What would Indy do? Resisting post-truth through the practice of annotation

Book Section

How to cite:

For guidance on citations see FAQs.

© 2017 The Author

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
What Would Indy Do? 
Resisting Post-Truth through the Practice of Annotation

Elton Barker

Entering the Commons
Among the many acts of resistance that welcomed the formal investiture of the 45th president of the United States was self-proclaimed white nationalist Richard Spencer, in Washington to celebrate the inauguration of his president, getting punched in the face. The video of the assault, posted on Twitter, subsequently went viral, as images of famous American heroes were shared with the hashtag #punchaNazi. One of the most popular was a still from Stephen Spielberg’s Indiana Jones and the Raiders of the Lost Ark, when Indy lands a knockout blow to one of his Nazi adversaries in the back of a truck (Duggan 2017).

Normally mentioning Indiana Jones in the context of archaeological work is enough to prompt weary sighs from colleagues, who are accustomed to having to patiently explain that being an archaeologist is not quite as glamorous as it is depicted in these films. But my purpose for invoking Indy’s name here has to do less with affirming (or denying) his archaeological credentials as with his battle against the fascist appropriation of cultural heritage. Along with giving encouragement to white supremacists, Donald Trump’s election has ushered in—and, to a certain extent, is clearly a manifestation of—a culture where expertise and informed opinion are to be rigorously denied, far less coveted. Two of the favorite phrases that Trump used during his presidential campaign, and that have subsequently come to define his administration, are “fake news” and “alternative facts.” When the challenge to expertise has gone so far that it is not only viewed skeptically but there is an active intolerance of it—when, indeed, the very idea of an epistemological truth or demonstrable fact is denied—then democratic culture is at stake.¹ As experts in our various disciplines, scholars need to

¹ Attacks on the sophists, who stand accused in much of the literature that survives from the democratic culture of fifth-century BCE Athens of peddling ways of making the weaker argument the stronger, reflect similar anxieties about the threat to democracy: see the evidence collected in Kerferd 1981.
join the resistance and counter fascist ideology with reasoned argument and empirical observation.²

Important here are questions of both infrastructure and agency: we need the web to be as open as possible in terms of its structure and delivery;³ but we also need to be as open as possible in our scholarship, with the ability, at the very least, to be able to point to the evidence on which we base our arguments and assertions.⁴ In the language of research communication, our data need to be FAIR: findable, accessible, interoperable, and reusable (Table 1). Or, to put that slightly differently, how can we enable people to discover our work, gain access to it, and (re)use it, particularly when academic publishing regards open access as a threat to its business model?

As researchers who (literally) dig into the past to better understand who we are and where we come from, archaeologists stand on the frontline of the battle. Many indeed have already embraced the challenge of putting their scholarship and research “into the commons.”⁵ Three indicative examples serve to illustrate the kind of work that is engaged in directly informing the public, as well as moving toward a new kind of scholarship: Open Context provides a process and a service for publishing research data (openly) on the web, preserving those data with leading digital libraries, and providing the means to analyze them through maps and other media (opencontext.org); ODATE, or The Open Digital Archaeology Textbook Environment, is an initiative to facilitate a collaborative practice in publishing, by enabling researchers to contribute their expertise to writing an archaeology textbook within an open publishing framework (o-date.github.io/); the Twitter hashtag #dayofarch invites archaeologists from all around the world to share their experience of an ordinary day working in the field, the results of

---

² See, for example, the pushback against the white supremacist appropriation of classical material on social media from some US classicists, in particular: Bond 2017, Futo-Kennedy 2017, and Zuckerberg 2017, all of whom have been predictably trolled for being women or liberals or both. The UK has not been immune from such controversies either, as evidenced by the debate over a BBC school video that featured a family in Roman Britain in which the father, a high-ranking soldier, was depicted as non-white (Beard 2017). On accountability and transparency as fundamental procedures of fifth-century BCE Athens, see Vernant ([1962] 1982, 51–52).

³ As the debates around net neutrality have made clear: e.g., https://www.freepress.net/issues/free-open-internet/net-neutrality.

⁴ In relation to the “painted statues” and “ethnic Roman Britain” stories, Neville Morley (2017) suggests that “what is surprising about these two arguments is that the substantive issues—ancient statues were painted, the Roman Empire (including Britain) was ethnically diverse—are such old hat,” and calls for “explanation, helping people move beyond the myths and half-truths they may have acquired at school or from the media, to understand how historians (and scientists) actually go about investigating and reconstructing the past.”

⁵ Marcel Detienne (1996, 95) explains the expression “into the middle” (es meson) as to set things in common, whether the referent is material goods (as when Achilles puts “into the middle” booty to be competed over in Homer’s Iliad, e.g., 23.507), or arguments and opinions (in political debate). Leslie Kurke (1999, 29) similarly describes Herodotus’s Histories as an “open agora [or market-place] of logoi [accounts],” which allow different voices to be heard; cf. Dewald 1987. On the use of the expression “es meson” in representations of debate in classical Greek literature: Barker 2009, 17–18 and the Index, s.v. “meson, es.”
which are then brought together for hosting on a common website (dayofarchaeology.com). Such examples could be multiplied; the point is that archaeologists are leading the way in the effort to do their research in the public domain.

