21 Textual Analysis and Conceptual Cartography

Charles H. Pence

Overview: At first blush, it might seem as though digital approaches could provide us with precisely the kind of input we need to perform something like conceptual analysis in the philosophy of science: querying the expressed intuitions of the "folk" (here, practicing scientists publishing in the journal literature) to see how they put various concepts to use, to which cases they believe they can be applied, and so on. In this chapter, I want to nuance this argument, both by clarifying what we might mean by "conceptual analysis" in this case and by tempering expectations about what digital approaches could be reasonably expected to give us. I claim that such a more moderate goal, which I'll call here "conceptual cartography," can still provide the philosophy of science with a number of advantages (which are otherwise difficult to attain), while avoiding the possibility of making promises that we can't fulfill.

21.1 Introduction

Digital philosophy (or, even more broadly, the digital humanities in general) is an impossibly broad subject. It could mean the creation of new digital philosophical content—for instance, preparing digital editions of manuscripts, or digital exhibitions on the life, correspondence, and works of philosophers. It might be about the use of computer modeling to represent processes of philosophical interest, like the development or spread of scientific knowledge (Mohseni, chapter 23, this volume). It could involve mapping the connections between documents or authors by way of citation analysis, revealing the network structure that binds them together (see Herfield & Doehne, chapter 22, this volume).

But one particularly appealing use might be in what Franco Moretti famously baptized "distant reading," attempting to use digital tools to do the same kind of extraction of semantic content from a corpus of text that a competent human reader normally takes themselves to be doing when they do philosophy (Moretti 2003, 2004a, 2004b). Surely, the reasoning would go, we can use these digital tools to look at large corpora of texts and, in turn, learn something about how scientists have used concepts within their published writings—approaching, as I will soon describe it, something like the much maligned "conceptual analysis" (see Koskinen & Alexandrova, chapter 5, this volume). Indeed, this was my port of entry into the digital humanities: for almost a decade and a half, I built and maintained a database of some two million scientific journal articles, largely in the biological sciences (Ramsey & Pence, 2016; Pence, 2016). I largely did so in the hopes of this kind of distant reading. What do biologists mean when they talk about fitness? How do they conceive of natural selection? With enough biological papers and enough ingenuity, the answers to these kinds of questions, I thought, would be forthcoming.

This chapter is a story of my disappointment. As it turns out, I will argue, this is precisely *not* what we should expect from the tools of the digital humanities. But that is not to say that they do not provide us with something useful. What I want to do here is pursue the causes of this failure, and describe what kind of (still, clearly, important and useful!) knowledge about the scientific process we might hope to obtain from the corpus analysis of scientific journal articles—knowledge that I will describe below as *conceptual cartography*. If you are more careful about your goals than I was, perhaps I can spare you from my frustration.[1]

21.2 Concepts in Philosophy of Science

Philosophers of science often take one of our essential duties to be the pursuit of a better understanding of the concepts used by scientists. The vague notion of "pursuing understanding" can be cashed out in a variety of ways: following Carnap, we might call it "explication" (Brun, 2015), or we could pick from any number of other choices, including (at least) the elaboration, clarification, systematization, or "tracing out" of science's conceptual commitments.

However exactly we might want to describe it, Ernan McMullin (1970) does not, I think, offer a particularly unorthodox opinion when he writes that, if we want to assess "the principal ways of understanding the patterns that occur in the complex of activities we call scientific research," there are at least two broad categories of approaches (pp. 51–52). The first is a psychological description of the behavior of scientists as they do science. Such a description might, at times, be essential to our understanding of a given episode.

But a distinctly *philosophical* approach, by contrast, would look quite different—even if we were to keep a particularly historical bent in undertaking it (in line with McMullin's broader commitment to the integrated history and philosophy of science):

> We can trace the gradual modification of a concept (like the concept of ether in Newton's thought), where it is possible to give plausible conceptual grounds ("reasons") for the modification's having occurred the way it did. The "pattern" here is a relation between ideas or can somehow be associated with such a relation. The techniques are those of conceptual analysis. (McMullin, 1970, p. 50)

As I just mentioned, I don't think that this appeal to, in at least some sense, "conceptual analysis" is particularly controversial. But fleshing out such an account still provides us with an interesting challenge. How exactly do we do this kind of work? To what kinds of tools or methodologies can we help ourselves in the process? How are the conclusions that we then draw to be justified? What is the relationship between this sort of work and scientific practice more broadly? One very quickly arrives not just at questions about the nature of the philosophy of science, but about the methodology and justification of philosophy as a whole.

It is here that one can see, I think, the reasons that I became initially enamored of digital tools. In much the same way that a metaphysician might appeal to "the public's" intuitions about time or motion, or an epistemologist might probe which cases "we" are willing to call knowledge, a philosopher of science can consult "the public"—in our case, practicing scientists—and their sincerely held beliefs, expressed at least in part in their journal articles (though see Lean et al., 2021).[2]

But this raises exactly the same questions that I posed above for conceptual analysis more broadly. How could computational approaches help us to engage with the kind of conceptual questions that McMullin gestures at above? When would we expect them to succeed and to fail? Can they offer us a complement to other, traditional philosophical methods, and in which circumstances? To start to answer these questions, we first need to get clear on how exactly we are supposed to clarify our understanding of these scientific concepts in the first place.

