

Part I

Computational Methodologies
for the History of Ideas

1 Introduction

Peter de Bolla

This book explores the ways in which computational techniques for text mining can contribute to the history of ideas. Traditional approaches to intellectual history are based on the careful, close scrutiny of texts and contexts, reading at a human scale in order to construct lines of transmission and genealogies of ideas. These approaches have served us well: the history of political thought, for example, has established deeply researched canons of texts and writers that are widely accepted as being in dialogue with each other. Over the last twenty years, however, the migration of analogue archives to digital corpora has opened up the possibility of reading at scale, creating the conditions for inspecting the transmission of ideas across hundreds of thousands of texts. Most of these texts have been deemed too obscure or insignificant for sustained close attention, and they remain at the outer margins of our standard histories of ideas. But when they are aggregated in large-scale datasets that capture a culture's ways of thinking, they open up the possibility of reconfiguring the history of ideas, expanding traditional accounts of ideas as personal property exchanged between independent minds so as to include within such accounts the dispersal and transmission of ideas across an entire culture. Moreover, heretofore the history of ideas has been modelled on a linear transmission, in which one thinker engages in dialogue with another or one text provides the foundation for its successors. Today, using computational techniques, we have the potential to create different models that can track the emergence and transmission of ideas in networks and clusters of concepts that are constellated in many dimensions thereby replacing a single linear thread of transmission with far more complex interwoven lines of connection. The use of such techniques opens up far more granular accounts of how transmission is often 'lumpy' in which long periods of stability are punctured by intense moments of change. Such models, based on computational methods of interrogating archives – now rendered as datasets – that often derive information from texts that are orthogonal to the canons of great texts and thinkers, have the capacity to substantially augment our traditional

4 I Computational Methodologies for the History of Ideas

histories of ideas as we move from the individual to the social or cultural in general, tracking the movement of ideas across the largest ‘bandwidth’ we can assemble in order to understand how such ideas operate as cultural entities.

Over the last ten or so years, digital humanities (DH) approaches to historical research have become more common as researchers take advantage of earlier waves of digitisation of archives. Projects that have turned to DH methods include, among others, the University of Stanford’s Mapping the Republic of Letters, an ambitious collaboration between a number of international partners, including the University of Oxford’s Cultures of Knowledge project; Groupe d’Alembert, hosted by the Centre National de la Recherche Scientifique; and organisations in both Holland and Italy, which set out to map the correspondence networks within which scientists, *salonnières* and travellers operated during the Enlightenment. At the University of North Texas, Andrew J. Torget set up the Texas Slavery Project, which explores the expansion of slavery between 1837 and 1845 in the territories that would eventually become the state of Texas. Although not solely focused on historical projects, the Helsinki Centre for Digital Humanities, hosted by the University of Helsinki, has undertaken a number of research initiatives that bear upon historical inquiry.

In common with many early adopters of digital methods, these projects and others draw upon statistical and computational methods first developed in the social sciences for mapping the movement and circulation of varying forms of information. Alongside many other projects in the field of historical research, they use geospatial information system techniques for creating layered representations of these flows. Some universities, especially in the United States, were early to the party – notable here is the University of Virginia, which developed methods in the digital exploration of humanities scholarship and trained many of the first generation of DH scholars – but in the more recent past centres and projects dedicated to digital methods have become commonplace in many national and international scholarly contexts.

In spite of this welcome flourishing of the field of DH across the disciplines, this book is the first to focus explicitly on the history of ideas in its narrow guise as the inquiry into the history of political thought. Although a number of collections of essays have been exemplary in introducing various methods and techniques for studying ideas in a general sense – here one might look to Chloe Edmundson and Dan Edelstein, *Networks of Enlightenment: Digital Approaches to the Republic of Letters*, or to Simon Burrows and Glenn Roe, *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth-Century Studies* – none have yet set out to