Within this context of enabling research methodologies and findings to be heard in the community’s “contest of voices” (Goldhill 1986, on fifth-century BCE Athenian tragedy), I want to focus on one small, but critical, element of digital practice: semantic annotation. The idea of annotation—of adding notes to a document—goes back at least as far as medieval manuscripts, out of which the vast majority of our classical corpus of Greek and Latin literature is reconstructed (Barker and Terras 2016). Figure 1, for example, shows the first page of the Venetus A manuscript of Homer’s *Iliad*, complete with scholarly annotations both in the margins and inline. Such annotations provide further information about a particular idea or element in the text to enrich its interpretation, often by pointing to information external to the text itself. Once frowned upon as a practice by the custodians of a printed book culture (viz., librarians), the annotation is making something of a comeback in the digital medium, whereby metadata (some

<table>
<thead>
<tr>
<th>Concept</th>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>To be Findable</td>
<td>F1. (Meta)data are assigned a globally unique and eternally persistent identifier.</td>
</tr>
<tr>
<td></td>
<td>F2. Data are described with rich metadata.</td>
</tr>
<tr>
<td></td>
<td>F3. (Meta)data are registered or indexed in a searchable resource.</td>
</tr>
<tr>
<td></td>
<td>F4. Metadata specify the data identifier.</td>
</tr>
<tr>
<td>To be Accessible</td>
<td>A1. (Meta)data are retrievable by their identifier using a standardized communications protocol.</td>
</tr>
<tr>
<td></td>
<td>A1.1. The protocol is open, free, and universally implementable.</td>
</tr>
<tr>
<td></td>
<td>A1.2. The protocol allows for an authentication and authorization procedure, where necessary.</td>
</tr>
<tr>
<td></td>
<td>A2. Metadata are accessible, even when the data are no longer available.</td>
</tr>
<tr>
<td>To be Interoperable</td>
<td>I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</td>
</tr>
<tr>
<td></td>
<td>I2. (Meta)data use vocabularies that follow FAIR principles.</td>
</tr>
<tr>
<td></td>
<td>I3. (Meta)data include qualified references to other (meta)data.</td>
</tr>
<tr>
<td>To be Re-Usable</td>
<td>R1. Meta(data) have a plurality of accurate and relevant attributes.</td>
</tr>
<tr>
<td></td>
<td>R1.1. (Meta)data are released with a clear and accessible data usage license.</td>
</tr>
<tr>
<td></td>
<td>R1.2. (Meta)data are associated with their provenance.</td>
</tr>
<tr>
<td></td>
<td>R1.3. (Meta)data meet domain-relevant community standards.</td>
</tr>
</tbody>
</table>

Table 1. FAIR principles. Source: Adopted from FORCE11 (2017).
Fig. 1. Marcianus Graecus Z.454 (= 822) (Venetus A), folio 12 recto. Homer Iliad Book 1, lines 1–1.25. http://www.homermultitext.org/urn:cite:hmt:vaimg:VA012RN-0013. This image was derived from an original (©2007, Biblioteca Nazionale Marciana, Venice, Italy). The derivative image is ©2010, Center for Hellenic Studies. Original and derivative are licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 License. The CHS/Marciana Imaging Project was directed by David Jacobs of the British Library.
kind of comment, explanation, or presentational “markup”) are attached to the online document to enhance its reading. Such information might relate to additional information other than the words on the page, including author, title, provenance, and date, or section demarcations, like book, chapter, paragraph (see the discussion below on TEI: a particular form of XML markup). Metadata can also relate to entities within the text, too, such as the people, places, or events mentioned. Information of this kind can give a text added value, by which the “hard copy” text (whether manuscript or book) is supplemented with additional semantic information.

For the rest of this chapter I want to reflect on the value of digital annotation through my own experience on two projects, in which we have used the concept of place as a starting point for analyzing historical documents.

**Spatializing the Humanities**

The idea of place and space is profoundly important in humanist thought, be it the spatial organization of narrative, the history of places, or location-oriented analysis of cultural objects. Space is not simply an empty container for historical action but a significant medium for the development and understanding of culture. It occurs, writes Michel de Certeau (1984, 117–18), “as the effect produced by the operations that orient it, situate it, temporalize it, and make it function as a polyvalent unity of conflictual programs or contractual proximities.” People are born, live, and die in places; events always happen somewhere. According to eminent geographer Yi-Fu Tuan (1977, 6), “What begins as undifferentiated space becomes place as we get to know it better and endow it with value.”

