1   Title

1. Hi everyone, glad to be here.

2. The paper I'm about to present draws from my dissertation, which ins't strictly about open science but is heavily inspired by challenges that digital archaeologists commonly face while engaging with open research data,

3. namely the awkwardness of opening a spreadsheet and entering a world of tacit assumptions and experiences that differ from your own, and the actions taken to reconcile these disparities.

2   Objectives

1. I consider uploading and downloading formally arranged information to and from a digital repository as just one kind of data sharing, which exists among a range of other additional collaborative actions.

2. Archaeologists share data all the time through alternative means, following social and technological protocols that don't necessarily involve making the records available 24/7 on the web.

3. My work basically tries to articulate these ways in which archaeologists already share data, which do not necessarily rely on the global information infrastructures that "open archaeology" is concerned with developing.

4. More specifically, I call attention to the collaborative commitments that acts of sharing entail, and the roles of mediating technologies in these interactions.

5. I think that many of the problems that we face when developing large-scale, web-based, open data infrastructures relate to our inability to grapple with archaeologists' discomfort and apprehension regarding the expansion if their collaborative commitments into new dimensions, for which the discipline has not yet established social and professional norms.

6. So I'm here to shed some light on some incongruities between how archaeologists tend to share data on a local and collegial scale, and the very different sociotechnical arrangements that are imposed by open data infrastructures, at least under the paradigm we have opted to pursue.

3   My approach

1. First a bit about my approach.

2. I examine archaeological research as cultural practice, as work that occurs within a social environment.

3. I'm less interested in what archaeologists are looking at than how they do it, and crucially, how the meanings they ascertain about things are validated through participation within a collective enterprise.

4. I'm specifically interested in examining how meanings are derived through collective action, how social and technolgical systems mediate and encourage behaviour targeted towards collectively-held goals, and the tensions that emerge through these cooperative systems.

5. So I set out to observe and interview archaeologists as they work, and articulate the social and technological circumstances that underlie the formation of a communal data stream.

4   Data collection and analysis

1. I recorded over 150 hours of audio and video footage, based on over 90 hours of observation, using three cameras that I placed in different parts of archaeological work environments, including on participants' foreheads.

2. I conducted over 30 interviews with more than 20 participants.

3. And examined the documents that archaeologists produce and handle on a regular basis.

4. I did this over several years, and three cases are represented in my work.

5. Observations and interviews were held in settings where people were excavating, collecting samples, processing and identifying finds, working at their laptops, giving site tours, and holding casual conversations.

6. However, most of the evidence I cite here today is derived from retrospective interviews, since this is the most direct way in which archaeologists expressed their attitudes on data sharing and collaboration.

7. This is just one slice of my much bigger project.

8. For the sake of keeping good time, I will not go into detail on my data collection and analysis methods, but I'm happy to discuss this in greater depth later on or during the break.

9. Also, all of the names presented here are pseudonyms.

5   Archaeological project collectives

1. So one of my points of departure was my puzzlement about how archaeological research could actually get done.

2. Archaeological project collectives:

   (a) comprise teams of people,

   (b) working in different settings,

   (c) in sequence and in tandem,

   (d) at different time scales,

   (e) using a variety of tools,

   (f) applying a diverse range of methods,

    (g) bringing their own unique outlooks and experiences,

    (h) and responding to local circumstances and challenges.

3. As with other miraculous efforts to integrate distributed collective efforts, such as wikipedia or large-scale development of open source software, archaeological projects implement social and technological structures to help ensure that work is done in a cooperative and productive manner.

4. Members of archaeological projects pool their data into communal data streams, using a combination of digital and analog devices and protocols, such as labelled tags and sample baggies, spreadsheets and todo lists, email attachments, text messages, artefact registries, and loads cheap USB thumb drives scattered all over the dig house.

5. And they moderate information flows by distributing work across the team.

6. A project director, who oversees all work, assembles the sociotechnical system by recruiting members from their professional networks and by establishing norms and protocols for sharing and validating information.

7. However, each individual comes with their own priorities and expectations, and directors try to ensure that all work is fulfilling for individual and the collective alike.

6 Territoriality in the field

1. In general, I observed that work is distribution in a territorial way.

2. Most people within a project are allocated a specific domain corresponding with their role.

3. In fieldwork, for instance, each supervisor retains control over the trench that they are assigned to excavate.

4. For example, Lauren, a trench supervisor at Case A, discussed her collegial relationship with Theo, who supervises work on another nearby trench.

