might have happened had Barthes continued to explore some of the specific *levels* of analysis noted in the interview.
[34] Emily Bender et al., 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜', in *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2021), pp. 610–23.
[35] Steven T. Piantadosi and Felix Hill, 'Meaning Without Reference in Large Language Models', *Arxiv (Computer Science)* (2022), p. 5. Available online: https://arxiv.org/pdf/2208.02957.pdf [accessed 1 August 2024].
[36] Diksha Khurana et al., 'Natural Language Processing: State of the Art, Current trends and Challenges', *Multimedia Tools and Application*, 82 (2023), 3713–44.
[37] Barthes, 'Semiology and Urbanism', p. 191.
[38] Barthes, 'Semiology and Urbanism', p. 195.

[39] Iwona Blazwick, 'Century City', in *Century City: Art and Culture in the Modern Metropolis* (London: Tate Publishing Ltd, 2001), pp. 8-15 (p. 8).
[40] Raymond Queneau, *Cent mille milliards de poèmes*, new edn (Paris: Gallimard, 1982).
[41] Barthes, 'From Work to Text', p. 61.
[42] Victor Hugo, *Notre-Dame de Paris*, trans. by Alban Krailsheimer (Oxford: Oxford University Press, 1993).
[43] See chapters 'Abbas Beati Martini' and 'This Will Kill That' in Hugo, *Notre-Dame de Paris*, pp. 181–91 and 192–206, respectively.
[44] Hugo, *Notre-Dame de Paris*, p. 193.
[45] Jacques Derrida, *Of Grammatology* trans. by Gayatri Chakravorty Spivak (Baltimore: Johns Hopkins University Press, 1976); *Writing and Difference*, trans. by Alan Bass (Chicago: University of Chicago Press, 1978); *Dissemination*, trans. by Barbara Johnson (Chicago: University of Chicago Press, 1981); *Limited Inc*, trans. by Samuel Weber and Jeffrey Mehlman (Evanston, IL: Northwestern University Press, 1988).
[46] Barthes, 'From Work to Text', p. 59.
[47] Xiaochang Li, *Divination Engines: A Media History of Text Prediction* [Doctoral Thesis]. Steinhardt School of Culture, Education and Human Development (New York University, 2017).
[48] Nicholas Mirzoeff, *How to See the World?* (London: Penguin, 2015), p. 17.
[49] Mirzoeff, *How to See the World?*, p. 6.
[50] Nathan Jurgenson, *The Social Photo: On Photography and Social Media* (London: Verso, 2019).
[51] Walter Benjamin, 'The Work of Art in the Age of Mechanical Reproduction', in *Illuminations,* trans. by Harry Zohn (London: Fontana Press, 1992), pp. 211–44 (p. 234).
[52] Siegfried Kracauer, *The Mass Ornament: Weimar Essays*, trans. by Thomas Y. Levin (Cambridge, MA: Harvard University Press, 1995).
[53] Barthes, 'Semiology and Urbanism', p. 192.
[54] Claude Lévi-Strauss, *Tristes Tropiques*, trans. by John and Doreen Weightman (Harmondsworth: Penguin Books, 1976), pp. 285–86.
[55] Lévi-Strauss, *Tristes Tropiques*, p. 286.
[56] Google for Developers (website), 'Machine Learning: Embeddings: Translating to a lower-dimensional space', https://developers.google.com/machine-learning/crash-course/embeddings/translating-to-a-lower-dimensional-space
[57] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (London: Profile Books, 2019).
[58] Barthes, 'Semiology and Urbanism', p. 196.
[59] Diana Knight, *Barthes and Utopia: Space, Travel, Writing* (Oxford: Clarendon Press, 1997), p. 62.
[60] Knight, *Barthes and Utopia,* p. 62.
[61] Barthes, 'Semiology and Urbanism', pp. 195–96.