Over the past decade the extraordinary growth of new technologies has reinvigorated the humanities’ spatial turn and given rise to a new discipline: the spatial humanities (Dear, Ketchum, and Richardson 2011). Dominating this field of largely historical-based research has been Geographic Information Systems (GIS). As the name implies, a GIS is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data. Mapping information can have a surprisingly powerful effect. As David Bodenhamer (2010, 28) suggests, simply by virtue of “locating historical and cultural exegesis more explicitly in space and time . . . [GIS] finds patterns, facilitates comparisons, enhances perspectives, and illustrates data.” Ian Gregory and David Cooper (2010, 89) discuss the capacity of GIS to enable scholars to explore “subjective geographies through the spatial representation of qualitative, or fuzzy, data.” For research into the ancient world, GIS provides the technical backbone to and interface for two important, open-web resources: Antiquity À-la-Carte (awmc.unc.edu/wordpress/alacarte/), which allows researchers to map and analyze historical data produced by the Ancient World Mapping Center; and ORBIS (orbis.stanford.edu/), which enables queries of the time and financial cost associated with different types of travel in antiquity.
Yet, even as such digital mapping technologies are enabling the emergence of the spatial humanities, critical geographers have been questioning the objectivity of the cartographic representations on which these technologies tend to rely. Conventional maps, in the words of David Harvey (1996, 4), are “typically totalizing, usually two-dimensional, Cartesian, and very undialectical devices.” Indeed, there has been a growing awareness that a history of space (and place) should also include “the history of representations” (Lefebvre 1991, 42), or that the spatial information embedded in non-cartographic material (like texts) have the potential to un-think the Cartesian tradition that has dominated understandings of space since the Enlightenment (Gurevich 1985). In the Greco-Roman world, maps appear to have been relatively rare commodities; even when they are mentioned in our sources, they are treated with suspicion or flatly criticized. In Aristophanes’s satire on the latest intellectual thinking, Clouds (423 BCE), for example, Strepsiades, who represents in some way the Athenian common man, is shown a map by a student of “Socrates”: as Sparta, Athens’s great enemy, is pointed out to him, he leaps back and exclaims, “So close to us? How about moving it a lot farther away from us!” (Clouds 215; ὡς ἐγγὺς ἡμῶν. τοῦτο πάνυ φροντίζετε, / ταύτην ἀφ᾽ ἡμῶν ἀπαγαγεῖν πόρρω πάνυ). It is a joke that plays on the fantasy that the map accurately embodies the space that it depicts (cf. Harley 1989). This is a rare example, however: Greco-Roman spatial knowledge is primarily mediated through poetry, drama, or historiographical narratives. These kinds of discursive representations present a challenge to traditional GIS, with its positivist emphasis on precise measurement and categorization. It is not only the case that GIS is notoriously poor at handling uncertain, incomplete, and ambiguous data (Gregory and Healey 2007), but also that text-based research studies of humanistic enquiry “raise fundamental epistemological and ontological issues for GIS applications” (Harris, Bergeron, and Rouse 2011, 228). Other ways of capturing and reimagining textual space are needed.

**The Forms of Place and the Spatial Form of the Histories of Herodotus**

Roughly contemporary with Clouds is Herodotus’s Histories, a rich example that I would like to pick through with reference to a project I directed: Hestia (hestia.open.ac.uk; see, for example, Barker et al. 2010; 2016). In this initiative we were interested in identifying and exploring two aspects of textual space: first, what places does Herodotus mention in his text and how does he describe them (the forms of place); and second, how his narrative is underpinned by and, to a certain extent, structures or organizes a sense of space (its spatial form). At the center of the Histories lies a similar trope that has just been seen in the example of Aristophanes, as Aristagoras, the tyrant of Miletus, arrives in Sparta seeking support for a revolt from Persian rule, armed with “a bronze picture, on which the whole world was engraved and all the seas and all the rivers” (Ἔχων χάλκεον πίνακα ἐν τῷ γῆς ἁπάσης περίοδος ἐνετέτμητο καὶ θάλασσά τε πᾶσα καὶ ποταμοὶ πάντες, 5.49.1). While Aristagoras exploits the representational form of this “map” to promote
the idea of easy conquest, Herodotus juxtaposes to this episode his own discursive representation of the same territorial scope (5.50–52), in a way that invites his reader to reflect on space as something lived and experienced and far more complex. Elsewhere, Herodotus laughs at the maps produced by his Ionian contemporaries that divide the world into two regions of equal size (4.36.2) and is also careful to put the claim that Asia belongs to the Persians and Europe to the Greeks in the mouth of certain Persian writers (1.4.4). It seemed to us working on Hestia that Herodotus’s representation of space was worthwhile exploring precisely because it was discursive and, in some shape or form, counter-cartographic (Purves 2010)—an anti-GIS, if you will.

Our approach was intentionally experimental and as innovative as possible, particular in our application and use of digital resources and mapping software, with varying degrees of success (Barker, Isaksen, and Ogden 2016); but at the heart of our investigations was the idea of semantic annotation. As a start, we reused a digital text of the Histories that was freely available from the Perseus Digital Library (via a CC-BY license). Encoded according to a particular XML annotation schema designed for textual analysis (TEI: tei-c.org), this text came with most of the places in the Histories already marked up. As such, it was easy to extract the place information and plot it in a GIS, which is what we did (Fig. 2). Even as a simple data cleaning exercise, this visualization was extremely useful, since we could easily spot “outliers” and correct their misattributions (such as “Priam”—the king of Troy—being incorrectly marked
as the town in the United States). But it was also useful for revealing at a glance two important ideas for interpreting Herodotus’s spatial representation: first, that, while the spatial scope of Herodotus’s narrative was broad, its focus was concentrated on the area of the Aegean; and, second, that the farther one moved from this center, the more likely physical features were used in description (rather than places). On the other hand, the representation of Herodotean places as “dots on a GIS map” simply served to reproduce the exact same Cartesian worldview that we had been at pains to challenge. Indeed, such a GIS visualization of Herodotus presented the danger not only of reproducing Cartesian principles but even of aestheticizing them (Wood 2010). Our thesis was that textual space was constructed by and through the stories told about it; point-based geometry was simply insufficient to capture this complexity, fluidity, plurality, ambiguity, and interconnectedness.

