

# New tools for new research questions and answers

Brett Greatley-Hirsch and James Baker talk to the British Academy Review about the potential of digital research in the humanities



Dr Brett Greatley-Hirsch is University Academic Fellow in Textual Studies and Digital Editing at the University of Leeds.



Dr James Baker is Senior Lecturer in Digital History and Archives at the University of Sussex.

‘Advances in computing have changed the nature and scope of the work we do in the humanities,’ says Dr Brett Greatley-Hirsch. ‘It’s not just a matter of efficiency – i.e., doing what we’ve always done, only faster. Nor is it simply a matter of scale – i.e., doing what we’ve always done, only bigger. Computing enables us to ask questions that were simply inconceivable for previous generations of scholars.’

Dr Greatley-Hirsch, University Academic Fellow in Textual Studies and Digital Editing at the University of Leeds, is among the first recipients of the British Academy’s new Digital Research in the Humanities (DRH) grants. Awarded at the end of 2018, in partnership with Jisc, the grants fund innovative research that applies new digital methods and tools to existing digital resources in order to yield new insights.

For his project, Dr Greatley-Hirsch will be using statistical procedures and machine-learning techniques associated with computational stylistics to analyse various samples of English drama, poet-

ry and prose from the late Elizabethan to the early Jacobean period. The aim is to find out whether a specific literary genre affects an author’s style, and, if so, how significant these differences are. He explains: ‘It’s about testing a longstanding assumption that genre affects style – an assumption that currently hampers authorship attribution study.’

It is only thanks to the recent development of new technology that Dr Greatley-Hirsch can attempt this kind of research. The sophisticated searches enabled by the combination of large-scale digitisation projects and natural language processing now allow scholars not only to trace the histories of generic forms, like the novel, sonnet or sermon, but also to investigate what Daniel Shore has termed ‘the genealogy of syntactic forms’. ‘The difference,’ explains Dr Greatley-Hirsch, ‘is whether you are studying the history of a specific phrase – e.g. “but me no buts” – or a broader phraseological unit – e.g. “(verb) me no (noun)” – or more general syntactic structure – e.g. “(verb) (pronoun) (adjective) (noun)”.’

Dr James Baker, Senior Lecturer in



© The British Academy

James Gillray's *La dernière ressource*, published by Hannah Humphrey in 1791, is one of over 12,500 prints described by Dorothy George in the *Catalogue of Political and Personal Satires Preserved in the Department of Prints and Drawings in the British Museum*. Image: British Museum (CC BY-NC-SA 4.0).

Digital History and Archives at the University of Sussex, is another DRH grant-holder. He and his colleagues are examining the catalogue of personal and political satirical prints produced by the historian Mary Dorothy George for the British Museum between 1930 and 1954. These prints cover political and social topics and everything in between. 'The political prints tend to be ephemeral responses to political dramas of the day,' says Dr Baker, 'while the social satires tend to cover topics with a longstanding appeal – for instance, jokes about gouty priests, ridiculous fashions, and roguish Irishmen that draw on stereotypes.'

Working with his co-investigator Dr Andrew Salway to combine traditional archive work with corpus linguistic methods, Dr Baker hopes to develop a generalisable toolset for understanding works of this kind. He is also looking at individual types and groups of words within Dorothy George's archive – analysing how the historian used the words, which words she used around them, and how she introduced various quotations.

So, what is it about Dorothy George's work that so interests Dr Baker?

'Every scholar who has used a "Golden Age" satirical print is indebted to Mary Dorothy George,' he says. 'Her descriptions of satirical prints are a substantive work of scholarship that elevated their subjects to canonical status among historians of long-18th-century British history. In turn, both the prints and George's descriptions of them have been integral to studies of print, culture, politics, and social life in Georgian Britain. Whether in book form, on microfilm, or online, George has been a constant interlocutor between the historian and this remarkable era of graphic reproduction.'

As with Dr Greatley-Hirsch's project, it is recent developments in technology that have made Dr Baker's research possible. These developments, particularly relating to image search and recognition, are of special interest to Dr Baker.