21.3 From Concepts to "Conceptual Analysis"

The very term "conceptual analysis" invites confusion. As Robert Hanna (1998) notes, it is sometimes used to refer to a particular historical philosophical movement, one that arose in England and was especially prominent in the middle of the twentieth century. But it can also, more generally, refer to something like a philosophical methodology, in which a search for propositions that define or connect concepts (such sentences often taken to be necessarily or analytically true) is the foundation of philosophical method.

Both the historical movement and the contemporary methodology are not often well regarded by contemporary philosophers; Karen Neander (1991) describes the term as having "plummeted in the popularity stakes" (p. 168). Historically speaking, the "movement" of conceptual analysis suffered from the pointed attacks on intensions and on analyticity launched by W.V.O. Quine (1951), and it was largely abandoned by the 1970s. And the methodology, to take only the critique perhaps most pertinent for our purposes in the philosophy of science, was seriously weakened by the arguments of Hilary Putnam, who noted that if concepts, especially scientific concepts, correspond to natural kind terms, then we have good reason to think that conceptual analyses are at best contingently true, as the behavior of natural kinds in our world is not a matter of metaphysical necessity (Putnam, 1973).

A number of philosophers have, nonetheless, attempted to develop versions of conceptual analysis that can escape these objections. I'll try to synthesize a few of their commonalities here, drawing largely from Frank Jackson (1998). Whether or not we believe, with Quine, that analytic truth is a meaningless notion, or, with Putnam, that "'meanings' just ain't in the *head*" (Putnam, 1973, p. 704), something like conceptual analysis seems nonetheless to play an important role in philosophical study. In Jackson's words, it "is the very business of addressing when and whether a story told in one vocabulary is made true by one told in some allegedly more fundamental vocabulary" (Jackson, 1998, p. 28).

An abbreviated argument for this claim goes something like the following. When we do metaphysics (and, I would urge, here we have significant overlap with conceptual projects in the philosophy of science), we regularly ask questions about whether some kind *K* exists, whether it is anything over and above some other kind *J*, or whether the behavior of the *K*s is fully determined by the behavior of the *J*s (Jackson, 1998, p. 31). The way in which we often do that is to query the nature of what Jackson calls the "ordinary conception" of *K*s and *J*s. What do competent users of those terms take them to be like? How do they make those judgments? And how do they judge what we might call "corner cases," where the application of the term is unclear? We answer these questions, in turn, by presenting competent users of those kind terms with various possible cases, to see whether the ordinary conception countenances them or no.[3]

Of course, there are a host of qualifications and nuances to this quick argument, which is just to say that this kind of querying needs to be done extremely carefully. (Who will be asked about their intuitions—put differently, who is "the public" or "the linguistic community" of relevance? When should we accept their uses of a concept at face value, and when should we consider them to be problematically uninformed?) But the fact remains that this method, familiar enough from a variety of philosophical enterprises, just *is* doing conceptual analysis: "Extracting the cases that count as *K*s from a person's response to possible cases" (Jackson, 1998, p. 36), or, on Neander's definition, "trying to describe the criteria of application that the members of the linguistic community generally have (implicitly or explicitly) when they use the term" (Neander, 1991, p. 170).

I'll set aside, for my purposes, further argumentation in favor of conceptual analysis, taking it as writ that this is something that philosophers of science at least sometimes engage in, and at least sometimes successfully. Before moving on, however, it will be important for what comes next in the argument—the relationship between conceptual analysis and digital philosophy—to note the role of *language* here. Thus far, there's been a bit of slippage (acknowledged by many of conceptual analysis's defenders, including both Jackson and Neander) between *concepts*, as the language-independent entities that we want to analyze, and the *terms* that we use to refer to those concepts. In Jackson's words,

> Our subject is really the elucidation of the possible situations covered by the *words* we use to ask our questions—concerning free action, knowledge, and the relation between the physical and the psychological, or whatever. I use the word 'concept' partly in deference to the traditional terminology which talks of *conceptual* analysis, and partly to emphasize that though our subject is the elucidation of the various situations covered by bits of language according to one or another language user, or by the folk in general, it is divorced from considerations local to any particular language. (Jackson, 1998, p. 33)

This is thus the sense of conceptual analysis with which I want to work in what follows. At least part of the effort of philosophers of science involves the attempt to see just what concepts scientists have in mind when they use particular terms. Of course, there are manifold potential objections and traps here, for which we must be cautious. But as Neander notes, the characteristics of scientific language might make these less pressing in the case of the philosophy of science. For most scientific concepts, "the relevant linguistic community consists of specialists, and the term under analysis is one of their specialist terms, and [it] is also abstract (nonperceptual) and embedded in well-articulated theory" (Neander, 1991, p. 171). Each of these features makes it more likely that the rigorized language of science would be fruitful terrain for conceptual analysis and will be less likely to run afoul of the objections raised against it.

21.4 Digital Conceptual Analysis?

As I've already gestured at several times, one could easily think that the tools of digital philosophy will be especially useful here. Our "public," practicing scientists, happens to have the habit of carefully expressing their considered opinions about the nature of scientific concepts—at least implicitly, and sometimes even explicitly—in their published writings. Surely, then, a replacement for at least some of the work that might otherwise be done by conceptual analysis, in the sense of the term that I laid out above, could be performed by analyzing the writings of scientists with an eye toward extracting their conceptual commitments. Eventually, I'm going to argue against this claim. But to make the case as robustly as I can, I'll start by presenting some particularly promising examples from one kind of textual analysis—topic modeling—that might seem to do just that.