5. They regularly share insights on each other's work, but it's clear that each person retains control over their respective trenches.

6. Similarly, Ben identifies his trench as an extension of him, as something that reflects his personality and his competency as an archaeologist.

7 Territoriality among specialists

1. With regards to specialist work – including analysis of archaeological finds, samples, and relationships – Basil, the project director for Case A, feels a need to ensure that work is distributed fairly and with clear intent.

2. He sees that specialists, especially those early in their careers, are anxious about retaining control over a particular slice of research materials, and he tries to assuage these fears by being very upfront about his expectations regarding what each person is responsible for.

3. The allocation of research materials, which can be leveraged to produce published research outputs, has tangible effects on people's careers, and Basil therefore takes this very seriously.

4. At the same time, he rewards seniority and uses the promise of access to research materials as a way of enticing more involved participation in the project.

8   Territoriality and sovereignty over projects

1. With regards to projects as a whole, Rufus, who co-directs Case B, identifies project directors as the most invested stakeholders, due to the effort they expend getting things off the ground.

2. Of course, their territory only pertains to projects that they lead themselves, which reifies the boundaries between projects.

3. This seems to mirror the structure of corporate firms, whose boundaries are defined by what is owned by the owners, or in these cases, the project directors.

9   Agency in contexts of fieldwork

1. I also noticed systematic behaviour regarding the distribution of agency within activity systems.

2. For instance, trench supervisors wield creative agency in their trenches, and leverage their assistants as tools to support the creation of context recording sheets and trench reports.

3. Assistants such as Ben, as in the example on screen, may be left in the dark about why they do things in certain ways.

4. However, when they acquire this understanding they may eventually take on creative agency over their own trenches.

5. Ben started off as a trench assistant, and in that role he simply did what he was told to do.

6. He effectively worked as his supervisors' digging instrument.

7. But when he came back the next season, he became a trench supervisor, and this new role came with a new imperative to fill out recording sheets.

8. He lamented that this also involved working his assistants in a similar way as he was worked the previous season.

9. He recognized that his new position is that of a quick and efficient collector of archaeological materials, and a producer of preliminary reports, which he then commits to the finds specialists and project director for further study and synthesis.

10 Agency in contexts of writing reports

1. Project directors draw from these trench reports and assemble overarching articles or monographs pertaining to the project as a whole.

2. For the director, each supervisor's contribution is genericized in the same was as the assistants are genericized for the supervisors.

3. Supervisors are rendered as tools that the director mobilizes in order to complete activities in which he is the creative agent.

11 Agency in contexts of data integration and reuse

1. We may extend this even further by identifying people who integrate data derived from multiple projects as creative agents in their own activity systems, and which leverage the work that project directors produce in a similar way.

2. To deal with the insecurity that project directors sense when sharing the fruits of their labour, George established Case C as a data sharing consortium built atop existing personal and professional collaborations among archaeologists working in a specific region.

3. In this consortium, everyone already knows everyone, which makes it very difficult to genericize anyone's contributions or co-opt others' work without their direct participation.

4. This consortium therefore maintains the collaborative commitments that back each project's autonomy.

12 Agency and activity systems

1. It is notable that in each of these "levels", the expectation of authorship, or general recognition of credit, gradually disappears.

2. Internal trench reports, published project reports, and synthetic studies, each exclude the people they depend upon, unless this effect is actively mitigated, as in the case of the data sharing consortium, which I just mentioned.

3. In any given activity, a certain set of individuals hold creative agency, or the power to assemble people and tools in a manner that suits their goals.

4. It is also noteworthy that only activities where upper management wields creative agency hold the potential to garner some legitimately valuable outcome.

5. For instance, excavators produce records that are not useful in themselves, and the products of their work must be integrated and analyzed by publishing members of the team in order to yield valued findings.

6. Once the information that is collected in the trench leaves that setting and enters the dig house or the database, it enters a different functional realm, no longer a product derived from intimate familiarity with a trench, and becomes a generic resource to be leveraged by others wielding analytical tools.

13   Formal records and situated representations 1

1. This is facilitated through the use of information systems comprising both digital and analog components, which bring together records produced from across the project.

2. These direct the flow of information towards specific ends, namely the integration of data collected across various research contexts.

3. For example, the observations recorded on a context recording sheet during excavation are linked up with analysis of finds or samples retrieved in that context.