explore the ways in which DH approaches and methods might contribute to the long tradition of scholarship that is closely associated with the monograph series published by Cambridge University Press and entitled *Ideas in Context*, that was initiated by Richard Rorty, Jerome B. Schneewind and Quentin Skinner over thirty-five years ago. The monographs published in this series have, of course, been varied in their topics and their authors more or less affiliated with what has become known as the Cambridge School in the history of ideas, an approach to intellectual history that has been at the core of the discipline – even if it has found critics and advocates of alternative approaches – for over forty years. In common with many robust and well-developed scholarly interventions into a disciplinary field, the Cambridge School has had its vicissitudes – indeed some scholars claim that there never was such a school. The present book does not seek to contribute to what at times, inevitably, may seem to be merely academic turf wars in the discipline of intellectual history. Instead, it takes as a given the utility of studying ideas ‘in context’ and asks what might happen when that context can be scaled up using digital resources, both databases and computational methods and statistical techniques. Given that it is just one volume, the explorations presented here can only be exemplary – essentially shining a light on just a few topics or areas of research within the larger terrain of the history of ideas. And given the restriction of space it was decided to softly limit the historical reach of the case studies that feature in Part II of the volume to the long eighteenth century. The purpose here is to simultaneously enquire into the particular features of this period as in some sense ripe for digital methods and into specific case studies built upon those methods. In part this was also motivated by the fact that another tradition in the history of ideas, which is associated with the German scholar Reinhart Koselleck and known as *Begriffsgeschichte* (or conceptual history), is grounded in the hypothesis that the period we direct our attention to here has a very distinctive outline in which, it is claimed, the character, shape or form of concepts in general began to shift into something that we might recognise as the structural composition of concepts in modernity. This argument is explored most carefully by Ryan Heuser in Chapter 12, the last chapter of the book, but we see all of the case studies as operating with this hypothesis in the background. The aim, then, is to create a layer of coherence across the chapters that would have been much attenuated if the aperture onto the history of ideas had been wider, that is if our historical sweep had been more capacious.

The methodology and tools that are used in most of the case studies in Part II of this book were developed by the Cambridge Concept Lab, which I directed from 2014 to 2018. The Lab set out to test a number of

6 I Computational Methodologies for the History of Ideas

hypotheses regarding the structure and compositionality of concepts and, using computational techniques, built a suite of tools that enabled us to discern the computational ‘signatures’ of discrete concepts. Chapter 2 sets out in detail how the Lab proceeded, and introduces the various methods and measures we developed. Once we were confident that we had created a robust method for generating such signatures we began to work on a number of case studies that tracked conceptual forms over time. The Lab experimented with a number of datasets, including small-scale resources such as the North American News Text Corpus, but we directed most of our attention to the very large text database Eighteenth Century Collections Online, whose affordances and problems are also outlined in Chapter 2. This book enhances that work and assembles twelve chapters split across two parts, authored both collaboratively (as all of the work of the Concept Lab was) in slightly changing formations or teams, as well as some sole-authored contributions. The first part is methodological and is intended to introduce DH methods, in general, to researchers in the field of intellectual history, and to provide detailed explanations of the specific techniques used in most of the chapters in Part II. In Chapter 2 readers will be introduced to the field we call Distributional Concept Analysis and to the ways in which it might contribute to more traditional histories of ideas. Chapter 3, a sole-authored chapter by Paul Nulty, provides a detailed account of some of the tasks researchers are confronted with when designing DH projects. This chapter explains a number of the decisions that are fundamental to the construction of a specific, targeted project that seeks to enquire into the history of conceptual forms and ideas and is intended to provide an explanation of why, in designing the tools in the Concept Lab, we took the decisions we did. It should be noted, however, that this window onto the technical aspects of designing a DH project is also intended to be exemplary, not prescriptive, for future projects in the digital history of ideas.

Part II collects nine cases studies on topics such as the idea of government in the long eighteenth century, the idea of liberty in roughly the same time period, and the conceptual foundations of the modern idea of economic growth. The design of the book is intended to address both researchers in the history of ideas who might have very little experience working with digital methods and those who may have more familiarity, as we point forward from the chapters in Part I to those in Part II, where particular computational techniques are taken up. And, conversely, in Part II we refer back to chapters in Part I that provide the technical material for understanding how the methodology adopted has been formulated. Thus, armed with the methodological apparatus set out in the