Topic modeling is by now a well-understood method in digital humanities, and one of the more common ones currently used in digital philosophy. Perhaps most prominent among these uses has been the historical analysis of the field. Christophe Malaterre and colleagues, for instance, have extensively analyzed a corpus of a number of philosophy of science journals, drawing out a number of publication trends across the eighty-year history of dedicated journals in our field (Malaterre & Lareau, 2022; Malaterre et al., 2019). Sometimes this kind of work, more properly described as HOPOS (history of the philosophy of science) study (about which see Lydia Patton's chapter in this volume, chapter 26), might have immediate conceptual upshot (e.g., in rethinking the importance of the concept of "organism" for our understanding of the early history of the philosophy of biology, see Nicholson & Gawne, 2015). But in general, this isn't its target.

Using the same tools, however, we can easily imagine work that is. Indulgently, I'll make use of an example from my own work (Lean & Pence, under review). Imagine a scientific concept that has been the target of significant philosophical attention, like *specificity*. Specificity sits at the nexus of two further, extremely complex biological concepts: *causation* and *information*, describing something like the extent to which a cause is "particular" to its effect (contrasted with nonspecific causes, which participate in the production of a wide variety of effects). It seems like scientists regularly talk about some causes as being more or less specific than others, and they use this notion (or something very much like it) in an argument for preferring one biological cause over other causes of the same phenomenon as the target of their analysis, as has been argued by James Woodward (2010). Such a notion then has been used as grist for various philosophical mills (see, e.g., Waters, 2007; Sarkar, 2005, on the relationship between development and genetics). Importantly, others have argued that there are other notions of "specificity" also at work across biology—like the "lock-and-key" or "one-to-one" notion common in discussions of enzymes (Bourrat, 2019), or the concept of "binding specificity" crucial in biochemistry (Lean, 2020).

But it is just here that we might ask ourselves: What *do the practicing scientists actually mean* when they use this term? How could we engage in a scientifically focused conceptual analysis of specificity? Here might be one way. Let's return to topic modeling. What exactly does topic modeling consist in? There are many different answers to this question, but perhaps the easiest to grasp is to consider it as proposing an entirely fictional model of document creation and then deducing the parameters of that model that best fit our corpus. Imagine that the process of writing went like this: We create a collection of *k* topics for our corpus. Each one of those topics is a probability distribution over all words, which describes the chance that a given word might be chosen in the context of that topic. (If we're talking about evolution, say, we probably have a higher chance of using the word "species" than the word "gunboat.") We then create a new distribution over the topics for our document: this tells us how likely each *topic* is to appear in that document. (If we're writing a paper on evolution, topics about genetics or biochemistry might also be mixed in; probably less so topics about nuclear chemistry.) We then imagine "writing" documents by sampling from these probability distributions. If a document has *n* words, then, *n* times, we first sample from the latter distribution to tell us which topic that word will come from, and then we sample from that topic's distribution to tell us which word from that topic will come next.

Once we've assumed this (obviously nonsensical) method for document construction, we can ask a computer to derive these probability distributions (one each for every article and one each for each of the *k* topics). Technically speaking, there are a number of algorithms for doing so. The most common is known as Latent Dirichlet Allocation, or LDA (Blei, 2012). Practically speaking, though, there are a number of tools that can make the creation of such models easier. Programmers proficient in Python can write quick scripts using the well-known Gensim library (Řehůřek and Sojka, 2010); similar packages are available for R (Jockers, 2014). Those looking for a more out-of-the-box solution can use software such as MALLET (McCallum, 2002). In all cases, after getting our textual data into a form that the computer can understand, models are "trained"—usually for a collection of various values of *k*, and then compared.

But let's pause for a second. Why would we want to spend all this effort to build models of "document writing," if we have represented that process by such an obviously fictitious abstraction? What makes topic modeling so useful is that, as it turns out, those *k* topic probability distributions that result very often *make sense to us*. Consider, for example, what happens when we look at the top ten words that a topic picks out as most probable (remembering that a topic is nothing more than a probability distribution over all the words in the corpus). We often note that these most probable words have a kind of thematic coherence—that they seem to describe, that is, one of

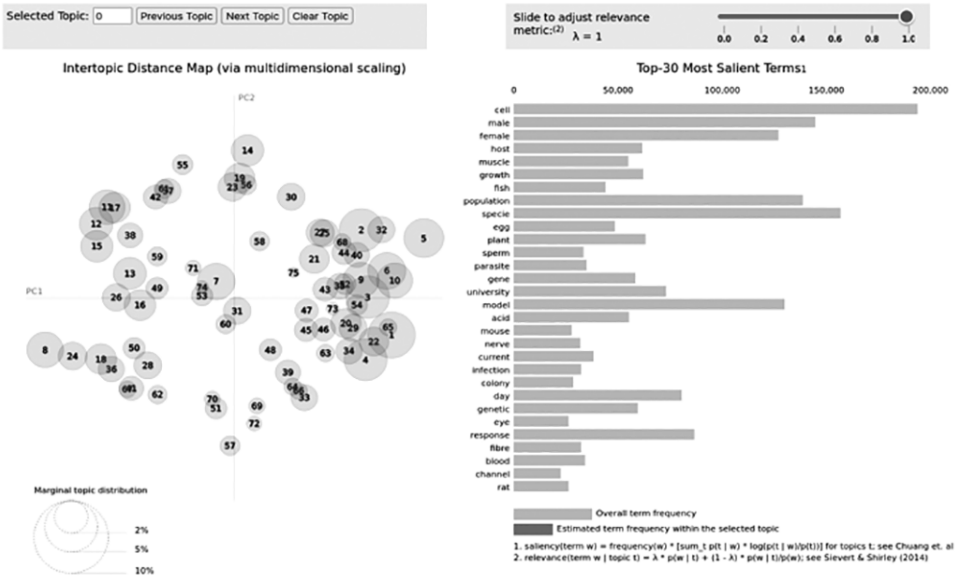

Figure 21.1

The output of pyLDAvis for our topic model used to explore "specificity". On the left, a map of the relative positions in a kind of "semantic space" of the various topics, and on the right, a way to explore which words in each topic are the most probable.

