



DIGITAL Enlightenment STUDIES

VOLTAIRE FOUNDATION
UNIVERSITY OF OXFORD

Review

De Bolla P. (ed.) 2023. *Explorations in the Digital History of Ideas: New Methods and Computational Approaches*. Cambridge: CUP.

Glenn Roe, University of Oxford, glenn.roe@humanities.ox.ac.uk

Roe G. 'Review: De Bolla P. (ed.) 2023. *Explorations in the Digital History of Ideas: New Methods and Computational Approaches*'. In: *Digital Enlightenment Studies* 3, 53–62.

DOI: <https://doi.org/10.61147/des.66>

1. A digital history of ideas?

Over the past half-century, intellectual history has been shaped by a set of methodological traditions that continue to define its central questions and evidentiary practices. Most influential among these have been the contextualist programme associated with the Cambridge School and the conceptual-historical approach developed within the German tradition of *Begriffsgeschichte* (Brett 2002). Despite their significant differences, both frameworks have shared a commitment to reconstructing the historical conditions under which ideas become meaningful, whether through the recovery of authorial intention within specific linguistic contexts or through the diachronic tracing of conceptual vocabularies across political and social transformations (Palti 2024). These approaches have generated an extraordinarily rich body of scholarship, particularly for the study of the early modern and Enlightenment periods, but they have also largely developed within interpretative frameworks premised upon close reading and relatively bounded textual corpora.

The rapid expansion of large-scale digitised archives over the past two decades has introduced new possibilities, and new challenges, for intellectual historians. Commercial resources such as Early English Books Online (EEBO) and Eighteenth-Century Collections Online (ECCO) have dramatically increased the accessibility of early modern print culture, enabling scholars to survey bodies of text at scales previously



unimaginable.¹ Yet the availability of such corpora has not automatically produced a corresponding transformation in historical method (Guldi and Armitage 2014). While digital search technologies have facilitated discovery and citation, they have more rarely been integrated into sustained methodological reflection about how ideas circulate, stabilise and transform within large cultural systems (Edelstein 2016). It is within this emerging methodological space that *Explorations in the Digital History of Ideas: New Methods and Computational Approaches*, edited by Peter de Bolla, and with contributions by Mark Algee-Hewitt, Ryan Heuser, Ewan Jones, Paul Nalty, Gabriel Recchia, John Regan, Nathalie Roxburgh and Claire Wilkinson, positions itself.

The volume represents a concerted attempt to articulate what a digitally informed history of ideas might look like when computational analysis is treated not simply as a tool of efficiency, but as a means of reframing historical explanation. Bringing together a group of scholars associated with the Cambridge Concept Lab, the book proposes that large-scale textual analysis can illuminate conceptual dynamics that remain difficult to capture through traditional text-centred approaches.² Rather than reconstructing intellectual lineages primarily through canonical authors or individual works, the contributors seek to model concepts as distributed cultural formations that emerge through patterned linguistic association across collections of texts.

The Enlightenment provides both the historical focus and the methodological proving ground for these experiments. As a period characterised by rapid conceptual innovation, expanding print networks and the consolidation of many categories central to modern political and social thought, the long eighteenth century offers a uniquely demanding environment in which to test claims about conceptual change at scale (Koselleck 1988). The central question posed by this volume, and one that guides the present review, is therefore not simply whether computational methods can extend the reach of intellectual history, but whether they can fundamentally reshape how historians understand the emergence, circulation and transformation of ideas within eighteenth-century print culture.

2. Ambition and architecture

Explorations in the Digital History of Ideas is structured into two main sections that collectively seek to move from methodological reflection to historical demonstration. Part I assembles a series of chapters devoted to computational approaches to conceptual history, focusing on the development, justification and critical examination of what the contributors term 'distributional concept analysis' (DCA). These chapters address the theoretical status of concepts, the operational choices required to model conceptual relationships

¹ On EEBO and ECCO, see <https://about.proquest.com/en/products-services/eebo/> and <https://www.gale.com/primary-sources/eighteenth-century-collections-online> respectively.

² On the Concept Lab, see <https://www.crash.cam.ac.uk/research/projects-centres/the-concept-lab-cambridge-centre-for-digital-knowledge/> and De Bolla 2013.



computationally and the interpretative challenges posed by large-scale aggregated text collections. Rather than presenting a single unified methodological manifesto, the volume offers a sequence of interrelated interventions that progressively elaborate both the possibilities and the limits of modelling conceptual structures through quantitative text analysis.