first part of the book, readers from both camps will be able to follow the arguments in the second. It is also important to note that the selection of the topics or ideas taken in each case study has been determined in two ways – once again seeking to provide exemplifications of how more traditional approaches to the history of ideas might begin to work computationally. The first approach is to identify a ‘hard case’ – as lawyers call them – in the history of ideas in the long anglophone eighteenth century, such as the idea of liberty, the topic of Chapter 4. This approach works, as it were, from the analogue to the digital, and it is important to register the reason for taking it: we mean to test the findings of our digital methodologies against the substantial literature within the traditional history of ideas. The second approach works the other way around: in accord with what we deem to be best practice in DH work based on massive datasets, we did not frame a question or topic before we had inspected the dataset and the information derived from it in the form of relative values of distributional probability function (*dpf*, explained in Chapter 2); rather, we let the data drive the formulation of the questions. This is an important aspect of work using such techniques and approaches: it fully takes account of the transformation of the analogue text base into a digital format that both creates new information and at the same time provides different access to the data, the most obvious of which is the scale at which one can read the archive. In this second approach, then, we mean to test the arguments put forward by the traditional history of ideas against the findings derived from the interrogation of digitised textual materials. Here one might compare Chapters 5 and 9, which take the second approach, with Chapters 4 and 6, which take the first.

It will be noted that the chapters in the second part of the book intersect with long-standing arguments about and characterisations of these candidate ideas in hopes that they will speak with immediacy to researchers in the field. Chapter 4, for example, takes its point of departure from Isaiah Berlin’s famous account of liberty and asks whether his notion of a divergence between positive liberty – ‘freedom to’ – and negative – ‘freedom from’ – maps onto a digitally constructed view of the anglophone late seventeenth and eighteenth centuries. And in its last section it places its findings alongside one of the most powerful accounts of liberty in the Cambridge School tradition of the history of ideas that has been consistently explored by Quentin Skinner over the last forty or so years. Similarly, Chapter 6 on republicanism engages with the various arguments intellectual historians have proposed concerning the ideological bases for the American Revolution, and using digital methods concludes that the idea of republicanism could not have been one of the foundation stones for the construction of the first modern republic in what became

8 I Computational Methodologies for the History of Ideas

the United States of America. Here the connections and links to traditional ways of thinking about ‘ideas in context’ are used in order to test and interrogate the new digital methods employed in the creation of contexts that would be impossible to construct without computational methods. While the underlying mode of inquiry, seeking to recover the precise contexts in which ideas circulated and were minted in the past, remains the same, the main purpose is to explore whether or not the computational construction of context at a massive increase in scale provides us with new evidence and new avenues of inquiry, thereby complementing and extending traditional intellectual history as we develop a new digital history of ideas.

Some of the case studies take a very narrow focus, such as Chapter 8, by John Regan, on the idea of commercial society as it was used by the late Istvan Hont in his posthumously published Carlyle Lectures, or Chapter 7, which seeks to explain how the components of the contemporary idea of economic growth can be traced back to Adam Smith in ways that unsettle the standard picture of Smith’s modelling of political economy. Similarly, Chapter 9 unsettles a very long-standing characterisation of the period from around the mid-eighteenth century to the mid-nineteenth as an ‘age of sensibility’ by uncovering the conceptual networks that coalesced around the idea of irritability. Its authors, Ewan Jones and Natalie Roxburgh, note in conclusion that perhaps the time has come to refer to an ‘age of irritability’. Chapter 10, on the South Sea Bubble, explores the ways in which that event has become a kind of place-holder for the idea of unstable financial markets – a bubble that was inevitably going to burst – and demonstrates how at the time it was conceptualised very differently. Here Claire Wilkinson notes that a kind of back formation seems to have operated, and it has coloured the history of both the event and the idea of speculative bubbles for generations of historians, both economic and cultural.

Chapter 11 introduces some novel methods of data mining that use a set of techniques developed in the field of word embeddings. Although the methods used here are different, rather interestingly the computational objects produced resemble in some ways those produced by the Concept Lab methods. Here the idea of revolution in the sense of mechanical movement and in the sense of political upheaval is tracked across the eighteenth century as Mark Algee-Hewitt presents an account of the ways in which ideas evolve over time. Chapter 12, the final chapter, takes up the challenge of thinking about conceptual history over the long *durée*, as Ryan Heuser engages with the work produced within the school known as *Begriffsgeschichte*. He speculates that the period of time taken as a

focus for most of our case studies was itself crucial with respect to the entanglements of concepts – their varying rates of change and alteration in structure and form – as he asks if digital methods can make some headway into a question that Reinhart Koselleck thought to be intransigent, namely whether the rate of change in social and political concepts distinctively increased in the period he called the *Sattelzeit* – roughly, 1770–1830.