the *topics* (now in the colloquial rather than the technical sense) that the documents in the corpus might be about. For instance, in the topic model that we prepared for our work on specificity, one topic's most probable words are "mouse," "cell," "immune," "response," "antigen," "antibody," and "dose." It doesn't require too much of an inductive leap here to think that this is a topic about mouse-based studies of immunology. More interesting visualizations can also be explored; figure 21.1 shows the output from the pyLDAvis software (a Python port of the original LDAvis package for R, see Sievert & Shirley, 2014), which includes a kind of relational map of topics along with a view for exploring their most probable as well as most "salient" terms—that is, terms that are probable in a given topic as well as improbable in other topics.

As is the case with any digital-humanities model, there are ways in which these analyses might fail. Consider, for instance, just the question of the choice of the value of *k*, the number of topics. The computer will happily produce a topic model corresponding to any value of *k* (if we have enough processing power, memory, and time)—that is, the computer will never tell us that we have chosen a bad parameter value. But if we try to divide a corpus into *too many* topics, we'll find that many of them are just

Table 21.1
The top twenty most probable words in each of the six topics identified within our corpus as being important to "specificity."

Topic	Most probable terms
MACROMOLECULE	protein, virus, cell, DNA, RNA, acid, enzyme, figure, sequence, amino, viral, activity, contain, peptide, show, gene, chain, synthesis, specific, gel
IMMUNOLOGY	mouse, cell, immune, response, antigen, antibody, dose, lymphocyte, immunity, specific, animal, anti, strain, tolerance, day, test, spleen, system, figure, control
MUTUALISM	plant, aphid, fungus, nodule, symbiont, fungal, host, root, gall, bacteria, soil, specie, symbiosis, symbiotic, cell, bacterial, tick, figure, pea, insect
PARASITISM	host, parasite, infection, pathogen, egg, virulence, infect, transmission, cuckoo, parasitism, immune, parasitic, strain, infected, disease, specie, figure, system, rate, soc
FETAL IMMUNITY	serum, rabbit, antibody, day, protein, pig, rat, sera, globulin, foetal, anti, sac, titre, embryo, yolk, foetus, guinea, table, result, antiserum
STRUCTURE	molecule, compound, group, enzyme, cytochrome, reaction, bond, hydrogen, atom, structure, bind, chain, chem, complex, band, form, spectrum, haemoglobin, protein, iron

duplicates, or that they start to pick out what we call "jargon" or "junk" topics—topics without any apparent semantic content that simply group together function words like "that" or "with." Meanwhile, if we don't divide the topic model finely enough, we can wind up with topics that cover too much, and thus become nearly impossible to interpret. While there are automated tools that can aid us in making these kinds of choices, such as quantitative measures of topic "coherence" (Röder et al., 2015), we should nonetheless never lose sight of the fact that many such choices are, and must be, subjective—the data do not speak for themselves, and the place of human judgment and decision-making is no smaller here than it is in "traditional" philosophical pursuits.

Now, to look at "specificity" in this model, we should ask which topics (of the seventy-five that were included in our model) have a comparatively high probability value for the word "specificity." As it turns out, there are six, which we identified with the short monikers MACROMOLECULE, IMMUNOLOGY, MUTUALISM, PARASITISM, FETAL IMMUNITY, and STRUCTURE. The top twenty most probable words for each of these topics are printed in table 21.1, which can already give a hint of the justification for our having chosen the labels that we did, as well as a feeling for what it is like to classify and offer explanations for the "content" of a topic. Further information can also be gleaned from the

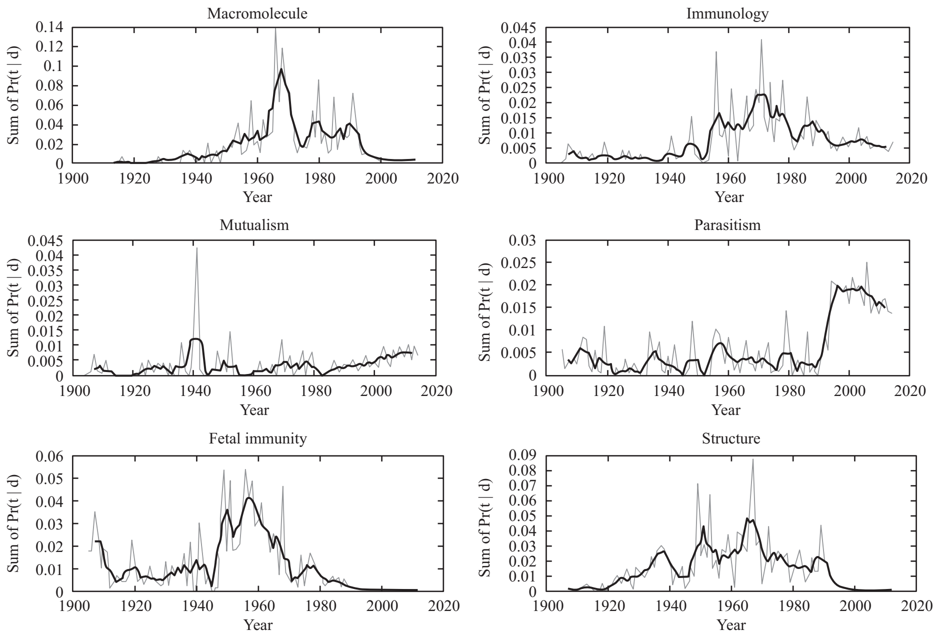

Figure 21.2
The prevalence of each of the six topics discussed above within the corpus over time. (Note that the vertical scales are different for each of the six subplots.) Thicker black lines are a five-year moving average; thin gray lines behind are the raw data.