14   Formal records and situated representations 2

1. Certain kinds of integrations are afforded based on the kinds of records that these systems are able to parse.

2. In my observations, only information formally arranged on recording sheets or in spreadsheets are transferred to a project's database.

3. Everything else, including tentative or ephemeral records, planning or strategy documents, and any representation that includes the perspectives of the people who created the record in a prominent role, are either dumped on a hard drive, discarded, or forgotten.

15   Field journals 1

1. However, field journals comprise one of only a few exceptional media that straddle this division, and help to illustrate some fundamental tensions pertaining to data sharing and reuse.

2. Field journals describe the gradual act of discovery, whereby the trench and its constituent excavation units gain significance in relation to the project's overall aims.

3. Journal entries switch between atomic and descriptive characterizations of specific elements within the trench, and more speculative associations that draw the trench within a broader understanding of the site as a whole.

4. In this way, field journals are explicitly discursive media in that they describe and discuss particular aspects of a project from situated perspectives, and contextualize and define an object's significance based on particular experiences with it.

5. Or, as one of my informants put it, journals are like a "stream of conciousness", which are meant to convey the thoughts that inform and influence more formally-defined data.

16   Field journals 2

1. In the cases I observed, field journals were never transcribed or codified into formal representations, or generally applied towards analytical tasks.

2. They are used as memnotic devices, to recall and reconstruct the mindset of the person who worked with the recorded material at specific moments in time.

3. This was exemplified by Liz's experiences collecting and integrating data from projects that she has not participated in, for her doctoral research.

4. She specifically recalls that the very first thing she does is read the notebooks, because that's where she comes to understand the thought processes behind the data, which she says helps makes the data useful.

5. Reading field journals draws Liz into the project, enabling her to share in the experiences of archaeological discovery, thus establishing a kind of parasocial relationship that affords her with a similar degree of understanding as an actual team participant might have had.

6. Obtaining access to field journal, thus, understandably, requires that trust be established.

7. Liz had to reach out to project directors to request access.

8. She had to make her case with a value proposition, and enter into a collaborative commitment in which she was placed in a dependent position.

9. Despite the fact that the information they contain is critical for making sense of formal records in contexts of reuse, it is extremely difficult to translate these records in ways that can be leveraged in formal representations.

10. Moreover, this information is communicated in trust, and is only accessible people who enter into a collaborative or pseudo-collaborative association.

17   Tensions with open data infrastructures

1. So what makes field journals particularly interesting is their capacity to bring back the voice of the excavator, which is lost as records are synthesized up the food chain.

2. Somehow, they cut across all of that.

3. They are distinct in that, as I recalled earlier, they are explicitly discursive records, whereas records stored in a database or published in a digital repository are presented as stable representations of reality.

4. Their instability, informality, and exclusivity are valuable features.

5. This is especially meaningful light of conclusions made in other studies of data reuse, which highlight the fact that reusers tend to seek out more contextual information and enter into try to enter into a discursive relationship with those who produced the dataset they are trying to access.

6. And on the other side of the archive, people who share data tend to express apprehension about relinquishing control over the data that they expended so much labour to produce.

7. They feel uncomfortable by the prospect of having the data, which they spent so much time and effort to produce, being misrepresented or misused in ways that they did not anticipate.

8. Archaeologists often deal with this by doing as much analysis as possible in-house, and publishing a husk of a dataset, with all its value already extracted.

9. In general then, archaeologists try to seek out or maintain collaborative commitment that they feel open data infrastructures do not currently support.

18   Where do we go from here?

1. I find it really ironic that the formal and transactional protocols meant to streamline mutual comprehension of a dataset reveal their own inadequacy for achieving their stated purpose, and the strengths of the system that they are meant to replace.

2. Lots of effort has been made to ensure that these infrastructures run as smooth as possible, less attention has been paid to account for the social relations that support deeper collaboration.

3. It's still unclear how we might go about this, but I believe that this must also involve normalizing new kinds of collaborative commitment.

4. I'm not quite sure whether handing out sucked oranges, or pre-assembled sets of legi with pieces either missing or glued together, is enough to qualify as being "open".

5. Instead, I am convinced, based on the findings produced by numerous studies of data sharing and reuse, including my own research, that we need to prioritize a shift to a discursive paradigm.

6. In other words, we need to foster new kinds of collaborative commitments that go beyond simply uploading and downloading csv files on the web, and which extend our professional relationships beyond the scope of individual projects.

19   Questions?

1. Thank you!