Part two turns from methodological exposition to a set of historically grounded case studies drawn primarily from anglophone political and intellectual discourse between the seventeenth and early nineteenth centuries. These chapters explore the conceptual trajectories of terms such as liberty, republicanism and political revolution, using large corpora derived principally from EEBO and ECCO. Together, they aim to demonstrate how distributional approaches can generate new historical insights by tracing conceptual associations across broad textual environments rather than within individual works alone.

While the chapters share a common set of corpora, analytical vocabulary and methodological commitments, the volume stops short of imposing a rigidly uniform research paradigm across all contributions. Instead, the case studies vary in both computational technique and historiographical emphasis, reflecting the experimental character of the broader enterprise. This diversity allows the book to showcase the adaptability of digital approaches to different kinds of conceptual problems, but it also introduces a degree of methodological heterogeneity that shapes the volume's overall argumentative rhythm.

Given the number and range of case studies included, the present review focuses on a selection of chapters that most clearly illustrate the volume's methodological ambitions and historiographical stakes, while referring more briefly to others in order to situate the collection as a whole.

3. Distributional concept analysis

The methodological centre of gravity of the volume lies in the development of 'distributional concept analysis' (DCA) as a concept, explained most fully in Chapter 2. The chapter advances a deliberately expansive claim about the nature of concepts and about the kinds of historical evidence capable of capturing conceptual change. Rejecting the straightforward identification of concepts with either individual words or stable semantic meanings, the authors instead define concepts as functional and relational structures that become visible through patterned lexical associations across large textual corpora. In this framework, conceptual history shifts from tracking the evolution of discrete terms to reconstructing networks of linguistic co-occurrence that reveal how ideas operate within broader discursive environments.

To articulate this shift, the chapter proposes a three-level analytical distinction between words, concepts and ideas. Words function as the material instantiations found within individual texts, while concepts represent recurring relational clusters that structure discourse across multiple texts and contexts. Ideas, in turn, emerge as higher-order argumentative formations composed of interacting conceptual components. By positioning concepts as an intermediate analytical layer between lexical usage and ideological argument, DCA seeks to mediate between two influential historiographical traditions. From the Cambridge School,



it retains a sensitivity to context and discursive function; from *Begriffsgeschichte*, it adopts the premise that conceptual vocabularies can be traced across extended historical durations. The resulting framework does not seek to adjudicate between these traditions but to provide a new empirical scale at which their questions might be addressed.

A defining feature of DCA is its explicit prioritisation of interpretability over computational optimisation. The authors deliberately distance their approach from black-box machine-learning models, such as neural word embeddings or topic modelling techniques, the internal operations of which often remain opaque to scrutiny. Instead, DCA relies on a transparent statistical measure designed to identify significant lexical associations while remaining traceable to the underlying textual evidence. This methodological choice reflects a broader historiographical commitment: computational results must remain open to verification, contestation and reinterpretation through conventional scholarly argument.

The chapter's application of DCA draws heavily on the ECCO corpus, treated (not unproblematically) as a proxy for anglophone eighteenth-century print culture. This aggregation enables the reconstruction of conceptual environments across thousands of documents, allowing historians to observe patterns of conceptual stability and transformation that would be difficult to detect through selective close reading alone. At the same time, the chapter acknowledges familiar limitations associated with large-scale historical corpora, including uneven genre representation, duplication of texts and the distortions introduced by optical character recognition (OCR). These caveats are not presented as fatal obstacles but as interpretative conditions that must be incorporated into historical analysis.

The principal historiographical tension introduced by DCA concerns the relationship between distributional patterns and historical circulation. The method assumes that repeated lexical association across a corpus reflects the cultural availability of a concept, rather than merely coincidental linguistic proximity. While this assumption enables the reconstruction of conceptual fields at unprecedented scale, it also raises questions about how far aggregated textual patterns can stand in for the heterogeneous communicative environments in which Enlightenment ideas circulated. Pamphlets, periodicals, translations and polemical texts all contributed to eighteenth-century conceptual debates in ways that may resist uniform statistical aggregation. The chapter recognises the need to complement computational analysis with close reading, but the precise boundary between functional conceptual structure and historically-situated meaning remains necessarily fluid and somewhat underdefined.

4. Operationalising conceptual structure

If Chapter 2 establishes distributional concept analysis as a viable methodological framework, Chapter 3 subjects that framework to sustained epistemological scrutiny. Rather than introducing new computational techniques, the chapter interrogates the assumptions embedded within any attempt to measure conceptual relationships quantitatively. Central to this argument is the claim that categories such as association,



similarity and relatedness are not natural linguistic properties but the result of specific operational decisions made during model construction. The choice of statistical metric, contextual window size or threshold for conceptual significance, the chapter argues, is inseparable from broader theoretical assumptions about language, culture and historical change.