graphs of those topics over time, as presented in figure 21.2—we see the rise and fall of various topics, as papers within the corpus make greater or lesser reference to them in different periods.[4] (I'll return to an analysis of this figure below.)

I don't have the space here to go into the full detail that would be required to explain our interpretation of each of these topics and their putative importance for our understanding of causal specificity. But I think we've seen enough to understand the superficial appeal of topic modeling for conceptual analysis: it's not hard to think that "specificity" as it shows up in IMMUNOLOGY is an immunological, likely one-to-one sense of specificity, while "specificity" in the topic MACROMOLECULE (which is identified by words like "protein," "virus," "DNA," and "RNA") is specificity in an informational, genetic sense.

But is this apparent interpretation really justified? There are a few reasons to be worried. First, recall that every document is *a probabilistic mixture* of these different topics. How should we react, then, if we see that a document has a nonzero probability for two or three of these topics that feature “specificity?” Any choice (multiple definitions at once? a hybrid definition? one definition being used beyond its usual range?) seems to require the close reading of numerous articles, a task that for large corpora will likely be somewhere between prohibitive and impossible.

Second, even if we can pick out, say, a group of documents that contain the word “specificity” and only make reference to one of the “specificity” topics, moving from here to a genuine *analysis* of the concept that can speak to the kinds of philosophical concerns that were introduced above is a tendentious inference. We assume, to be sure, that the “immunological” sense of specificity has something to do with lock-and-key, one-to-one type specificity, as this is the same way in which biologists (usually) talk about the relationship between antibodies and their targets. But as every philosopher of science well knows, this is at best tacit knowledge, not directly present within their journal articles, and at worst will amount to a question about which the practitioners have never thought and might have no clear opinion. As Hanna Lucia Worliczek describes in her chapter in this volume (chapter 24), for example, a term like “descriptive research” might become less popular or even disappear without any change in practice, if trends surrounding the “popularity” of various methodologies drive changes in article content.[5] Ascribing a *conceptual analysis* of the notion of specificity to the mere presence or absence of these topics within documents seems suspect.

Lastly, and while this takes me beyond my example (as this wasn’t a problem that we encountered in this case), these topic models still sometimes pose serious interpretive challenges. In recent work analyzing a corpus of literature in taxonomy (Pence & Conix, under review), we found that topic models often produced topics that simply picked out the organism of study (topics for butterflies, for fish, and so on)—and they did so by seizing on peculiar words that make reference to organismic anatomy. Unfortunately, these words are often used in different senses in different organisms (leading to an occasionally confusing merging of topics); even when they are used univocally, understanding the topics that result may require specialist knowledge in anatomy that is beyond the reach of most philosophers of science. In the former case, an interpretation of a concept within a confused topic might be impossible; in the latter case it might, once again, require more close reading than is feasible.

At the same time, it seems as though we nonetheless learn *something* about the concept of specificity from this topic model. We now understand in which subfields of biology it seems to be used (at least within the limitations of our corpus). We also know something about the other terms around which it is often to be found; if J. R. Firth (himself following the late Wittgenstein) is right that “you shall know a word by the company it keeps!” (Firth, 1957, p. 11), then we have likely also learned *something* of the meaning of the concept, though what exactly remains to be seen. How can we best understand what's happening here?

21.5 Toward Conceptual Cartography

In the rest of this chapter, I want to argue that the kind of attenuated knowledge about concepts that we get from digital philosophy of science is perhaps best understood as what I will here call *conceptual cartography*. Briefly and vaguely (defects that I'll remedy in a moment), what I think we get from turning to digital philosophy is more like relational information about concepts: both their links with one another and their links with other, nearby concepts in philosophy and in science. Just as in "proper" cartography, such a map is useful insofar as it gives us an unexpected view into our source material, making those relationships more perspicuous than they would have otherwise been (think here, for instance, of Harry Beck's famed map of the London Underground—see also Brunet, chapter 9, this volume).

The idea of "conceptual cartography" is not foreign to the philosophical literature, but it is not common enough to have yet received a standardized definition. For Robert Smithson, it is "the project of determining the necessity and contingency of the various features of our conceptual schemes" (Smithson, 2021, p. 98); for instance, whether or not a given concept must be primitive or derived from other, more fundamental concepts (like the relationship between "green," "blue," and "grue"). For Nathaniel Goldberg, the notion is marshaled in a defense of the study of the history of philosophy. Drawing on Gilbert Ryle's use of the terms "geography" and "cartography," he argues that "conceptual cartography is the practice of mapping how concepts generally (including philosophical views) relate conceptually (including logically and extralogically)" (Goldberg, 2017, p. 123), where here he has in mind things like determining the extent to which various philosophical views of empirical properties ground them in subjective and objective features of knowers.