Although the book as a whole does not set out to promote a single line of argument – indeed the intention is to present interlocking but independent examples of the ways in which a small selection of DH methods might reinvigorate and contribute to more traditional histories of ideas – there are, nevertheless, some common themes and preoccupations. Each of the case studies, for example, explore in their own ways the utility of modelling concepts and ideas as distinct but overlapping modes of thinking or cognition. And, in their different ways, they all seek to illustrate and interrogate the methodologies proposed by this book, which use computational and statistical techniques for establishing a more fine-grained account of the linguistic contexts in which concepts scale up into ideas. This can also be understood in a number of other ways, as, for example, the movement from concreteness to abstraction (as Chapter 7 remarks with respect to ‘improvement’), or from the local to the general (as, for example, the singular local event of the crash of the speculative market in South Sea stock became the idea of speculative markets and their tendency to ‘burst’ in general). The same type of transformation is mapped in Chapter 5, in which we explore how the concept of government, used as a way of thinking about a *type or form* of government, ‘scales up’ as it came to be used as a way of thinking about government in general as a kind of universal category intended to make the hard distinction between order (government) and chaos (anarchy, despotism). And in Chapter 11 the intriguing hypothesis emerges that the political idea of revolution was created by the same forces of conceptual upscaling as the term transformed from its technical, mechanical sense into an abstract idea whose precise components (the necessary and sufficient conditions for something to be conceived as a political revolution) remained unarticulated.

This is to note that each case study identifies a curious phenomenon in conceptual and ideational behaviour whereby the scale at which one might think is significantly increased, or to put that another way, the utility of thinking in general categories coupled with the discovery of what might be called epistemic fundamentals becomes very apparent. Our case studies, then, track a kind of transformation in epistemology, noting that within the period to which we direct most consistent

10 I Computational Methodologies for the History of Ideas

attention – say from the mid-eighteenth century into the early decades of the nineteenth – it became both possible and desirable to think on a new scale, to ground human knowledge in what were thought to be the irreducible elements of human understanding. To some extent this might be considered to be a very rehearsed story of the aims and ambitions of Enlightenment inquiry, or a confirmation of Koselleck's notion of the *Sattelzeit*, but if it does indeed correspond to those long-standing historicisations of the European project in knowledge – which we know under the clumsy rubric of 'the Enlightenment' – we mean to make our argument at a level of granularity that only becomes possible using digital methods.

Thus, one might note the following similar patterns across the chapters: John Regan notes that the idea of commercial society was scaled up from its early formulation as a *segment* of the social in which commercial activity took place, to a characterisation of society in general: modern societies in the age of expansion were reconceived as through and through commercial. In Chapter 9 a similar alteration in scale is outlined by Ewan Jones and Natalie Roxburgh in which something that was understood to occur to an individual body – irritable nerves – was scaled up to aid understanding of a general category, the human. In Chapter 7 we outline the ways in which our modern idea of economic growth has its deep ideational roots in the entanglements of 'growth' and 'improvement', which led Adam Smith to experiment with the notion that 'growth' might become autotelic, that is in and for itself, as distinct from its application to something. In Chapter 11 Mark Algee-Hewitt shows how the concept of revolution looked both ways across the century, towards the specifics of mechanical forces and to the ideational explosions that were associated with political upheaval. Here the time frame becomes extremely interesting as we begin to note differences in the temporal envelopes of what we might think of as the event-time of conceptual and ideational change. Some of the transformations we explore across the book have very impacted or explosive event horizons: the publication of Whytt's essay on the involuntary motion of animals marked a decisive intervention into the conceptualisation of the human. Others are more like slow burns, such as 'culture', as outlined by Ryan Heuser in Chapter 11, or 'liberty', as tracked by Paul Nulty in Chapter 3. In these cases, a set of affordances lie deep within a concept's architecture, surfacing and resurfacing over time before eventually settling into one or another stable form, or coming to be replaced by a new configuration and articulation better suited to the contexts in which it is to be used.

These observations of themes and continuities across the book are intended to provoke and encourage further exploration. There are numerous ways in which such suggestive lines of inquiry might be extended and subjected to rigorous evaluation. That is for future work, which we hope will become fully conversant with and contribute to a digital history of ideas that has the potential to transform the discipline of intellectual history.