Drawing on the operationalist tradition associated with Percy W. Bridgman and later adapted within digital literary studies, the chapter reframes computational modelling as a form of historiographical interpretation (Moretti 2014). To define a concept through the operations used to measure it is to acknowledge that methodological design inevitably encodes theoretical commitments. This reframing represents one of the volume's most significant contributions, encouraging historians to treat computational parameters as explicit elements of scholarly argument.

Particular attention is given to the distinction between paradigmatic and syntagmatic forms of association. Paradigmatic association, grounded in substitutability or semantic similarity, captures relationships between terms that can occupy comparable conceptual roles. Syntagmatic association, by contrast, reflects patterns of co-presence within shared discursive contexts. The chapter demonstrates how different statistical approaches privilege one form of association over the other, thereby generating divergent representations of conceptual structure. For intellectual historians, this distinction has direct methodological implications, since it determines whether computational models emphasise semantic equivalence or discursive articulation.

The chapter also provides a sustained reflection on the epistemic status of visualisation. Tables, network graphs and spatial embeddings are presented both as tools for presenting data and as modelling choices that impose distinct assumptions about conceptual organisation. Spatial visualisations, for example, suggest metric proximity between concepts, potentially smoothing over polysemy or internal conceptual conflict, while network representations foreground relational connectivity but often require thresholding that discretises continuous linguistic variation. By foregrounding these representational constraints, the chapter encourages readers to approach visual outputs as interpretative constructions rather than transparent reflections of historical reality.

A particularly significant theoretical intervention lies in the chapter's distinction between mental and cultural conceptual structures. Whereas many computational linguistic models aim to approximate cognitive representations of meaning, the authors argue that intellectual history must instead model concepts as cultural artefacts shaped by institutions, publication practices and collective discursive habits. This shift opens the possibility of designing computational parameters grounded in historically specific communicative environments rather than in general theories of cognition. At the same time, the chapter leaves largely unresolved how such historically situated constraints might be operationalised in practice, pointing towards a productive area for future methodological development.

Taken together, the methodological chapters establish the intellectual seriousness of the volume's computational ambitions. Their most significant achievement lies in demonstrating that digital methods



do not bypass theoretical reflection but instead relocate it within the design of analytical procedures and statistical measures. At the same time, the very sophistication of these methodological reflections highlights a persistent challenge: the need to translate computational modelling into historically grounded interpretation. This tension, already visible in Part I, becomes the central testing ground for the case studies that follow.

5. Liberty, 1600–1800

Chapter 4, devoted to the conceptual history of liberty, stands as the volume's most persuasive demonstration of the historical payoff of distributional concept analysis. Drawing on both the EEBO and ECCO corpora, the chapter re-examines one of the most extensively studied concepts in early modern political thought, engaging directly with influential interpretative frameworks associated with Isaiah Berlin and Quentin Skinner. Rather than adjudicating these debates through the close reading of canonical authors, the chapter reframes them as questions about the circulation and distribution of conceptual associations across anglophone print culture as a whole.

The chapter proceeds through a carefully staged analytical sequence that mirrors the methodological principles articulated in Part I. Initial frequency comparisons establish that 'liberty' and 'freedom', often treated as interchangeable terms in modern political discourse, occupied distinct conceptual environments during much of the seventeenth and early eighteenth centuries. Liberty appears predominantly in constructions emphasising capacity or entitlement, while freedom demonstrates a more variable distribution across both positive and negative articulations. The analysis then moves to distributional modelling, identifying stable clusters of bound lexis surrounding liberty, including associations with slavery, servitude and bondage, before undergoing a significant late-century transformation in which rights, government, republicanism and despotism become increasingly prominent.

The chapter's most striking historiographical claim lies in its identification of a relatively abrupt conceptual reconfiguration in the final decades of the eighteenth century. According to this account, liberty undergoes a shift from an attribute primarily attached to persons toward a concept increasingly associated with forms of political order. Republican discourse emerges not as an ideological rival to liberalism, but as a catalytic framework through which liberal conceptions of liberty are reformulated. By modelling these shifts through aggregated lexical association, the chapter challenges narratives of gradual conceptual evolution, suggesting instead that conceptual transformation may occur through punctuated phases of rapid reorganisation across broad discursive environments.