My aim here is a sense more similar to the latter than the former. It is all too clear to philosophers of science who explore scientific practice that terms like—to draw only from my own areas of expertise—"natural selection," or "genetic drift," or "fitness" are used by scientists in a host of superficially dissimilar ways. It's quite possible that this superficial difference actually hides various kinds of underlying similarity. It's possible that it doesn't. It's possible that this similarity correlates with various other features of that scientific practice—again, to continue the example, with organisms of interest, with sub-fields of study, with scientific "schools," with geographic location, and so on. Again, it's possible that it doesn't. In my sense, conceptual cartography is the mapping of this kind of landscape: an understanding of the relationships between and features of scientific concepts as those concepts are expressed within scientific practice. And it is conceptual cartography in this sense, I think, for which digital-humanities methods are a particularly powerful tool.

Let's return to our example and to figure 21.2 above. When we set the six topics in which "specificity" is found against one another, and explore how their prominence unfolds over time, we see something like a coherent "map" emerge: specificity explodes as a topic of scientific discussion between 1940 and 1970 in lockstep with the similar explosion of molecular biology; discussions are dominated by the topic having to do with biological macromolecules and the translation of DNA and RNA into proteins. As these have become less important topics in contemporary molecular biology (moving outward, one might speculate, into more dedicated biochemical journals), the concept is now dominated by its use in parasitism (with some contribution from immunology).

This map doesn't tell us what specificity *means*, and in that sense it cannot be a conceptual analysis. But it does tell us *where specificity lives*, and in that sense the analogy with mapping seems apt. If we want to understand the concept, we now know where we need to look—what other biological and philosophical concepts we need to understand in order to situate it. Some of these, of course, are expected (e.g., DNA to RNA to proteins), but some of them might be less anticipated (e.g., immunology) and have interesting and important philosophical consequences (e.g., with Pradeu 2012, concepts of the self and identity).

There are still cautions and objections for which we need to look out, and open questions that remain to be solved. Let's look at four examples. Goldberg, in exploring the use of conceptual cartography for the history of philosophy, raises a number of objections. One of these, in particular, applies extremely well to the case of the philosophy of science. He writes:

> Using the history of philosophy *productively* in conceptual cartography requires choosing complex concepts, or views, selecting relevant features, and systematically transforming those features into markers in conceptual space. [. . . In his example, this entails] ignoring the rest of [numerous philosophers'] views on metaphysics (and much else), and ignoring other views altogether—which, one might object, does all such views injustice. (Goldberg, 2017, p. 134)

There is a general response to this worry, which Goldberg notes: all maps are inherently selective, and as we learn from Borges's (1998) fable of the "Map of the Empire whose size was that of the Empire, and coincided point for point with it" (p. 325), any map must make these kinds of simplifications.

But we can add a more subtle point in the case of digital philosophy. The kinds of representational simplifications that are required in order to transform the full scientific texts with which we begin into digital representations like the outputs of a topic model *just are*, I claim, the changes that Goldberg has in mind as involved in "systematically transforming those features into markers in conceptual space" (p. 134). Put differently, one remains, of course, free to dispute the utility of conceptual cartography as an enterprise. But to do so would be the very same thing as disputing the utility of these methods as a whole.

A second objection concerns what we might call the "success conditions" for such a conceptual cartography. Even if the method now is the subject of some disrepute, we at least had some idea what a good conceptual analysis would look like: it would capture the important features of the intuitive or folk use of a concept and help us understand which further aspects were "just details." What are the analogous conditions for cartographies? Of course, a minimal bar is easy to establish: if, say, specificity is *not* actually used in immunology, our cartography had better not say that it is. But a map has a much more pragmatic orientation than a conceptual analysis. We are not looking for "the one true map"—a contradiction in terms in any case—but rather we are making something like a *promise of utility*: this map will prove useful if only you'll give it a try. (Harry Beck's London Tube map was rejected by his bosses in 1931, before later finding wide public acclaim.)

Evaluating whether or not we've created a good map, then, will inherently be a diachronic, pragmatic affair. As Hasok Chang has described such a verification process (here, for science as a whole in his pragmatist philosophy of science), we are looking for something like the "operational coherence" of our cartography with other scientific and philosophical affairs, the way in which we make "elements of our activities fit together harmoniously so that our aims may be achieved" (Chang, 2022, p. 4). This opens up any would-be conceptual cartographer to many of the same objections that might be raised against a pragmatist theory of truth (see, e.g., Chang, 2022, pp. 197–203). But I don't see how such a conclusion can be avoided, given this pragmatic orientation.

Finally, there are two open questions to be raised concerning philosophy's relationship to other, nearby disciplines. More negatively, we might worry that philosophers, in pursuing this kind of cartographic work, are simply becoming poor amateur sociologists of science.[6] After all, these are well-established disciplines with extensive methodological canons. Trespassing on their territory without sufficient training is certain to end badly. But I think we can avoid much of the force of this objection by keeping our focus on the *conceptual* aspect of conceptual cartography. After all, we are not here focused on what have often become the stock in trade of sociologists (or historians) of science: networks of collaboration, or training, or institutions, or financing, and the like (see Kusch, chapter 20, this volume). We remain, as philosophers of science, focused on science's conceptual content.[7] And since, as we've just discussed, any map will include choices about which aspects of the terrain to foreground and which to background, ours will be the choices of the philosopher, not those of the sociologist.