We came back, then, to the question of how Herodotus relates the places mentioned in his narrative: how, in other words, the narrative form constructs a sense of space. In taking this approach we were drawing on earlier work that had conceived of Herodotus’s representation of space as *hodos*, a “way” or “path” (or *hodological*: Purves 2010, 15; cf. Janni 1984), as well as more broadly the notion of space (or *topos*) as conveying “movement” (*kinesis*; sic. *topokinesis*: Turnbull 2007). This idea of movement (through space or of places) was inspired by our reading of Herodotus’s description of his narrative journey: “I’ll go forward with the further part of my account, going through small and great cities alike. For those that were once great have now become small, while those that were great in my time were before small” (*Histories* 1.5.3–4).

Subsequently, we found Franco Moretti’s (1998, 5) work on literary maps helped to provide a useful guide for bringing out the “place-bound nature of literary forms” and the internal (spatial) “logic of narrative.” I would now add to this Caroline Levine’s (2015, 16) recent emphasis on tracking “the forms of the content, the many organizing principles that encounter one another inside as well as outside of the literary text . . . structuring and patterning experience . . . but also competing and colliding and rerouting one another” for drawing attention to the possible political implication underpinning Herodotus’s spatial mapping.

Within this more self-consciously theoretical framework, we annotated the text ourselves, by hand (see Bouzarovski and Barker 2016). Our aim was not only to identify spatial entities in a more nuanced and comprehensive manner than what had been possible by relying solely on the data capture from the digital text; above all, we also wanted to mark the relations between these spatial entities in a semantically meaningful way—that is, at the sentence level, where one entity (or more) was the subject of a verb and one (or more) the predicate. After a close reading of one book—Book 5 of the *Histories* was chosen on the basis that it stands at the center of the narrative and represents the coming together of the Greek and Persian worlds—we decided on two kinds of annotations that best reflected our aims. First, in addition to
places we marked “proxies”—the individuals, social groups or even non-human agents that belong to the geographical imagination described by Herodotus; place-names alone were insufficient to capture the full complexity of spatial entities depicted in Herodotus’s *Histories*. Second, we identified the *textual relatedness* of the places and/or proxies according to their directionality (which place or proxy was the subject and which the predicate) and semantic association: that is, whether the verb joining them denoted movement (*hodology*) and/or transformation (*topokinesis*). Based on these qualities, we determined four kinds of classification:

1. Positioning (including verbs of inclusion, proximity, comparison, origin)
2. Movement (in, through, away from)
3. Intervention in place (e.g., governance, alliance, awareness, commemoration)
4. Intervention with movement (e.g., invasion, exchange, communication, settlement)

Sample diagrams depicting the connections that we drew up are given in Figures 3–5.

The first thing to note about these graphs is their density and complexity. They are difficult to read. And they are difficult to read in two senses. First, we were faced with the challenge of simply enabling the data to be viewed more easily: Figures 3–5 represent the same graphs (the “total network” of the *Histories* Book 5), but using different software or formatting choices to render their information more readable (essentially by making decisions what to highlight and what to leave out). Second, we also needed to render the visualizations in such a way that could help inform our analysis. Although the product of intense and laborious line-by-line analysis of an entire book, the graphs do not depict “results” in any conventional understanding of the term. They demand interpretation.

Both aspects of the (problematic) readability of network visualizations have been raised by historian Irad Malkin. A pioneer of network theory to rethink ancient Mediterranean space (and reconnect Greek peoples to their far-flung multicultural places), Malkin is highly skeptical of the utility of graphic visualizations of Mediterranean networks. Such “two-dimensional representations of connectivity,” he complains, “mostly turn out to be messy ‘spaghetti monsters,’” which not only need “very long verbal explanations” but also omit entirely “the networks’ multidirectionality, multidimensionality, and multitemporality” (2011, 18). In short, “It is not enough to draw a few points on the map, identify them as ‘Phokaian,’ and come up with a diagram of a Phokaian network” (2011, 149). As we discovered, even once we had slayed the “spaghetti monsters,” our visualization was still difficult to read. Generally, humanities scholars lack the visual literacy to appreciate, and analyze, the complexity and nuance of graphic data; moreover, were our graph self-explanatory, it would still lack any of the metrics or rules that help enable network analysis. Rather, our graphs are more like the baby-steps of Moretti’s network picture of Shakespeare’s *Hamlet* than
Fig. 3. The total graph network of place (and people) relations in Book 5 of Herodotus’s Histories, visualized in Gephi.
Fig. 4. The total graph network of place (and people) relations in Book 5 of Herodotus’s Histories, visualized using the Science of Science (Sci2) Tool. Indiana University & SciTech Strategies (http://sci2.cns.iu.edu), with the support of Scott Weingart. The size of the node shows the strength of the relations of each place: the bigger the node, the more relations a place enjoys.
those used in the science of network theory (Moretti 2011). In fact—and this is our main point—our graphs are better understood as provocations rather than illustrations, as stimulations for further analysis and reflection. They are tools for argument, not the proof or replacement of it.

