*FLAME* Database: an integrated system for the study of Archaeometallurgy

Perucchetti L, Bray P., Felicetti A., Sainsbury V., Howarth P., Saunders, M.K., Hommel P, Pollard M.

ABSTRACT

This paper discusses the reorganisation of archaeometallurgical legacy data for future research. When archaeometallurgical research aims to answer questions that involve significant movements of raw material or metal objects, it needs to rely on large sets of data. These data are available but scattered across several publications, where they are differently organised. The *FLAME-D* database aims to collect this corpus of data and include it in a versatile structure that also maintains the information about the original data organization. The database is complemented by a series of online tools that make data available to answer new questions.

Keywords:archaeometallurgy, database, Web-GIS, Bronze Age, Eurasia, CIDOC-CRM.

INTRODUCTION

The *FLAME* (Flow of Ancient Metal across Eurasia) project, supported by the European Research Council, tackles the question of how metallurgy operated within Bronze Age society. This area of research has been at the heart of academic archaeology since the discipline coalesced in the early 19th century, with metal artefacts standing variously as proxies for time periods, social identities, economic systems, and networks of trade and exchange. The underlying aim of *FLAME* is to revisit the vast analytical legacy dataset with new tools that were not available when the data were first produced. Over two hundred years of work has created a daunting corpus of artefact chemical composition. We have assembled over 60,000 analyses of copper-alloy artefacts dating from 3rd and 1st millennium BCE, covering the diverse societies found in Europe, Russia, West, Central and East Asia. We believe this represents a signficant proportion of the data that exist from across Eurasia.  This paper is designed to discuss the format, structure, design of and philosophy behind the construction of a GIS Database that integrates this corpus of data together with archaeological and geographic data.

For much of the early history of the subject, the prevailing aim of archaeometallurgists has been to capture some kind of “chemical fingerprint” that would link an ancient metal object to a specific ethnic group (Göbel 1842), a time period (Berlin 1852; Wocel 1853), and the ore from which metal originated (Fellenberg 1860; Schrötter 1861). A series of pilot projects (Bray and Pollard 2012, Perucchetti *et al.* 2015; Cuénod *et al.* 2015; Hsu *et al.* 2016) demonstrated that rather than simple chemical fingerprints, ancient technology often produces long, complex, overlapping and mixed signals within the data. *FLAME*, therefore, aims to characterise the structure of these large and rich legacy datasets. This involves understanding all the processes that affect their final chemical and isotopic composition, rather than merely seeking provenance signals. These interwoven patterns often reveal processes of use, re-use, mixing, and alteration of copper-alloys across archaeological cultures, through time (Bray *et al.* 2015; Pollard 2018*)*. Crucially, the scientific analytical data must be fully placed into its archaeological context, through the linking of all available datasets.

Such an approach requires the creation of a reliable database, which must meet a number of important criteria. The database must be able to collect and organise information from a series of different sources (maps, tables, other databases, journals, books, personal communications) and academic fields (typo-chronology, field excavation, laboratory analysis). These data have been produced over centuries of archaeological research, during which the field of archaeometallurgy has profoundly changed. There have been constant developments in analytical equipment and protocols, and shifting of research agendas. The *FLAME* database also needs to be flexible over the course of the five-year project to accommodate new ideas, research questions, and collaborations. Obviously, a range of interpretative scales is needed to cover two thousand years of metal use over most of Eurasia. Where the dataset is richest, such as Early Bronze Age Central Europe, the database structure should facilitate detailed research on individual hoards, cemeteries or workshops. Simultaneously, the database must be able to incorporate large regions, such as Mongolia and central Asia, where metal finds are more disjointed, sporadic and chemical analysis is rarer.

Finally, the database must be accessible both to current researchers outside of the project, who will have a wide range of interests and skillsets, and also try to anticipate the needs of future generations (Bowker and Star 2000). The experience of the Oxford team showed how reinterpreting legacy data can lead to new information and approaches, not predicted by those who originally carried out the analyses (Wylie 2017). The same applies moving forward, and so the database needs to be flexible enough to allow the development of “different technical and conceptual scaffolding” (Wylie 2017) in approaches to the data. This paper shows how the FLAME team addressed these challenges.

Databases and archaeometallurgy: a historical perspective.

