Hello and thank you for having me here today. Thanks also to Karina and Machee for their great talks. And also to Machee for stylo, which I have used in the past. It is great to be able to speak with you today about my research and I am extremely grateful for the opportunity. Of course, I am also cognizant of the fact that we are separated by distance but brought together by technology. I am grateful for this as a safety measure during the ongoing pandemic but regret the loss of proximity. It’s almost the inverse of Wordsworth’s residency in London, a situation where instead of living ‘Even next-door neighbours, as we say, yet still / Strangers, and knowing not each other’s names’ we live apart, but very much have our Zoom-names attached.

However, there is a certain poignancy to the digitally mediated situation in which we find ourselves for my talk today. For while the perhaps best known phrase for computational literary methodologies is “*distant* reading”, I am going, today, to talk to you about the ways in which computational and digitally mediated reading practices can bring us *closer* to texts; to show how such approaches need not alienate us from interpretative paradigms, but can furnish us with fresh evidence to feed our interpretations. I’m going to talk to you about one aspect of my “simple” computational methodological interventions, which I bill under the broad heading of “registers of knowledge”, a phrase that speaks to how literary reading techniques can be used to provide us with access to a set of differing epistemologies that all take inscriptive forms: historical, scientistic, digito-factual, and literary.I’m going to talk to you today about my work on David Mitchell’s *Cloud Atlas*

Although we like to think of this as a contemporary debate, it has a long pedigree/lineage. Vernon Lee proposed a “statistical experiment”—a quantitative analysis—on literature in her 1923 *The Handling of Words* – in which she said that such “statistical labours had brought to light remarkable and suggestive differences in the use to which those words were put” in Henry James and Hardy, even when the statistical labours themselves were potentially “fruitless”. That is, Lee tells us that the count may not be important, but the act of counting may be. William Empson wrote, in *Seven Types of Ambiguity*, of how, despite the serious challenges of a more scientistic approach to art, he sits on the side of reason in criticism.

**[SLIDE]**

For various reasons, *Cloud Atlas* is a spectacularly good novel for digital studies. It’s a work that moves through six different time periods, from 1850 through to a far-flung future dystopian science fictional world. Each segment of the novel is written in a generic pastiche. The Pacific Diary of Adam Ewing, for instance, pays homage to Herman Melville’s seafaring narratives and is supposedly written in the language of its time. By contrast, the chapter entitled Half-Lives is written as a pacey airport thriller, the only section of the novel to adopt a present-tense voicing. Sloosha’s Crossin’ the final chapter, adopts a supposedly regressed future language, based on Russell Hoban’s 1980 novel, *Riddley Walker*. This stylistic variety gives a rich playground for digital aesthetic critique and literary-linguistic profiling.

So far, so straightforward. Mitchell plays another trick, though, by splitting each of the chapters in two, so that we cascade halfway through each narrative into the future, next section, before the second half of the book tumbles gradually back towards 1850, one chapter at a time, in reverse order. In this way, *Cloud Atlas* is often described as a Russian Doll novel. Finally, I should note that there are two editions of *Cloud Atlas* in circulation with significant textual variation between the editions due to an elaborate publishing history, on which my work is the authoratative source.

However, importantly for my story today, this is also a novel that spurns naive or straightforward historiographic narratives of improvement; what we might generally gesture towards as the “style of history writing that told a story of constitutional and social progress” as Stefan Collini concisely summarises ‘whig history’ . By shunning a simple temporal unfolding in narrative towards progress – and also by giving his most futuristic setting as one that is more like the iron age – Mitchell appears to pass comment on those historiographic movements that would proclaim unending progress as a move from darkness into light.

**[SLIDE]**

Or so I thought. Except that it turns out that the novel is betrayed in this ambition by its own language – if we pay close attention. I discovered this, though, only through a set of formalist questions, answered by a computational approach. Having seen Mitchell’s performance of the language of 1850, I was curious to know: how much does he get wrong? I set out, therefore, to write a computer program to detect terminology that should, in theory, have been inaccessible to Mitchell’s narrator. (I should add that there are many complexities to this, including the fact that the novel implies that the diary object has been edited by around 1895, but I am glossing over that for concision today.)