More positively, I think we have more work to do in exploring the relationship between conceptual cartographies and other areas of the philosophy of science. If nothing else, conceptual cartography is, and avowedly must be, a largely descriptive affair.[8] How, then, should it be related to enterprises like the stipulative definition of scientific concepts (Neander, 1991), or, more recently, conceptual engineering (Isaac et al., 2022; Burgess et al., 2020; Haslanger, 2000)? It is equal parts clear and facile to note that in order to re-engineer a concept, it helps to start by understanding where it is to be found and to which other concepts it might be related. I suspect that making that observation more rigorous is likely to quickly become quite difficult. But unless we want to entirely renounce normative interventions on scientific practice, I think such clarification will be necessary.

21.6 Conclusion

Those enamored, as I am and remain, with the possibilities of digital philosophy of science still have much work to do. In addition to offering more examples of completed, successful, useful digital projects, the relationship between the work that we might like to do and other philosophical projects is, for the moment, relatively unclear. I have argued here that, despite some superficial similarities, we should avoid the trap of thinking that digital tools can directly provide us with something like conceptual analyses. On the contrary, they offer us a weaker, but no less interesting, sort of approach to scientific concepts in use, something that I've baptized here "conceptual cartography."

Much other work in digital philosophy, of course, will not pursue conceptual cartography. But insofar as one of digital philosophy's biggest advantages, I think, is in offering us access to methods and materials that were previously unimaginable, conceptual cartography offers us a promising use of such tools. Fully mapping a concept—beyond the handful of well-known citations for an idea to which philosophers would traditionally appeal—is, I think, just this kind of application. With luck, it will allow us to build better situated, more sophisticated philosophical understandings of scientific practice.

Notes

1. It is interesting that in recent years, similar meta-level discourse has also broken out surrounding the merits and demerits of experimental philosophy, which might similarly be somewhat too alluring at first blush (Shepherd & Justus, 2015; Koch, 2019; Kozlov, chapter 19, this volume).

2. The caveat and citation here are important. As is a repeated theme throughout this volume (see Ankeny & Leonelli, chapter 17, Worliczek, chapter 24, Currie, chapter 10, Patton, chapter 26, and Walsh & Pulkkinen, chapter 25), the nature of scientific articles—in particular, the *evidentiary value* of scientific journal articles—is an extremely complex and difficult question. I have argued (with Oliver Lean and Luca Rivelli) in the paper cited above that philosophers need to have clear and forthright answers to these kinds of questions—especially to questions about the role of journal articles in scientific publishing and the nature of the semantic claims that we can derive from syntax—in order to most fruitfully use digital tools. I'll thus send the reader to that article if they are interested in my take on those questions.

3. As Jackson notes, this thin description doesn't even clearly distinguish philosophers from other fields where researchers engage in effectively identical methodologies: think of psychologists trying to understand how children develop cognitively, or political theorists trying to understand how voting behavior works, or linguists investigating how some particular dialect differs from its neighbors.

4. The usefulness of conceptual cartography in a temporal sense is an important aspect to which I cannot dedicate enough space here. The conceptual maps that digital philosophy gives us—because texts are inherently temporal objects—is a four-dimensional map, on which time plays an equal role to other aspects of conceptual relationships.

5. While I lack the space to discuss its impact in full detail, see also Ankeny and Leonelli's chapter in this volume, chapter 17, which considers the impact of nontraditional publication formats in philosophy, another challenge to document-centric conceptual analysis.

6. Thanks to Sophie Veigl and Adrian Currie, with whom I worked out the substance of this objection.

7. Rightly, I think; at least someone should be.

8. A subtle kind of normativity might arise via the kinds of "free choices" that we make in the construction of those cartographies, but this will not be normative in nearly the same sense or with nearly the same force as the other parts of the philosophy of science I discuss in this paragraph.

References

Ankeny, R. A., & Leonelli, S. (2025). Research design for philosophy of science in practice. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Borges, J. L. (1998). On exactitude in science. In A. Hurley (Ed.), *Collected Fictions* (p. 325). Penguin Books.

Bourrat, P. (2019). Variation of information as a measure of one-to-one causal specificity. *European Journal for Philosophy of Science*, *9*,11. https://doi.org/10.1007/s13194-018-0224-6

Brun, G. (2015). Explication as a method of conceptual re-engineering. *Erkenntnis*, *81*, 1211–1241. https://doi.org/10.1007/s10670-015-9791-5

Brunet, T. (2025). Corresponding domains: Advice on applying mathematics in the philosophy of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Burgess, A., Cappelen, H., & Plunkett, D. (Eds.). (2020). *Conceptual engineering and conceptual ethics*. Oxford University Press.

Chang, H. (2022). *Realism for realistic people: A new pragmatist philosophy of science*. Cambridge University Press.

Currie, A. (2025). Thinking with case studies in the philosophy of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. In J. R. Firth (Ed.), *Studies in linguistic analysis* (pp. 1–32). Blackwell.

Goldberg, N. (2017). History of philosophy and conceptual cartography. *Analytic Philosophy*, *58*(2), 119–138. https://doi.org/10.1111/phib.12097

Hanna, R. (1998). Conceptual analysis. In E. Craig (Ed.), *Routledge encyclopedia of philosophy*. Routledge.

Haslanger, S. (2000). Gender and race: (What) are they? (What) do we want them to be? *Noûs*, *34*(1), 31–55. https://doi.org/10.1111/0029-4624.00201

Herfeld, C., & Döhne, M. (2025). Empirical network analysis as a method in philosophy of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Isaac, M. G., Koch, S., & Nefdt, Ryan. (2022). Conceptual engineering: A road map to practice. *Philosophy Compass*, *17*, e12879. https://doi.org/10.1111/phc3.12879

Jackson, F. (1998). *From metaphysics to ethics: A defence of conceptual analysis*. Clarendon Press.