From the earliest days of formal experimentation, chemists have studied copper alloy artefacts from the Bronze Age (Pollard 2013). In the 19th century, the first true ‘archaeometallurgical’ research emerged with publications such as Berlin (1852), Göbel (1842), Mallet (1852), and Wocel (1853) (Pernicka 2011). The 20th century saw the establishment of projects with an increasingly ambitious perspective. These projects concentrated on the areas of most interest to Europeans trained in the classical archaeological tradition: the Near-East and Europe. Key texts include Otto and Witter (1952), the work of Pittioni and the group of researchers from Vienna (Preuschen and Pittioni 1939; Pittioni 1957; Neuninger and Pittioni 1963), the *Studien zu den Anfängen der Metallurgie* (Studies on the origins of metallurgy, *SAM*) project by the University of Stuttgart (Junghans *et al.* 1960; Junghans *et al.* 1968; Junghans *et al.* 1974), Coghlan and Case (1958) and the overlapping research projects of the University of Moscow led by Chernykh over three decades (Chernykh 1992).

These efforts have produced thousands of datasets for Bronze Age copper-alloy artefacts, typically recorded and presented as tables. For instance, the publication of the *SAM* project consists of several volumes of closely set tables of data, covering just over 22,000 artefacts in total, with accompanying statistical analysis, interpretation, and mapping. Typically, each analysis had a lab number, a brief description of the object analysed, the museum location and a series of fields for each element analysed.

It should be noted that, although these tables are known as “chemical datasets”, they usually included not only information on chemical composition, but also some geographical and archaeological data. In other words, the datasets had a heterogeneous nature, even if this was not explicitly stated by the authors. This additional information was sometimes provided in a column in the tabular publication of the datasets or in the organization of tables by specific cultural groups. These named cultural groups can be linked to a consensus date range and geographic region, though these parameters have been continually shifted and refined. A key point is that whereas some information is often provided by the authors about the analytical data (typically analytical technique and laboratory protocols), very little is usually given about how the artefact was assigned a cultural association and typological category and what that meant in terms of chronology.

Another notable aspect that has had consequences for subsequent research is the frequent use of symbols or shorthand codes to describe analytical uncertainties. For example, this is the case when an analytical instrument is able to detect an element, but the result is too close to the minimum detection threshold to provide a reliable quantification. In these cases, the authors have often indicated the element as present in “traces”. Sometimes the meaning of the symbols used in the table is stated in the main body of the text or in a key, but this is not always the case. A clear example is the use of the symbols “+” “++” and “+++” to indicate, in turn, the detectable presence of an element, a low amount, and finally a great deal of it. In the case of the SAM publications, these symbols were used for gold, bismuth and iron; but in the case of the group of researchers from Vienna, these symbols were used to indicate the quantity of *all* the elements present in the artefacts. Pittioni justified using such semi-quantitative measurements by arguing that the pattern of the presence/absence of certain elements was the most useful property of the chemical results, not their absolute level (Pittioni 1957, 3). However, what he and his team meant by “presence” and where the threshold was between “+” and “++” is not clear.

In the last decades of the 20th century, there was a trend towards analytical programs whose focus was more limited in time and space. We can cite here Ottaway (1982), Rychner (1989) and Barker (1971). In some cases, chemical analyses on metal artefacts have been produced as part of a “toolkit” of analyses specific to a particular archaeological site (eg: Angelini, 2007, 2004). Usually, these smaller projects compared their results with the existing overarching work of the aforementioned authors.

As a result of the accumulation of chemical data over time, several projects concentrated on collating previous efforts and targeting under-represented areas and periods for further chemical analysis. This is the example given by some studies of de Marinis on Italy (2005; 2006b; 2006a), with the most ambitious case being the Stuttgart Metal Analysis Project (SMAP). Krause (2003) continued the legacy of SAM by continuing to add large collections of new analyses to the Stuttgart corpus.

Archaeometallurgists often published new data alongside collations of extant datasets produced by someone else and the introduction of personal computing and digital databases has hugely increased this approach, with *FLAME* being one of the largest ongoing examples. The aggregation of data from a variety of laboratories and sources introduces another level of heterogeneity and demands a focus on the comparability of data. This has long been a concern and has been discussed from the 1980s (Ottaway 1982) to the present day (Pearce 2016). The comparison of data resulting from different techniques or laboratory protocols was the reason for establishing round-robin analysis programs (e.g. Chase 1974; Rychner and Northover 1998; Heginbotham *et al.* 2010). In an attempt to ensure that new data can be trusted to be compatible between different laboratories, a set of standards has been created that reflect the typical composition of ancient metal objects (Heginbotham *et al.* 2015).

