Literature in the Digital Age: From Close Reading to Digital Reading

Video transcript

What do we learn from distant reading?

[PHILIPP SCHWEIGHAUSER] Thanks, Hugues, for taking the time to talk about the Euterpe project.

[HUGUES MARCHAL] Thank you, Philipp.

[PHILIPP SCHWEIGHAUSER] You've spent years analysing scientific poems. And my first question is, isn't that very dry stuff? Aren't there much more interesting texts to read, say novels? What first got you interested in scientific poetry?

[HUGUES MARCHAL] Well, some years ago I was working on contemporary French poets who were very interested in the representation of the body. And many of them would quote texts from biologists or medical texts. And this was considered by most critics as something very original, very modern.

But I was not completely satisfied with this story, this version of the history, because I knew that some poets in the 19th century were also very – had an active dialogue with sciences. So this is why I discussed this with colleagues and we all realised that this whole previous tradition was not known well enough. And we began the project. But I want to correct what you just said. Some of these poems, some of these scientific poems are really, really great texts, are really interesting. Thank God.

[PHILIPP SCHWEIGHAUSER] OK. Good to hear. Good to hear. As beautiful as they are, you don't seem to be that interested in analysing, you know, single scientific poems and seeing how a single poem works and how it fits into its time and so on. If I get you right, you're more interested in sort of collecting metadata about hundreds of poems, information about the author's name, about the size of the book, about the contents, and so forth, and in sort of categorising, grouping, analysing them with almost scientific or rather scientific methods, really. And so my question is, what do we get out of these poems with these new methods that we couldn't get out of them with traditional literary scholarship, that boring people such as myself practice?

[HUGUES MARCHAL] Oh, you're not boring. I'm also a very boring person to that extent, because we really read those books in very traditional ways. We did some close reading. We were very interested in understanding the interplay between poetry and scientific texts themselves, scientific discourse, so we

looked at this very closely too. And so in that framework, using distant reading was just one of our tools. We used several tools. But we had to understand, a key question for us was the question of the disappearing of the genre as a genre.

So this is why we needed ways to be able to observe a whole trend of our century, because we had to, we were going against a still dominant discourse for which the genre disappeared at the beginning of the 19th century. And we needed evidence to show that. And we needed objective evidence to try and do so.

[PHILIPP SCHWEIGHAUSER] All right. So basically, for you distant reading and close reading are complementary practices.

[HUGUES MARCHAL] Completely. They are complementary. And as Moretti underlines very well, it's a question of scales, actually. There is no, distant reading is not very, in my opinion at least, it's not something which is very useful when you just have one text.

But it turns into something very useful when you have a huge amount of texts. Moretti gives the example of the 60,000 novels that were published during the Victorian time. And of course, you cannot read that if you are alone. What we did, of course, our corpus was not as big, we only have 500 poems, but we use them in ways that allow to see phenomena that you would not see with other elements. And our graphs also raise questions. For instance, when we can really try and study who these poets were, what they were doing, what was their profession? Were they poets? Were they writers at large? Were they scientists? Were they amateurs? Who were they exactly?

And distant reading also is, on a very basic level, it's something that has since 15 years roughly, it went into, in my opinion again, every scholar's practice, because when you use Google Books, for instance, to try to identify the source of a given quotation, or if you want to know how this quotation was used in several books at the same time, now you have wonderful tools just using BnF (Bibliothèque nationale de France), big online libraries, or Google Books. And to put it roughly, you don't need to be, you don't need any erudition any more. Of course you need it still, but in a certain, up to a certain extent, the computer does it for you now. It does the reading and you only get the results. You think from the results, but you don't read the book by yourself.

[PHILIPP SCHWEIGHAUSER] Ah. So there's also lighter forms of distant reading. Still, when Moretti published Graphs, Maps, Trees, this created an uproar among many literary scholars. Many felt that he wasn't really doing a proper job. He wasn't really analysing literary texts. He wasn't doing what literary scholars are supposed to do. And others also said that many of his findings were basically self-evident or banal even. And so my question is, what is your take on this controversy?

[HUGUES MARCHAL] Well, you know, Philipp, there are many rooms in the house of a literary critic. And Franco Moretti is not preventing people from using previous tools, previous existing tools. He brings a new tool that we can add to our set of tools. Still, I'm not advocating distant reading in a completely uncritical way. There are many limits to the method. There are limits to the data visualisation that you create. These visualisations are supposed to help you observe things that you couldn't see otherwise. And sometimes the charts are not up to the task. You have to select them and you have to, they have to work to let you know things that you wouldn't see otherwise, because creating a database can be painstakingly long. It takes a long time to read all the books, to register all the information that you need and to create very boring big files in which you split the information.

Second big thing is that you have to let people know about the limits of your results. Let me take an example. For the Delille project, we decided to create charts in which we only take into account the first editions. But it means that a very famous poet whose books were maybe published in dozens of editions only appear one time, the book will appear one time at the same level of a very obscure book that maybe no one read. So it's not, there are clear limits to what you can do.

And Moretti gathers together in his book data and results from several teams. And I don't know if you can really do that, because the way it was collected is different. And so maybe you create artefacts. Maybe you create things that seem to be objective but are not so objective.

I'm also not very much convinced by the trees in which Moretti tries to retrace the history of the development of certain literary devices. They seem a little too crude to me. And I'm not entirely convinced by the underlying biological model that there is in this kind of infography.

But still, when you use distant reading to contrast, collect, gather, and synthesise big data, you create some sort of composite portraits for a genre, say the novel at a given time or scientific poetry in our case. And we can use this kind of composite portrait to contrast it with individual works. And so you go from close reading to distant reading and then back to close reading with new questions. And this is what I think is very interesting.

[PHILIPP SCHWEIGHAUSER] Yeah. It sounds interesting. Hearing you talk about your project, I hear a lot about data and data visualisation and so on and so forth. And you know, a practical question, and it's really my final question here, would be: you were trained as a literary scholar. You were not trained as a computer programmer, yet you're using models and methods from the natural and the social sciences. What were some of your main challenges in making this project work?

[HUGUES MARCHAL] Well, in that respect, Euterpe was a very low-key project. It may seem very complex, but it's not. Creating the database and I would say creating our problems, our questions, was the

big task. Then the whole data visualisation thing is quite easy. And also, obviously for this kind of work, you have to work in a team. And I was not directly responsible for, no one was directly responsible for all the tasks.

For instance, it's one of my colleagues, Muriel Louâpre, from Université Paris Descartes, who was in charge of all the data visualisation. So she really developed the expertise for that. And if I may just add this, you just spoke of how there was a kind of a very, very easy to do, like some distant reading that you can do at home. And this is, to me it's also a key element. People need not be afraid of the techniques. It's here. It's easy to do, and how disciplines, like any disciplines, have to find ways of taking advantage of big data. It's a question for any discipline now, and it's the same for humanities.

We need to find ways of using all this big data that are now available online.

[PHILIPP SCHWEIGHAUSER] Thanks so much for sharing your insights into what really sounds like a wonderful project, and also for giving some advice to our learners how they could do this themselves.

[HUGUES MARCHAL] Well, thanks a lot, Philipp.

OUTTAKES:

Thanks for taking the time to talk about your Euter– [LAUGHTER] OK.

Thanks for taking the time to talk about your Euterpe project, Hugues. [LAUGHTER] Sorry. Come on.

Thanks for taking the time to talk about your Euterpe project. Thank you, Philipp. Well, you know, you spent years analysing scientific poems. Isn't that very dry stuff? [LAUGHTER] Hey, come on. [LAUGHTER]