As a methodological demonstration, the chapter is highly effective in showing how large-scale textual evidence can reveal patterns of conceptual stability and mutation that remain difficult to capture through canonical texts alone. At the same time, its conclusions rely heavily on the assumption that distributional frequency reflects cultural circulation. While this identification enables analysis at an unprecedented scale,



it also raises questions about how heterogeneous eighteenth-century communicative contexts – ranging from pamphlet debates to periodical commentary and translated political theory – are flattened within aggregated corpora. The chapter acknowledges the politics of archival exclusion, particularly in its discussion of slavery's conceptual relationship to liberty, but the broader implications of such exclusions for conceptual modelling remain only partially explored. Even so, the chapter provides the clearest illustration of the explanatory potential of digital intellectual history and serves as the benchmark against which the volume's other case studies can be evaluated.

6. Republicanism and the founding of America

Chapter 6 offers the volume's most overtly revisionist contribution, recasting a canonical historiographical question concerning the ideological foundations of the American Revolution as a problem of lexical and conceptual availability. Beginning with the observation that the term 'republicanism' appears only rarely within eighteenth-century anglophone print culture, the chapter advances the provocative claim that the concept may not have functioned as a coherent motivating ideology for the founding generation in the form later attributed to it.

Rather than treating this lexical absence as a methodological obstacle, the chapter uses distributional analysis to reconstruct the conceptual environments of related terms, including republic, republican and republicans. The results reveal a gradual shift over the course of the century from associations centred on political actors and partisan identities towards a vocabulary increasingly orientated around institutional forms of governance, constitutional order and political rights. From this perspective, republicanism emerges less as a stable ideological programme than as a retrospective abstraction drawing together diverse political practices that were initially articulated in more disparate conceptual terms.

The chapter's argument has significant historiographical implications, challenging influential narratives that position republican ideology as a coherent intellectual framework guiding revolutionary political action. By foregrounding conceptual improvisation and retrospective ideological stabilisation, the chapter demonstrates how computational analysis can destabilise established interpretative models. At the same time, the argument also exposes the limits of distributional inference. The identification of lexical rarity with conceptual absence raises the broader question of whether political concepts can circulate through rhetorical, performative or institutional practices without acquiring stable terminological expression. While the chapter acknowledges the existence of non-print political discourse, its conclusions remain grounded primarily in textual corpora, leaving open the possibility that certain forms of conceptual articulation fall outside the evidentiary reach of computational analysis.

Placed alongside the liberty case study, the chapter highlights the uneven temporality of conceptual development during the eighteenth century. Whereas the idea of liberty undergoes rapid and measurable transformation, republicanism appears to resist lexical consolidation until relatively late in the century.



This contrast underscores one of the volume's broader insights: digital methods may be particularly well suited to tracing asymmetrical conceptual trajectories, including the emergence, mutation and even absence of historically consequential ideas.

7. Embedded ideas: political revolution

Chapter 11 expands the methodological scope of the volume by addressing the formation of compound concepts, focusing on the emergence of political revolution as a historically transformative idea. Recognising the limitations of collocation-based approaches that rely on single focal terms, the chapter introduces diachronic word-embedding models aligned across successive temporal intervals. These models allow conceptual change to be visualised as movement within multidimensional semantic space, enabling analysis of how distinct conceptual domains converge over time.

The chapter's central empirical finding concerns the asymmetric convergence between the terms 'revolution' and 'political' during the eighteenth century. Rather than moving directly towards political discourse, 'revolution' is shown to pass through the conceptual terrain of Newtonian mechanics and scientific theory, where it becomes associated with cyclical motion, systemic transformation and law-governed change. Only in the later decades of the century does 'revolution' re-enter explicitly political discourse, now conceptualised as a systematic process rather than as episodic upheaval. The chapter argues that this scientific mediation played a crucial role in making political revolution intelligible as a theoretical and practical category, contributing to the simultaneous emergence of political science as a disciplinary field.

Methodologically, the chapter demonstrates the interpretative possibilities opened by embedding-based approaches, which allow historians to model conceptual convergence, divergence and mediation beyond simple lexical adjacency. At the same time, the interpretative burden placed on geometric distance within semantic space raises important historiographical questions. The chapter provides compelling visualisations of conceptual movement but is less explicit about how historians should translate spatial proximity into historically meaningful explanation. Moreover, by emphasising convergence and conceptual compounding, the analysis risks smoothing over the rhetorical instability and ideological contestation that frequently characterised eighteenth-century revolutionary discourse (Edelstein 2025).

Despite these limitations, the chapter offers one of the volume's most innovative demonstrations of how computational methods can capture the emergence of complex conceptual formations that do not correspond neatly to individual lexical units. When read alongside the chapters on liberty and republicanism, it highlights the methodological diversity of digital intellectual history, showing how different computational techniques can illuminate distinct forms of conceptual change, from mutation and absence to mediated convergence.