However, very little attention, if any, has been given to the *organization* of the data’s heterogeneity. In fact, most of the researchers ,who produced these collections of new and old data maintained the same structure as the researchers, who only collected their own data. Typically, a table with a code for the analysis, a description of the object, a date, the chemical elements analysed and the location, which the object is from;the only difference being the occasional addition of a further field, the bibliography, which records within the publication where to find the original analytical information. However, sometimes the producers of new collections changed the original information to adapt it to their framework. For example, changing the terms of the types of artefacts, or changing the date of the objects in the light of new information.

A further complication arises from the tendency in archaeometallurgical research to use statistical analysis to analyse patterns within datasets. For instance, cluster analysis and principal component analysis have been routinely used to create “copper groups” with the attempt of relating these to an ore source (Ottaway 1982; Krause 2003; Merkl 2011). The *SAM* project grouped together artefacts whose chemical composition would plot onto a log-normal distribution..

In order to undertake any style of numerical analysis, any associated uncertainty or limit needs to be accounted for. As mentioned, the published data are often not a simple number, and may include a variety of symbols, often used in inconsistent ways Researchers who needed numbers for their statistical analysis, typically recorded in their database what was, effectively, their own interpretation of these symbols. In some cases they did not report the original data (e.g. Ottaway 1982); in other cases, they reported the original data, but not their “interpretation”, or how they changed the data into numbers to then process them statistically (e.g. de Marinis 2006a).

The translation of symbols into numbers was only one of the issues created by the integration of different datasets. As the overall scientific corpus grew, the complexity of the datasets has similarly increased. Several levels of heterogeneity have been introduced: heterogeneity of the nature of data (not only analytical but also archaeological and geographical); heterogeneity of publication styles and formats; and a combination of data as reported, or as processed and interpreted. All these levels of heterogeneity need to be properly recognised. We are arguing here that a simple table is not sophisticated enough for this task. What is required is a suitably structured database. This was the purpose behind the creation of the FLAME database, FLAME-D, which is presented here.

METHODOLOGY

The construction of the database: what are data for?

The first process when designing a database is to clarify its aims and applications. In the case of *FLAME-D*, the database was created for the *FLAME* project, which aims to understand the circulation of prehistoric metal in Eurasia. Two ambitious tasks were to be completed:

* Producing one of the biggest *collections of data* of analyses on ancient metal artefacts.
* Creating a research tool that will allow the engagement with research questions that require a heuristic approach to metal, as described in Bray *et al.* (2015) and Pollard (2018).

To expand on this second point, hypothetically, a metal object that was smelted, cast, and then deposited, would have a very different chemical profile from an object made of the same copper but that had undergone mixing, or extensive cycles of recasting before being deposited. Our interpretative methodology aims to highlight the individual characters of units of metal and link these with the archeological and social context of the technology, helping us infer the motivations and processes that were undertaken on metal as a corpus. As one example of how this is done, ‘Impurity’ elements are sometimes vulnerable to being lost from the base copper through oxidation during remelting. Arsenic within copper undergoes profound inverse segregation and is lost relatively quickly, while silver and nickel tend to remain stable. By tracking shifting distributions of these elements relative to each other in the chemical assemblage, it is possible to identify technological processes, for example linear reuse of metal as it is traded down the line through Europe. If groups of artefacts in neighbouring regions show an identical chemical profile it is more secure to infer that they underwent very similar ‘technological biographies’. Given the vigourous debate over how to interpret the chemical and isotopic datasets for copper-alloys, it is vital to remember that the FLAME-D database is future proof. It preserves, without bias or manipulation, the original raw data, allowing any new developments in approach to easily be built into the database. It does this while being fully integrated into the new approach of the FLAME research group.

A useful procedure is *Ubiquity Analysis*, which calculates the percentage of the presence of a specific composition over an assemblage of artefacts. The assemblage can have an archaeological definition for example ‘all the artefacts belonging to an archaeological culture’, or any other meaningful unit, particularly spatial.