It is as provocations, I contend, that our graphs can have a strong analytical value by challenging dominant Cartesian-based assumptions of Herodotean space. Such “X-ray” maps (Moretti 2011, 4) bring to the fore the underlying ways in which Herodotus constructs space in his Histories in terms of relations between peoples and places rather than according to topographic reality. In Figure 5, which represents our best current attempt at representing the “total network” graph for Book 5, those places

Fig. 5. The total graph network of place (and people) relations in Book 5 of Herodotus’s Histories, visualized using the Science of Science (Sci2) Tool. Indiana University & SciTech Strategies (http://sci2.cns.iu.edu), with the support of Scott Weingart. The size of font shows the strength of the relations of each place: the bigger the font, the more relations a place enjoys.
(or peoples) most connected are depicted at the center of the graph; the bigger the font, the more times the place is mentioned in relation to others; arrows show the direction of the relation, whether one place is acting on, or being acted on, by another. What this graph reveals, we argue, is a picture of two emerging spheres: an Eastern world centered on and around Persia, and a Western, Greek one with Athens and Sparta dominant. Such a division not only refers back to, and picks up on, Herodotus’s own initial distinction between Greeks and barbarians (1.1.1); it is one familiar to us, as implied by the use of the cardinal descriptors “Eastern” and “Western,” which come with so much ideological baggage (Said 1978). Yet, the picture is in fact far more complex than a simple geographically based ideological distinction between East and West (cf. Pelling 1997). In addition to the two spheres interlocking in various ways, a number of Greek places—from islands such as Cyprus, Naxos, and Chios, to broader regions such as Ionia, the Aegean Sea, Macedonia, and even the notion of “Greece” itself—have become attracted into the Persian sphere. This picture not only represents the growing power and influence of this super-state over the course of Book 5; it also suggests the alterity of these Greek places, caught between the Greek mainland proper and the fringes of the Persian Empire, and caught up in the growing opposition between the two. In fact, it is the Greek city of Miletus that acts as the fulcrum around which the Greek and Persian worlds pivot and collide: the even bidirectionality of the arrow between Miletus and Persia shows that agency flows both ways in this relationship and reminds us of the activity of Aristagoras, tyrant of Miletus, in engineering conflict against the Persians. The Greek sphere is even less homogenous. Characterized by strong internal dynamics, this sphere reveals the strongest relationship in the whole graph: Sparta’s actions on (and against) Athens. This history (and not only at this point) is also about the internal conflicts among Greek cities and peoples, in ways that both anticipate and presage the outbreak of the disastrous war of Herodotus’s time (Barker and Pelling 2016).

Such diagrammatic forms, then, offer alternative pathways through Herodotus’s space that subvert our long-standing cartographic tradition and suggest new ways of thinking about pre-modern narrative spatial thinking. They present, among other things, (1) ways of thinking about how Herodotus constructs space (or, rather, how events in his Histories are embedded in and structured by certain places); (2) ways of contesting essentialist views of space (“Persia,” “Athens,” etc.), by refocusing attention on the relatedness, intersections, and movement of places and peoples; and (3) ways of identifying and analyzing how space and place are constructed by historical agents (Pred 1983). Fundamentally, such graphs need to be viewed not as illustrations of the data so much as provocations to analyze them further, precisely to elicit the “long verbal explanations” of which Malkin (2011, 18) complains. Depicting network graphs means little without understanding how the networks operate, in particular how these links “jeopardize, stabilize, or reroute bounded unities” (to again quote Levine 2015, 120), or produce, in the words of Doreen Massey (2005, 158), “constellations of
trajectories which, though interacting and undoubtedly affecting each other, have very different rhythms.” Thinking of maps as provocations, however, presents a further challenge to their visualization of data. As Irad Malkin (2011, 18) has asserted, “We are in need of new modes of graphic representation that would avoid the more precise, yet messy, ‘lines connecting dots’ and avoid the pitfalls of dazzling oversimplifications.” Missing, for example, from our series of network graphs is the concept of time, whether we mean chronological “real” time or narrative time, the unfolding of the narrative or chronotope (Bakhtin 1981). Yet, since relations do not remain immutable but change according to historical circumstances (even over the course of a single book), they need to be located in time as well as space.

What we are talking about here is the need for a new kind of map that better enables exploration and analysis on the part of the viewer, or, as David Bodenhamer (2015, 23) suggests, is “alert to the basic ways in which we analyze space—movement, direction, proximity, and connection, among others—and discover the patterns that occur within and among places.” Such a map would be a deep map. The deep map has been recently defined as “afford[ing] open-ended exploration of a particular place” by being “more fully situated archives in which one might find myriad traces of evidence about a site, and from these begin to build stories and arguments” that “would represent the pathways that weave places together, creating urban life by their interplay and motion” (Ridge, Lafreniere, and Nesbit 2013, 177–178). It combines “highly structured, mathematically precise cartographic exercises with more subjective and discursive mapping methods” in ways that defamiliarize topographic space to reveal “a multi-layered network of connections between people and places, past and present” (Gregory-Guider 2004, 1–2). Thus deep maps are emerging as a concept for managing heterogeneous data because of advances in technology, by virtue of which they can be “inherently unstable, continually unfolding and changing in response to new data, new perspectives, and new insights”—in essence, “the means by which we represent the contested meanings of space and place, as well as the dynamics that produce them” (Bodenhamer 2015, 21).

