

The Data Interface

written by Jeremy Huggett | 11/10/2024



Detail from *Woman at a Window*, by Casper David Friedrich (1822)

We understand knowledge construction to be social and combinatorial: we build on the knowledge of others, we create knowledge from data collected by ourselves and others, and so on. Although we pay a lot of attention to the processes behind the collection, recording, and archiving of our data, and are concerned about ensuring its findability, accessibility, interoperability, and reusability into the future, we pay much less attention to the technological mediation between ourselves and those same data. How do the search interfaces which we customarily employ in our archaeological data portals influence our use of them, and consequently affect the knowledge we create through them? How do they both enable and constrain us? And what are the implications for future interface designs?

As if to underline the lack of attention to interfaces, it's often difficult to trace their history and development. It's not something that infrastructure providers tend to be particularly interested in, and the Internet Archive's **Wayback Machine** doesn't capture interfaces which use dynamically scripted pages, which writes off the visual history of the first ten years or more of development of the Archaeology Data Service's ArchSearch interface, for example. The focus is, perhaps inevitably, on maintaining the interfaces we do have and looking forward to developing the next ones, but with relatively little sense of their history. Interfaces are all too often treated as transparent, transient – almost disposable – windows on the data they provide access to.

Nevertheless, a recent paper by Seaton *et al.* (2023) on barriers to archaeological data reuse suggests that users were reasonably satisfied overall with their interfaces, although they might experience some technical issues and frequently mentioned a desire to enhance their functionality. But improving how an interface works is a challenge that extends beyond the purely technical and functional into the social and political, which makes an understanding of their development and change all the more important.

For example, Gray (2023) has recently written of interfaces configuring users through scripting and ordering interactions, that they can be political in the way they are designed to produce particular kinds of users and to steer users in particular ways. He describes interfaces as “zones of contestation” negotiating our relationships with the infrastructure and its data (Gray 2023, e10-6), which stands in marked contrast to the commonplace conception of the interface as little more than a non-controversial tool enabling access to our data. Interfaces are not transparent windows on our data in the commonly-understood sense – as Arns (2011, 256) has emphasised, the computer science perspective of transparency is one entailing invisibility and concealment, which, as she says, can be useful to hide underlying complexities, but means that users should not assume that what is revealed is complete or even real. And it is not difficult to find instances amongst our well-established archaeological interfaces where we encounter the effective hiding of data, the concealment of implementation, and interface limitations placed on our access and retrieval, which in combination mean we operate within a series of **filter bubbles**.

Some years ago, I wrote something about the implications of this – the way that we are, often unknowingly, constrained by the interfaces that we rely on for our access to data. At the time, I highlighted Burdick’s desiderata for interfaces (Burdick 2015, 31), and still find them compelling today (Huggett 2023a, 30). She argued that interfaces should not be black boxed, so their inner workings are available and capable of being reconfigured if desired. She also proposed that an interface should not have a default worldview, that it should enable multiple perspectives and make it clear that the current view is one that is a choice rather than the primary or favoured option. Related to this, she argued that interface should be observer-dependent: that interface views should be context-specific, observer-dependent, partial, and situated. As archaeologists we’re accustomed to seeing our data in these terms – our interfaces should also reflect this. Likewise, our interfaces should reflect the ambiguity in our data, again something we’re familiar with in our data but which is lost in translation in our interfaces.

So how might this be practically implemented in interfaces going forward? Some inter-related suggestions might include:

- *Use exact search by default* to avoid unexpected outcomes. At present, current interfaces use varying degrees of fuzzy search methods. For instance, given a term such as ‘gold coin’, the **ARIADNE** portal will explicitly treat this as two separate keywords and match them wherever they appear in the metadata. Elsewhere, **ArchSearch** appears to both ‘AND’ and ‘OR’ such keywords – a search for ‘motte and bailey’ castles will also return mottes which have no baileys and baileys without mottes, for example. In neither case can this behaviour be changed by the user.
- *Ensure the interface clearly reflects the underlying tools* being applied. For example, although an **ARIADNE** help hint explains how the keyword search works, the interface itself does not make it clear; similarly, **ArchSearch** doesn’t explicitly reveal its handling of search criteria.
- *Include a facility to use Boolean operators*, brackets for grouping and precedence, and so on in order to create more precisely tailored queries. Such a device is not uncommon elsewhere (query builders in GIS, for instance) and they have been used in archaeological interfaces for over thirty years at least to provide a more flexible approach to searching. This would enable the kinds of treatments highlighted above to be over-ridden if desired as well as allow multiple search criteria to be combined as

desired. It would logically appear as an Advanced search, one that would be much more useful than either the current Advanced blade in ArchSearch or the Advanced tabs in ARIADNE. Desire for a Boolean search facility is also highlighted in the recent SEADDA survey (Seaton *et al.* 2023, sect.3.7; 4).