Jockers, M. L. (2014). *Text analysis with R for students of literature*. Springer.

Koskinen, I., & Alexandrova, A. (2025). Conceptual analysis plus. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Koch, S. (2019). Carnapian explications, experimental philosophy, and fruitful concepts. *Inquiry*, *62*(6), 700–717. https://doi.org/10.1080/0020174X.2019.1567381

Kozlov, A. (2025). Experiments in philosophy of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Kusch, M. (2025). Integrated philosophy and sociology of science: The example of the sociology of scientific knowledge. In S. Veigl & A. Currie (Eds.). *Philosophy of science—A user's guide*. MIT Press.

Lean, O. M. (2020). Binding specificity and causal selection in drug design. *Philosophy of Science*, *87*(1), 70–90. https://doi.org/10.1086/706093

Lean, O. M., & Pence, C. H. (under review). Exploring specificity in the biological literature with topic models.

Lean, O. M., Rivelli, L., & Pence, C. H. (2021). Digital literature analysis for empirical philosophy of science. *British Journal for the Philosophy of Science, 74*(4), 875–898. https://doi.org/10.1086/715049

Malaterre, C., Chartier, J. -F., & Pulizzotto, D. (2019). What is this thing called *Philosophy of Science*? A computational topic-modeling perspective. 1934–2015. *HOPOS*, *9*(2), 215–249. https://doi.org/10.1086/704372

Malaterre, C., & Lareau, F. (2022). The early days of contemporary philosophy of science: Novel insights from machine translation and topic-modeling of non-parallel multilingual corpora. *Synthese*, *200*(3), 242. https://doi.org/10.1007/s11229-022-03722-x

McCallum, A. K. (2002). MALLET: Machine learning for language toolkit. http://mallet.cs.umass.edu/

McMullin, E. (1970). The history and philosophy of science: A taxonomy. In R. H. Stuewer (Ed.), *Historical & philosophical perspectives of science* (pp. 12–67). Minnesota Studies in the Philosophy of Science (Vol. 5). University of Minnesota Press.

Mohseni, A. (2025). Methods for modelers of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Moretti, F. (2003). Graphs, maps, trees: Abstract models for literary history—1. *New Left Review*, *24*, 67–93.

Moretti, F. (2004a). Graphs, maps, trees: Abstract models for literary history—2. *New Left Review*, *26*, 79–103.

Moretti, F. (2004b). Graphs, maps, trees: Abstract models for literary history—3. *New Left Review*, *28*, 43–63.

Neander, K. (1991). Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science*, *58*(2), 168–184. https://doi.org/10.1086/289610

Nicholson, D. J., & Gawne, R. (2015). Neither logical empiricism nor vitalism, but organicism: What the philosophy of biology was. *History and Philosophy of the Life Sciences*, *37*(4), 345–381. https://doi.org/10.1007/s40656-015-0085-7

Patton, L. (2025). History of philosophy of science. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Pence, C. H. (2016). RLetters: A web-based application for text analysis of journal articles. *PLOS ONE*, *11*(1), e0146004. https://doi.org/10.1371/journal.pone.0146004

Pence, C. H., & Conix, S. (under review). Mapping controversy: A cartography of taxonomy and biodiversity for the philosophy of biology. To appear in *Advances in Experimental Philosophy and Corpus Methods*.

Pradeu, T. (2012). *The limits of the self: Immunology and personal identity*. Oxford University Press.

Putnam, H. (1973). Meaning and reference. *Journal of Philosophy*, *70*(19), 699–711. https://doi.org/10.2307/2025079

Quine, W.V.O. (1951). Two dogmas of empiricism. *Philosophical Review*, *60*(1), 20–43. https://doi.org/10.2307/2181906

Ramsey, G., & Pence, C. H. (2016). evoText: A new tool for analyzing the biological sciences. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *57*, 83–87. https://doi.org/10.1016/j.shpsc.2016.04.003

Řehůřek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of LREC 2010 Workshop: New Challenges for NLP Frameworks* (pp. 45–50). University of Malta. http://is.muni.cz/publication/884893/en

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* (pp. 399–408). Shanghai: WSDM 2015: Eighth ACM International Conference on Web Search; Data Mining; ACM. https://doi.org/10.1145/2684822.2685324

Sarkar, S. (2005). How genes encode information for phenotypic traits. In *molecular models of life: Philosophical papers on molecular biology*, 261–283. MIT Press.

Shepherd, J., & Justus, J. (2015). X-Phi and Carnapian explication. *Erkenntnis*, *80*, 381–402. https://doi.org/10.1007/s10670-014-9648-3

Sievert, C., & Shirley, K. E. (2014). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, (pp. 63–70). Baltimore: Association for Computational Linguistics. https://doi.org/10.3115/v1/w14-3110

Smithson, R. (2021). Conceptual cartography. *Inquiry*, *64*(1–2), 97–122. https://doi.org/10.1080/0020174X.2020.1779805

Walsh, K., & Pulkkinen, K. (2025). Philosophy of science in the archive. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.

Waters, C. K. (2007). Causes that make a difference. *Journal of Philosophy*, *104*(11), 551–579. https://doi.org/10.5840/jphil2007104111

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318. https://doi.org/10.1007/s10539-010-9200-z

Worliczek, H. (2025). Historicizing epistemology—Historiographic practices for the uninitiated. In S. Veigl & A. Currie (Eds.), *Philosophy of science—A user's guide*. MIT Press.