**Linking Information through Common References to Places**

The challenges of visualization still need to be addressed, though progress is being made particularly in the design of user-led exploration of networks (e.g., Dörk, Comber, and Dade-Robertson 2014). For the final part of this chapter, however, I want to return to the idea of semantic annotation and its powerful use as a means of linking the kinds of heterogeneous data I’ve just described as potentially forming the basis of a deep map. Thus far we have explored place in one document, the narrative history of Herodotus of Halicarnassus. Yet places have a much richer history than can ever be captured in any one representation, cartographic or otherwise. Places “gather things in their midst”—both animate and inanimate entities, as well as experiences
and histories. “The process of gathering (things, emotions, people, memories, etc.),” according to Edward Casey (1996, 24, quoted in Cresswell 2015, 51), “suggests that there is a relationship between the inside of a place (which gathers) and an outside (from where things are gathered). It underlines the relational nature of place—the necessity of place being related to its outside.” Or, following Doreen Massey (2005, 130) again: “One way of seeing ‘places’ is as on the surface of maps . . . But to escape from an imagination of space as a surface is to abandon also that view of place. If space is rather a simultaneity of stories-so-far, then places are collections of those stories, articulations within the wider power-geometries of space. Their character will be a product of these intersections within that wider setting, and of what is made of them. And, too, of the non-meetings-up, the disconnections and the relations not established, the exclusions. All this contributes to the specificity of place.” Both critics point to the importance of reading places as fundamentally relational, as both connected to other external entities and linked in within wider power-geometries of space.

One opportunity that is emerging for this kind of analysis is through the method of linking data through semantic annotation. By means of example, and by way of returning us to the ideas about openness, transparency, and accountability with which I began, I end this chapter by discussing a second initiative that I have been helping to direct: Pelagios.

Using the linking potential of the web, Pelagios has been pioneering a means of “mutual contextualization,” whereby any online record—be it a text, image, or database—can be connected to another simply by virtue of having something in common with it. The Pelagios method is to achieve linking through the semantic annotation of place references in web documents. That is to say, whereas other mechanisms for linking data require agreement at the level of ontology (essentially a standard data representation) or seek to aggregate data all in one place, Pelagios works through a decentralized modus operandi that uses the method of semantic annotation to enrich metadata in the form of “stand-off” markup.

The method of annotation is twofold (see Simon et al. 2017). First, locate the place(s) in an online document and create an annotation; then align or “resolve” these annotations to a digital authority that provides the means to identify and disambiguate between different places. Digital authority files for places are online gazetteers: as well as providing information about places that are useful for their understanding—items such as geographical coordinates, type of place, political affiliation, bibliography, etc.—online gazetteers also provide what are called uniform resource identifiers (URIs). These URIs essentially operate like social security numbers to help disambiguate places of the same name or identify the same place with different names or typography (Simon et al. 2016a). By annotating the places in a document with URIs, not only can they be correctly identified and placed on a map, if desired, but also, because other data providers are also using the same identifiers, disparate
documents can be linked through their references to common places. By these means, Pelagios is able to link two different kinds of knowledge communities—those who hold information about places (the gazetteer community) and those whose documents refer to those places (the research community that depends on place identification and disambiguation).

Originally a trial collaboration to interlink various online data providers who had the technical know-how to align datasets, Pelagios now offers any researcher, data curator, student, or general enthusiast a simple, intuitive means to encode place information in their work, whether or not they are curating data or are interested in producing it as linked data. This do-it-yourself, open-access, and open-source (geo)annotation platform is called Recogito. Thus, where in the Hestia project we had to annotate place and space information by hand, essentially by inputting the information into a spreadsheet (a CSV file), any user is now able to annotate the text itself directly, and a CSV file is only one of a number of output options for the data. In addition, Recogito lowers the barrier of entry for annotating a document. Figure 6 shows the Recogito page for uploading a document for annotation, which can be a text (plain text or TEI encoded), an image (e.g., a map), or a database. Figure 7 shows the annotation screen, where those entities in the document to be marked as places are identified and then aligned or resolved to the appropriate entry in the appropriate gazetteer. Figure 8 shows the map view in which all the newly annotated places are presented in a map interface, including the option of displaying “snippets” of text (if appropriate). This built-in visualization means that anyone can now identify and plot places in a narrative to reveal their spatial footprint and create maps out of texts. Of far greater significance, however, Recogito also provides other download options—KML, RDF, GeoJSON, and CSV—for use in other applications, such as GIS or network analysis tools, as well as for publishing the annotations as linked data. It should also be noted that Recogito provides full capability to annotate image documents (such as pre-modern maps), with tools including zoom, rotate, and transcribe, all with IIIF (International Image Interoperability Framework) support.

What I have been describing is very much the technical structure of Recogito in terms of the options that it provides for annotating and then making use of the annotations. But there is an important social element too. On the one hand, Recogito has been designed to enable social annotation, by which means one can share documents with others and work on the same document synchronously, which has a significant role to play both for collaboration and in the classroom. On the other, while Recogito was

---

6 To access Recogito, simply log in to recogito.pelagios.org/. For all the (open-source) code produced by Pelagios, see github.com/pelagios. For more information about Pelagios and its community at the time of going to press, see pelagios.org and medium.com/pelagios.