- *Build for flexibility* – as far as possible don't presume what the user will want. For example, the current reliance on the *What Where When* triple in searches is limiting in terms of depth and scale as currently implemented in ArchSearch and the flexibility introduced by the faceted classification quickly runs out further down the hierarchy (for instance, the limited categories of *Physical Evidence* or *Excavation* identified within the *What* facet).
- *Expose all metadata and paradata* including relationships between data. For instance, the *What Where When* triple could be usefully extended to include *Why* (the reasons behind the production of the data) and *How* (the methodologies used in the production of the data). It would also help to reduce the problem of the disassociation of associated records, for instance.
- *Enable alternative mappings* of the ontologies or terminologies used in data retrieval to improve flexibility and allow for different usages across time and space.

Access to an API would no doubt enable some of these suggestions to be implemented, but would require a degree of technical know-how, whereas implementation in the interface would open up the possibilities to a much wider audience.

One area remains largely unexplored in terms of archaeological interfaces – that of the visual. Elsewhere, I've argued for the recognition of the profoundly visual nature of archaeology in our digital methodologies (Huggett 2023b). I'm not thinking here of the visual search tools we currently see: whether it's the map search found in ArchSearch or in ARIADNE, or the interactive timeline in ARIADNE, for example. Instead, I'm thinking of the way that photographs, field drawings, CAD files, GIS layers, and so on, capture, communicate, and archive in visual form much of our archaeological knowledge, and this is not reflected in our interfaces which remain resolutely textual in outlook, and if they incorporated visual media at all, it would more than likely be via the textual metadata associated with them. It's interesting to consider whether artificial intelligence tools might have a role to play in this regard, given the (not uncontroversial) use of generative AI to perform image classification and recognition. However, most AI work is currently focused on photographic images – and largely modern rather than historic photographs – not diagrams, plans and the like which form a large proportion of the archaeological record. These remain a largely un-investigated area of study in AI terms, and certainly not one that has been addressed from an archaeological perspective.

Ultimately, interfaces provide certain perspectives, subordinating others, concealing motivations and agendas: as Drucker puts it, "Point of view is structured into interface design but never exposed or marked conspicuously." (2020, 104). The interface provides users with the illusion of control: the search criteria can be changed, the search refined, the results displayed and extracted, but Drucker argues that "we are always most deluded when most convinced of our capacity for agency" (2020, 110) and that when we engage with this illusion, we align ourselves with the particular limitations and restrictions imposed by the interface.

In the end, what I'm arguing for is that we should aim to present alternative perspectives in our interfaces, to allow for other research questions, other kinds of enquiries. Indeed, I would suggest

that it is increasingly vital that we do so, as we are perhaps starting to see a form of standardisation developing across our archaeological interfaces. After all, once something becomes standardised, even institutionalised, it risks becoming conservative and resistant to change, with implications for future research opportunities.

[This post is based on part of a presentation given in the Swedigarch lecture series on 9th October 2024. My thanks to Helena Hulth, Fredrik Gunnarson, and their colleagues for the invitation, their warm welcome, and for the discussion which followed.]

References

Arns, I. (2011). Transparent World: Minoritarian Tactics in the Age of Transparency. In C. Andersen & S. Pold (Eds.). *Interface criticism: Aesthetics beyond buttons*. Aarhus University Press, pp. 253-276.

Burdick, A. (2015). Meta! Meta! Meta! A Speculative Design Brief for the Digital Humanities. *Visible Language: The Journal of Visual Communication Research*, 49(3), 13-33.

Drucker, J. (2020). *Visualization and interpretation: Humanistic approaches to display*. The MIT Press.

Gray, J. W. Y. (2023). What do data portals do? Tracing the politics of online devices for making data public. *Data & Policy*, 5, e10. <https://doi.org/10.1017/dap.2023.7>

Huggett, J. (2023a). Deconstructing the Digital Infrastructures Supporting Archaeological Knowledge. *Current Swedish Archaeology*, 31, 11-38. <https://doi.org/10.37718/CSA.2023.01>

Huggett, J. (2023b). Extending Discourse Analysis in Archaeology: A Multimodal Approach. In C. Gonzalez-Perez, P. Martin-Rodilla, & M. Pereira-Fariña (Eds.), *Discourse and Argumentation in Archaeology: Conceptual and Computational Approaches* (pp. 159-189). Springer International Publishing. https://doi.org/10.1007/978-3-031-37156-1_8 (OA version).

Seaton, K.-L., Laužikas, R., McKeague, P., Almeida, V. M. de, May, K., & Wright, H. (2023). Understanding Data Reuse and Barriers to Reuse of Archaeological Data. A quality-in-use methodological approach. *Internet Archaeology*, 63. <https://doi.org/10.11141/ia.63.8>