'I was lucky enough to be invited a couple of years ago to do some work for another British Academy project run by a colleague in Edinburgh called Anouk Lang,' he says. 'She was running some workshops on digital methods, and she asked me to do something on computational image recognition. I found then that we had just hit the point where there were open source toolsets that I could use and install. They were very good at

finding images where there were two prints of very similar things. They could sometimes pick out generic landscape scenes, and they could just about work out that most landscapes had a church on the left or the right and a tree in the middle, and that kind of stuff. That was actually quite exciting at the time because, at the very least, it was finding matches.

‘But what is interesting to me is that this technology is changing rapidly. Every time I step away from it for a few months, I look back at it again and find that there are massive advances.

‘For me, it is exciting to think about how we can take the work we are doing with language to describe an image and join it up with the work that is going on at the computer vision end around how you look at an image and figure out which parts are which.’

The development of this kind of technology has occurred in tandem with the rise of artificial intelligence, especially ‘deep learning’, a concept which rose to wider fame in 2015 when Google’s Alpha-Go became the first computer program to defeat a Go world champion. Alpha-Go’s deep learning model mastered Go by competing in thousands of games against human players. In just a few years, it has introduced a series of innovative winning moves and overturned centuries of received wisdom.

But as impressive as this may be, Dr Greatley-Hirsch points out that the use of artificial intelligence presents a challenge to those working in the humanities that is perhaps more disciplinary than technical.

‘Neural networks of the sort used in deep learning exemplify the so-called “black box” problem,’ he says. ‘We can

observe and study the inputs and outputs, but the inner workings of the model that transforms them are essentially unknown. Humanities scholars will no doubt find this lack of explanation deeply frustrating.’

There are also challenges associated with capturing and processing information in the digital age. As a society, we create more information in one day than we have ever created before, which, Dr Baker says, means the historians of the future will have to be well-versed in digital methods and up to date with the newest toolsets.

‘There are going to be some real challenges,’ he says. ‘Though there is a large amount of information in public which is at risk – things on Twitter, Tumblr and Geocities, for example – there are projects trying to capture that. But the task of capturing personal archives is going to require an enormous amount of investment in infrastructure. Archivists have been doing an amazing job in very embattled circumstances for over a decade, trying to develop methods for capturing personal archives: for the person who walks in with a laptop, not a stack of papers. There is a lot of information on that individual device, but because of the difficult nature of working with that device, they are not able to take too much of that material.

‘We need time and energy to be spent thinking about how we support the creation of broad digital archives of personal papers of that kind.’

So, what do the researchers think lies in store for humanities research generally? Will we see a pivot towards digital research methods, and to what extent might these methods displace more traditional

approaches?

‘“Displace” may be the wrong word,’ says Dr Greatley-Hirsch. ‘It perpetuates a fear of mechanisation, of being replaced by machines, which I’m convinced goes some way to explain the resistance to digital scholarship from certain quarters. Something closer to “augment” is more accurate, because there will always be a space and need for traditional approaches. Put simply, there are things that machines can do that human beings cannot – but the reverse is also true, especially when it comes to interpretation of language and expression, thought and emotion.’

More broadly, Dr Greatley-Hirsch believes the adoption of digital research methods offers an important opportunity to encourage truly interdisciplinary research.

‘Understanding digital research methods in the humanities is also a necessary step towards preparing ourselves, and our students, for a changing world in which the digital is fast becoming the primary means of cultural creation, dissemination, and preservation. This means that our disciplines should value and promote numeracy as well as different types of literacy.’

So are digital history posts likely to increase as the years go by, and, more to the point, would that be a good thing?

Dr Baker isn’t sure. ‘I would partly be putting myself out of a job by saying digital history posts should not be there!’ he says. ‘But I have always imagined that in time they would not be.

‘History has always had a range of different methods that people have used. A recent methodological turn was towards oral histories, but people do not describe themselves as being an “oral history historian”. They still refer more to their areas of historical interest.

‘History and art history will change as disciplines when we start working with more contemporary sources. The introduction of home computers in the 1990s was huge – and since then our archives have become increasingly more digital. Historians will need to be more adept at using large web archives, or a personal archive that might not be a series of boxes but a hard drive instead. That will be the most natural tipping point.’

Brett Greatley-Hirsch and James Baker were talking to Joe Christmas.

**Understanding digital research methods in the humanities is also a necessary step towards preparing ourselves, and our students, for a changing world**