[62] Barthes, 'Semiology and Urbanism', p. 196.
[63] Barthes, 'Semiology and Urbanism', p. 196.
[64] Barthes, 'Semiology and Urbanism', p. 201.
[65] Knight, *Barthes and Utopia,* p. 1.
[66] Barthes cited in Knight, *Barthes and Utopia,* p. 1.
[67] Barthes, 'Semiology and Urbanism', p. 201.
[68] Barthes, 'Semiology and Urbanism', p. 197.
[69] Barthes, 'Semiology and Urbanism', p. 197.
[70] Roland Barthes, *Empire of Signs*, trans. by Richard Howard (London: Jonathan Cape, 1982), p. 31.
[71] Barthes, *Empire of Signs*, p. 30.
[72] Barthes, *Empire of Signs*, p. 4.
[73] Lévi-Strauss, *Tristes Tropiques*, pp. 169–70.
[74] Mustafa Suleyman (with Michael Bhaskar), *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma* (New York: Crown, 2023), p. 56.
[75] Ashish Vaswani et. al. 'Attention is All You Need', *31st Conference on Neural Information Processing Systems, NIPS 2017*. Available online: https://arxiv.org/abs/1706.03762 [accessed 1 August 2024].
[76] Suleyman, *The Coming* Wave, p. 56.
[77] Barthes, 'Semiology and Urbanism', p. 198.
[78] Barthes, 'Semiology and Urbanism', p. 199.
[79] For a visual representation of Queneau's *Cent mille milliards de poèmes*, see Les Mathématiques, C'est Pas Automatique, *Cent Mille Milliards de Poèmes*. Available online : https://youtu.be/26-_Cf8llh0?si=h3U_WKbzUD4_60Kl [accessed on 1 August 2024].
[80] The 1961 edition of Queneau's *Cent mille milliards de poèmes* includes an epigraphic quote from Alan Turing: 'only a machine can appreciate a sonnet written by another machine'. Jonathan Baillehache, 'The Digital Reception of *A Hundred Thousand Billion Poems*', *Sens public*, (2021), 1–13. Available online: https://id.erudit.org/iderudit/1089666ar
[81] Tom Savage, 'The Potential Literature Horseshoe', 9 January 2024. Available online: https://sav.phd/posts/oulipo [accessed 1 August 2024].
[82] Savage, 'The Potential Literature Horseshoe'.
[83] From Claude Berge's 'For a Potential Analysis of Combinatory Literature', in *The New Media Reader*, ed. by Noah Wardrip-Fruin and Nick Montfort (Cambridge: MIT Press, 2003), p. 171. Available online: https://nickm.com/classes/resources/12-oulipo-p.pdf.
[84] Savage, 'The Potential Literature Horseshoe'.
[85] Barthes, 'Semiology and Urbanism', p. 199 (emphasised added).
[86] Barthes, 'Semiology and Urbanism', p. 200.
[87] Suleyman, *The Coming* Wave, p. 53.
[88] Suleyman, *The Coming* Wave, p. 48.
[89] Barthes, 'Semiology and Urbanism', p. 201.

[90] Knight, *Barthes and Utopia,* p. 62.
[91] Roland Barthes, *The Eiffel Tower and other Mythologies*, trans. by Richard Howard (Berkeley: University of California Press, 1997), p 5.
[92] Roland Barthes, *Mythologies*, ed. and trans. by Annette Lavers (London: Vintage Books, 2009), p.131.
[93] Barthes, 'Semiology and Urbanism', p. 201.
[94] Roland Barthes, *Roland Barthes by Roland Barthes*, trans. by Richard Howard (New York: Hill and Wang, 1975), p. 46.
[95] Fabbri, et al., 'On Narrative: An Interview with Roland Barthes', p. 163.
[96] Steven Pinker, *Words and Rules: The Ingredients of Language* (New York: Basic Books, 2015), p. 16.
[97] Jonathan Basile, *The Library of Babel*. Available online: https://libraryofbabel.info [accessed 1 August 2024]
[98] Benjamin H. Bratton, *The Stack: On Software and Sovereignty* (Cambridge, MA: MIT Press, 2016)
[99] Umberto Eco, *Six Walks in the Fictional Woods* (Cambridge, MA: Harvard University Press, 1994), p. 90.
[100] Cf. James Lovelock refers to the basic unit of information, the 'bit' (having a value of zero or one, true of false), as 'primarily an engineering term, the tiniest thing from which all else is constructed'. Following which, the future world is one where 'the code of life is no longer written solely in RNA (ribonucleic acid) and DNA, but also in other codes, including those based on digital electronics and instructions that we have not yet invented'. James Lovelock, *Novacene: The Coming Age of Hyperintelligence* (London: Allen Lane, 2019), p. 92.
[101] 'Documentation Congress Step toward Making "World Brain"', *The Science News-Letter*, 32.861 (1937), 228–29.
[102] 'Documentation Congress Step toward Making "World Brain"', p. 228.
[103] Michel Foucault, *The Archaeology of Knowledge*, trans. by A.M. Sheridan Smith (London: Routledge, 2002).
[104] Hugo, *Notre-Dame de Paris*, p. 205.
[105] Roland Barthes, 'Inaugural Lecture, Collège de France', p. 463.
[106] Roland Barthes, 'Inaugural Lecture, Collège de France', p. 463.

ABOUT THE AUTHOR

Sunil Manghani is Professor of Theory, Practice and Critique at University of Southampton and a Research Fellow of The Alan Turing Institute. He is managing editor of *Theory, Culture & Society*, and Co-Editor of *Journal of Visual Art Practice*. His books include *Image Studies*; *Zero Degree Seeing*; *India's Biennale Effect;* and *Farewell to Visual Studies*. He curated *Barthes/Burgin* at the John Hansard Gallery, and *Building an Art Biennale* and *Itinerant Objects* at Tate Exchange, Tate Modern.

COPYRIGHT INFORMATION