To cut a long story short, after much computational work, there are just three terms that, I feel, can be said with certainty to have been absolutely inaccessible either to Mitchell’s historic author or editor: spillage, from ~1934; latino, from ~1946; and lazy-eye, from ~1960. Focusing on these latter two terms, the word “latino” did not actually come to prominence until after the Second World War. But the use, here, of a racial epithet has an important different effect for the construction of a stylistic imaginary of the nineteenth century, to which I will turn shortly. Secondly, Mitchell gives us a “parlour […] inhabited by a monstrous hog’s head (afflicted with droop-jaw and lazy-eye, killed by the twins on their sixteenth birthday”. The sources that I consulted give this slang term for amblyopia as first appearing in the middle of the twentieth century.

So far so formalistically interesting: this is a very good attempt at linguistic mimesis within a work of purported historical fiction. But it is also clear that readers are poor at identifying terms that are anachronistic. I had no idea that “spillage” came from the 1930s. Indeed, I am unsure that, if asked, readers would be able to point to these words as the markers of the language seeming not to “ring true”. How, then, does one create a linguistic styling that appears mimetic of 1850 when working under the assumption that readers will not know when words were coined?

I began to question whether Mitchell might simply be using uncommon language to create the perception of a stylistic affinity with Victorian-era prose for the twenty-first-century reader.

**[SLIDE]**

This formalistic question pushed me back to the text and, specifically, to examine more closely the “archaic overloading” of Mitchell’s novel. So I asked a second formalistic question that I addressed with computational methods: which words does David Mitchell use in this “1850” section that are *not* present in a magazine corpus from the year of the novel’s actual publication, 2004? In other words, what vocabularies build our historical imaginary of nineteenth-century style?

Some of the answer to this question proves gimmicky. As with Pynchon’s *Mason & Dixon*, Mitchell replaces all uses of “and” with an ampersand in this section to yield a type of imagined typographical synthesis with the historical period. This use is, itself, not strictly accurate for the time period. Looking at the first passage of *Moby-Dick* as a correlative text, for instance, and we can see that ampersands do not appear instead of “and”, and several passages would be totally acceptable in contemporary spoken English (were they not so well known already): “Call me Ishmael”, “There is nothing surprising in this”.

However, the more disturbing and perhaps political linguistic conclusion at which I arrived is simply that Mitchell’s language of 1850 is replete with terms of racial abuse and insensitivity to disability. While the text redacts the word “damn” for fear of offence, it prints outright the n-word and uses the terms “blackamoor” and “latino”. The latter appears as though, therefore, it fits within the nineteenth-century context, because it sounds like another “dated” term of racial categorisation from the era of Empire. The same goes for “lazy eye”, which sounds like a pejorative term for a disability.

There are two questions, really, that arise from this close attention to language, which I identified through the use of computational methods. First, is it ethical for Mitchell to use such techniques, which rely on specific linguistic appropriations, in order to create an aesthetic outcome of historical effect? Second, and perhaps more importantly, doesn’t this contradict the more common reading of an anti-Whig historiography in the novel? The reason that readers might think that the Ewing diary “sounds like” it came from 1850 is because of a belief that these terms are outmoded now and are not used; a belief, fundamentally, that things are better now than they were at that time. My computationally distant but linguistically and politically close reading of the text yielded to me an oppositional feature of Mitchell’s language that went against almost all other historiographic – and, therefore, also political – readings of the novel. *Cloud Atlas* constructs its historical imaginary of nineteenth-century style through a reliance on a narrative of linguistic progress and improvement that contradicts its own meta-historical narrative. The computational approach gave me a re-integrative synthetic approach to distance and depth.

While, then, most people *imagine* computational methods to mean a move away from critical evaluative approaches, my work here actually gave me a way to re-approach literary value. This type of reading involves the success or otherwise of the novel’s ‘reality effects’, as Barthes put it. The question becomes one of literary *success* and whether we should read the Diary segment of this novel as realist, pastiche, or parody – with different success criteria mapped against each.