If we imagine space as divided into a sequence of polygons, or a net, we can define a series of metal assemblages as being all the metal artefacts that have been found in each of the polygons of the net. Mapping the difference of the Ubiquity Analysis among all the polygons of the net gives us the variability of “importance” of one composition through space. Ubiquity analysis can quickly show the relative use of one ‘impurity compositional group’ compared to the others. It can also evaluate the percentage of objects containing an element, for example tin, in order to understand alloying behaviour. From there, we can begin to explore the perception and use of metal in ancient societies, such as when they began to use alloying elements, how quickly these were adopted, whether certain alloys were used for particular object types (Cuénod *et al*. 2015; Perucchetti *et al*. 2015).

After assessing the aims of the database, the second step is the creation of an ordered structure where data, with all the fuzziness and heterogeneity discussed above, can be stored. Only from a tidy, logical structure will it be possible to extract data to answer archaeological and archaeometallurgical questions. FLAME-D was designed for the FLAME needs, but a well-structured database should not only be able to answer the questions that we envisage within the *FLAME* project, but be open and robust enough to answer further questions that may arise in future research. The creation of an ordered database structure was undertaken following the path drawn by CIDOC-CRM, a registered ISO standard 21127:2006 (<http://www.cidoc-crm.org/>).

The database was implemented in PostgreSQL, a widespread, open source object-relational database system, whose SQL implementation conforms to the ANSI-SQL:2008 standard. The presence of a spatial dimension within the Oxford group approach suggested the use of a GIS, a Geographic Information System that allows mapping the information onto a geographic space, and also facilitates a range of spatial analysis approaches (Bailey and Gatrell 1995). Therefore, the PostgreSQL database has been linked to ArcGIS 9.4, and the user interface is provided through a WebApp built on an ArcGIS portal.

Managing archaeological data – CIDOC-CRM

There is a vigorous and ongoing debate within data science over how to organize and manage archaeological and cultural heritage datasets, considering the fuzziness, uncertainties and heterogeneity of the information related to these fields of knowledge (Cooper and Green 2016). CIDOC-CRM offered an interesting approach to the question.

The principal idea in the creation of CIDOC-CRM was not to propose a “standard database” model that everybody should follow, but to create an ontology to map the decisions taken by any database creator. CIDOC-CRM is a framework or a tool that allows the mapping and recording of all the explicit and implicit concepts and relationships within a database. Its implementation, CRMinf, provides more control for mapping and integrating information from different fields, including who made assertions and how reliable they are judged to be. Meanwhile, CRMsci has been designed to specifically map information coming from the scientific analysis of cultural heritage material.

Having a database that is compatible with the CIDOC system has clear advantages. It assures stability, as it is compliant with decades of research and expertise in Database Management Systems (DBMS). It also guarantees a structure compatible with databases from other institutions. The general classes of the CIDOC CRM include elements typical of any domain, following the well-known principle according to which in each domain it is always possible to identify Actors ("Who"), Places ("Where"), Object ("What"), Time spans ("When"). This gives CIDOC the ability to describe any phenomenon that can be found in a dataset (Vassallo and Felicetti, in press). It is therefore possible to “talk” with other systems, allowing our information to be easily retrieved and used. Equally, we can query information from other systems that is of interest and include it in our database.

Finally, it should be underlined that standardization is also important for the issue of legacy data, and its continuing accessibility outside the short duration of funded projects. Our database can be stored, retrieved and resumed in the future by other researchers without the need for our database manager’s expertise, as all the metadata is explicitly provided. Such a structure allows for the straightforward reorganisation of the database to meet new priorities and to add other modules specific to future research. This is particularly important because we cannot assure database maintenance and modification after the end of the project.

RESULTS AND DISCUSSION

The database structure: the essential backbone

The FLAME-D structure is made up of a total of 69 tables (Perucchetti 2019a). The creation of the *FLAME-D’*s structure was informed by two main principles. The first is recognising that we are working, in both the data collection and the data analytical phase, in a multilevel framework (Figure 1). The second is the heterogeneous nature of the information that needs to be stored (Figure 3).

The available information is organised in a series of inclusive entities (tables) linked through a series of one-to-many relationships. The core of the database, to which most of the information is related, and most of the entities are linked, is the metal artefact. From this, it is possible to rise up or go down levels. Going down, the database recognises with specific entities that multiple samples can be taken from the same object, and each of them may be analysed multiple times.

There is also the need to acknowledge that a metal artefact can be a single object, a component of a composite object or an object comprised of different components. This has been done through internal links that recognise each part of a metal artefact both as an object in itself and as part of a “bigger” object. On the next level up there is the context in which the metal artefact was found, that, in turn, is part of an archaeological site, which is recognisable from its location as belonging to a region.