7 This two-step process—identifying the place (target) and resolving it to the gazetteer (body)—is essentially producing RDF triples, the data structure for linked data.
Fig. 6. Screenshot of an upload document page for Recogito.

Fig. 7. Screenshot of an annotation page in Recogito, showing a place ("Thessaly") being annotated (with alignment to the Pleiades gazetteer) in Herodotus’s *Histories*.

Fig. 8. Screenshot of an output page for Recogito, showing the map of annotations for Herodotus’s *Histories*, along with a text snippet for the highlighted place (the river Strymon).
initially designed as a digital tool to extend a technical infrastructure (linked data), first
and foremost it is easy to use and puts the onus on the input of the domain expert rather
than a data technician. Not only does the Recogito user not need to know anything
about linked data; most in fact use Recogito not to produce linked data at all but rather
to annotate (and analyze) entities within documents. This aspect of public use is critical.
Not only have we put the focus on easy use as a way of embedding this kind of practice
in current workflows—to make it a “practice of annotation,” as it were; 8 we have done so
in a way that makes it abundantly clear that every annotation is essentially an assertion
or interpretation. Indeed, this is why we believe that semantic annotation should be
in the hands of the domain experts. In addition, all annotations are captured in the
metadata of Recogito to achieve maximum accountability and transparency. 9

Producing a technical platform that can be used by anyone, and the results of
which are open and transparent to all, needs to be one of the primary concerns of
the digital humanities if it is to mature as a discipline that can develop humanistic
research and pedagogy in new areas and in new ways. As Lorna Hardwick (in Barker
et al. 2012, 191–92) has written, “The potential offered by technology for demolition
and constant reconstruction of clusters and systems of knowledge opens sometimes
startling associative possibilities. This is both threatening to academic authority and
potentially liberating if—and only if the students and other users are equipped to
reflect on the new clusters that they create; that is, they do not merely accept these as
reflecting a supposed ‘natural order’ of things but, like Plato in the Theaetetus, probe the
definition of epistemology as ‘true belief together with an account.’” The problematic
faith in technology has as much to do with the new tools and methodologies being
used, as with the big data that demand “distant reading”: as humanists adopt scientific
methods to try to make sense of these data, “we also need to be aware of what is
at stake in using those digital tools [and] how the tools themselves can distort the
research” (Barker et al. 2012, 191–92). As Martin Foys (2013) puts it, “We are still in the
age of the digital incunabulum—the time when the new media form has not yet figured
out how best to realize its own inherent logic.” One of the critical opportunities here is
for the humanities to shape the digital research agenda. Arguably the most important
advance would be, as Bethany Nowviskie (2013, n.p.) contends, to have better control
over the data.

8 With this phrasing, I wish to recall Pierre Bourdieu’s (1990) work on habitus, both for the sense of agency
in the development of infrastructure and for the reproduction of social values in everyday practice. Both
points, I believe, are essential for the embedding of Pelagios’s linked data infrastructure in the everyday
practice of scholarship.

9 While Recogito stores the documents that users upload and the annotations they create, it is not a publication
platform. Users are, however, free to make their annotations open, as well as the documents on which they
are working, provided that these comply with copyright regulations. Throughout, users remain in control:
at any time, they can delete their documents (which will also delete the annotations and everything else
connected to the document) or indeed their account from Recogito (which will delete all their document
data, metadata, annotations, etc., in compliance with GDPR). See Pelagios, n.d., “Privacy Policy”.


If Recogito enables this through putting the power of annotation into the hands of the non-technical user, so that places and spatial entities can be researched in ways far more comprehensive and nuanced than before, how to make use of the resulting linked data annotation network is still only at an early stage of conception and development. The Pelagios team has made an initial foray into this space by developing a prototype search tool called Peripleo (Pelagios. n.d. “Peripleo.”; Simon et al. 2016b). As the search engine of linked data maintained by partners of Pelagios, it aims to better enable the discovery and analysis of the connections between scholars, projects, and institutions studying places of the past. Currently most of our partners publish information about ancient places and physical objects, such as archaeological finds, but other kinds of data, including geo-tagged literature or data transcribed from historical maps, as well as other kinds of entities (people and time periods, for example), are also being annotated.

Peripleo enables various search features. First of all is a basic search for places, objects, and other types of data. Using a search box to search by keyword, Peripleo provides suggestions for search terms and specific records matching each query. Search results are coded based on item type, using colors and symbols based on places, objects, time periods, people, and datasets. Second, search results can be narrowed down by using filters. The most prominent part of the filter panel is the time histogram, which shows how search results are distributed over time; but one can also filter by data source, item type, or geographical area. Third, as well as searching for data, summary information about a whole dataset can be displayed. Figure 9 shows an example dataset in Peripleo of an archive, which includes a list of all results contained within its dataset and its geographical footprint on a map.