8. Conclusion

Taken as a whole, *Explorations in the Digital History of Ideas* succeeds in demonstrating that computational textual analysis can produce genuinely new explanatory perspectives for intellectual history. Its most significant contribution lies in its reconceptualisation of ideas as distributed cultural formations, the development of which can be traced through patterned linguistic association across large textual environments. By modelling conceptual change as a process characterised by stability, mutation, absence and convergence, the volume moves beyond frequency-based histories of ideas toward a more structurally orientated understanding of conceptual transformation. The case studies examined here illustrate how such approaches can generate historically meaningful interventions in longstanding debates, particularly in the analysis of political vocabulary central to Enlightenment discourse.

Equally important is the volume's insistence that digital methods must remain theoretically self-conscious. The contributors repeatedly emphasise that computational modelling requires theoretical modelling assumptions to be built into the design of statistical measures, corpus construction and visual representation but that this does not displace interpretation. In doing so, the book raises the methodological bar for digital intellectual history, demonstrating that large-scale textual analysis can operate as a form of historiographical argument rather than as an auxiliary research technique. This insistence on transparency and interpretability represents one of the volume's most enduring strengths, particularly at a moment when increasingly sophisticated machine-learning and AI-based approaches risk obscuring the relationship between computational output and historical evidence.

At the same time, the volume highlights several challenges that remain unresolved within the current practice of computational conceptual history. Most prominently, the reliance on aggregated corpora such as ECCO inevitably introduces tensions between scale and historical specificity. While large datasets allow historians to reconstruct broad patterns of conceptual association, they also tend to flatten distinctions between genres, rhetorical contexts and forms of textual mediation that have long been central to Enlightenment scholarship. Multilingual conceptual circulation, translation practices and the uneven geography of print publics similarly remain difficult to incorporate within models that treat corpora as relatively unified cultural environments.

These limitations do not undermine the value of the volume's methodological innovations, but they do suggest that future work in digital intellectual history will need to engage more directly with the material and institutional conditions that shape conceptual transmission. If the book occasionally risks treating distributional patterns as proxies for cultural circulation, this tendency is less a methodological failure than an indication of what the next stage of research development must be. Indeed, the volume's greatest achievement may lie in clarifying both the interpretative possibilities opened up by computational modelling and the historical complexities that such modelling has yet fully to accommodate.



Read in the context of rapidly expanding generative language technologies, the volume's methodological commitments acquire additional significance. While large language models can simulate historically informed discourse with remarkable fluency, they typically prioritise predictive plausibility over interpretative transparency. The approaches developed in this collection therefore offer a complementary model for digital scholarship, one that emphasises operational clarity, evidentiary traceability and historical accountability. The future of digital intellectual history will likely depend less on replacing traditional interpretative practices than on integrating computational modelling into historically informed analytical frameworks.

By demonstrating how concepts can be studied as dynamic cultural formations while maintaining methodological self-awareness, *Explorations in the Digital History of Ideas* establishes an important foundation for future research. If the digital history of ideas is to mature as a field, it will do so most productively by continuing to treat the Enlightenment not simply as a repository of textual data, but as a demanding conceptual terrain through which new historical and methodological questions can be articulated.

References

- Brett A.** 2002. 'What is intellectual history now?'. In: Cannadine D. (ed.) *What is History Now?* London: Palgrave Macmillan, 113–31.
- De Bolla P.** 2013. *The Architecture of Concepts: The Historical Formation of Human Rights*. New York: Fordham University Press.
- Edelstein D.** 2016. 'Intellectual history and digital humanities'. In: *Modern Intellectual History* 13:1, 237–46.
- Edelstein D.** 2025. *The Revolution to Come: A History of an Idea from Thucydides to Lenin*. Princeton: Princeton University Press.
- Guldi J. and Armitage D.** 2014. *The History Manifesto*. Cambridge MA: Cambridge University Press.
- Koselleck R.** 1988. *Critique and Crisis: Enlightenment and the Pathogenesis of Modern Society*. Cambridge: MIT Press.
- Moretti F.** 2014. "'Operationalizing": or, the function of measurement in modern literary theory'. *Stanford Literary Lab Pamphlet* 6. <https://litlab.stanford.edu/assets/pdf/LiteraryLabPamphlet6.pdf>.
- Palti E. J.** 2024. *Intellectual History and the Problem of Conceptual Change*. Cambridge: Cambridge University Press.