Peripleo functions rather differently from other map-based search interfaces by virtue of the data in Peripleo being networked. That means items are internally connected through links. One major consequence is how the map works: one dot on the map is not always the same as one search result. Instead, one dot may represent many results connected to the same place, and dot size varies accordingly. Vice versa, one result can appear as many dots. For example, an archaeological artifact might be linked to a findspot as well as to a place of production, while a work of literature might contain references to hundreds of places. Fundamentally, Peripleo enables discovery of data that might be relevant or interesting, based on the idea of place. What it does not do is provide any semantic reasoning for linkages; it is important to note this limitation. As Andrew Prescott (2013) has complained: “There is a tendency to think that if we link a random group of resources together, somehow the magic of linked data will give us instantly new perspectives and new understandings for a particular place or period . . . Sadly, scholarship is much harder than this. Linking of data can be a very useful scholarly technique, but we need to be clear about why we are linking data, what sort of data we are linking, and our aim in doing so.” So far as Pelagios is concerned, computers may be linking data, but humans must do the thinking about them.
A final point relates to this idea and brings us back, once more, to the community aspect of digital work. We have already seen the social practice of annotation facilitated by Recogito. More broadly, Pelagios’s linked data methodology works by linking people as well as data. It primarily seeks to achieve this through our social media account and website, where we are in the process of documenting our resources, especially with regard to how people (from different communities) can use them. Furthermore, a major focus of the current phase of Pelagios has even been on redistributing financial resources (from the grant package that we currently receive from the Andrew W. Mellon Foundation) in the form of Resource Development and Working Groups grants. The former provides grants for the development of resources that will be of benefit for the broader community, while the latter sponsors events for addressing specific issues of a wider concern to the community. One example of a Resource Development Grant has been the redevelopment of the mapping tiles for the Digital Atlas of the Roman Empire (dare.ht.lu.se). Switching from raster-based technology, which sets limits to the area and zoom levels served and consumes a lot of resources, to vector tiles (the information for which is layered, not “baked in”), it has been possible not only to update
this map to include different periods but also to make it downloadable for people to use and combine in their own projects.\textsuperscript{10} As for working groups, one has been dedicated to multilingualism to translate Pelagios resources into different languages, another to embedding Pelagios tools (especially Recogito) into classroom teaching, while another called “Linked Pasts” has tried to map out the emerging digital ecosystem and what would be needed for this new kind of digital scholarship.\textsuperscript{11}

### Conclusion

Perhaps due to its fragmentary nature, the study of the ancient world has always been open to a holistic approach that has read archaeology through a historiographical lens and vice versa. At the same time, the growing professionalization of academia, as well as the exponential growth in scholarly output, has resulted in ever more specialized fields of research and ever-sharper disciplinary boundaries being drawn. New digital technologies carry an even greater risk of fragmentation, since the different models and standards used to structure and represent digital data present substantial challenges to their discoverability, accessibility, and (re)usability. And yet, not only can the web be a powerful means of bringing disparate data together for both visualization and analysis, but also scholars of the past now have more vital roles to play in informing the public through this medium.

In this chapter I have explained the process of semantic annotation through two pioneering digital initiatives, the Hestia project and Pelagios, and considered the potential uses of this kind of semantic annotation, including the repurposing of data in other systems (such as GIS or network visualizations) and the linking to other resources. While many challenges remain, I hope to have shown that, contrary to much current thinking, digital tools can in fact enhance, and even demand, the close reading of data in ways that facilitate a more nuanced understanding of place as represented both in our historical sources and through the evidence on the ground. More than this, however, I hope to have highlighted that the social practice of annotation represents one method by which we as experts in our field can communicate information to broader publics with the sufficient robustness, transparency, and accountability necessary for preserving the health of the democratic commons.

\textsuperscript{10} For a fuller description, see Åhlfeldt 2017a. For a list and description of all Pelagios resource development grants, see Pelagios, n.d., “Resource Development Grants” and “About Working Groups”, whose reports are available online via the dedicated Pelagios blog, medium.com/pelagios.

\textsuperscript{11} Karl Grossner has produced a working “white paper” as a way of making use of the results from the initial work undertaken by the Linked Pasts Working Group (LPWG) of Pelagios Commons. See https://docs.google.com/document/d/1h9pvrswZjktiN1OUn03K4smedjwD4OeW6hfwxijhUW0/edit.
Acknowledgments

I would like to thank Chris Roosevelt and ANAMED for the opportunity to speak before and learn from such an engaged and informed audience. I would also like to thank my colleagues on Hestia (Stefan Bouzarovski, Leif Isaksen, and Chris Pelling) and my investigative team colleagues on Pelagios (Leif again, Rebecca Kahn, Simon Rainer, and Valeria Vitale). The Hestia project was made possible with funding from the UK’s Arts and Humanities Research Council (AHRC), in the form of an Early Career Research Grant (2008–2010, ID. No. AH/F019459/1), and with subsequent “follow on” funding (2013–2014), where we shifted from geospatial experimentation to consolidating the technological outcomes and sharing knowledge of those methods and tools with cognate disciplines, interested cultural heritage groups, and the general public. Pelagios has received funding from: JISC (under its Geospatial Engagement and Community Outreach, and Resource and Discovery schemes); the AHRC (Digital Transformations); the Open Knowledge Foundation (Open Humanities Award); and, in three phases of funding, the Andrew W. Mellon Foundation. I benefited also from an Alexander von Humboldt Foundation Research Fellowship for Experienced Researchers (2012–2013) to conduct research with Klaus Geus at the Historische Geographie des antiken Mittelmeerraumes, Freie Universität Berlin and Gerhard Heyer at the Institute of Informatics, University of Leipzig, and an A. S. Onassis Foundation Visiting Fellowship for research at the National Hellenic Research Foundation, Athens (2014–2015).
Works Cited