Metal items, contexts and sites also have geographical attributes, making it possible to map them. It should be stressed here that not all fields of the database need to be filled in. Detailed geographic information about single artefacts is rarely available and generally only comes from modern excavations, where the director may decide to georeference and map individual finds. But the aim of the database is to have a structure that allows this information to be recorded wherever it is, or may become, available.

The multilevel structure not only provides a better organization of information, it also allows versatility in research that may be on very different scales. It is possible to investigate why several analyses from the same object give different results: whether it is due to different sample locations or because of the use of different analytical techniques (see for example the study of Pearce (2016)). At the same time, the database can also be used to ask questions on a very large scale, such as the distribution of leaded tin bronze artefacts in the first millennium BCE, highlighting how this production technique was important in China but not as ubiquitous in any other region (Figure 2).

Another principle of the database structure is recognizing the multidisciplinary nature of our data. The information about the artefact is dual: one being related to scientific observation and the other the archaeological information ascribed to it (Figure 3). Both need careful management that cannot be delegated to one field, as happened in some examples in the past reported above.

Scientific information

The term ‘scientific analysis’ can cover a wide range of different processes, including chemical analysis, isotopic analysis, metallographic imaging, radiocarbon analysis, and use-wear analysis. In addition, analyses of the same nature can be undertaken using different analytical techniques, for example Energy-Dispersive Scanning Electron Microscopy (SEM-EDS) and Optical Emission Spectroscopy (OES) both produce quantitative chemical data but through different procedures and to different precisions. Nevertheless, different types of analysis can be included in one encompassing logical entity, linked to a sample by the property of being the result of scientific observation. Samples are usually taken from metal artefacts. However, in the case of radiocarbon analyses, the sample can also be from the context in which the metal artefact was found. All analyses have common properties but a key attribute for the correct interpretation of analytical results is their quality.

Quality of data may be inferred from the literature or can be decided as the data are entered. For example, it is well known that the quality of analysis obtainable with X-Ray Fluorescence analysis performed by a portable instrument (pXRF) on unpolished surface is significantly inferior to the one obtainable sampling the core of on an object and analysing it with other intruments such as SEM, XRF or or ICP-MS, providing that the instrument is well-calibrated. However, it is not data entering the moment when data should be excluded, but it should be given to the user all the information to decide a correct use of the data. For example, the FLAME team decided not to use data from pXRF when constructing compositional groups. This database includes also early dissolution atomic emission spectroscopy (AES), which is precise but can be inaccurate due to the chance of contamination. For the FLAME database, techniques are being ranked on the basis of their acuracy and precision. This will allow for specific types of analysis, and unreliable techniques, to be filtered out from different applications and conclusions. The key point in any use of this database is being explicit about *who* inferred this reliabilty, and why.

The scientific analytical results are recorded in FLAME-D values as they are published. This includes all the symbols present in previous publications, as discussed above. We preserve this information and coding, and separate it from the section of the database where further interpretation of data is performed, for example assigning these symbols a numerical value to allow statistical analysis. There is not, and probably there cannot be, a universally recognised way to do this, therefore the same uncertainty can be interpreted in different ways. One example was provided by Ottaway (1982) who substituted “traces” and “<” with half the value of the minimum detection limit of the instrument (Ottaway 1982, Appendix XXIII). Krause (2003) simply substituted traces with “0”, and removes the less than sign from “<n” to leave n, with n being the percentage in weight of an element. FLAME-D is able to apply all these different possible schemes to the same dataset and track the impact of different approaches. The key points to maintain are: first, original data and interpreted data are both recorded, and clearly separated into two different entities. Second, the choices taken to transform symbols into numbers are being explicitly stated.

Another example of reinterpretation of data is with radiocarbon analyses. In FLAME-D, the result of a radiocarbon analysis is in the table of analyses, and the calibration in a different table. It would be good practice to recalculate calibrated dates every time a new calibration curve or a new version of OxCal (Ramsey 2017) is released. This can easily happen due to the ordered, properly separated database structure.

Archaeological information

The right-hand side of Figure 3 is dedicated to archaeological information. These types of data are often very different from scientific analyses. Deciding if an object typologically corresponds to a certain period requires an evaluation made by an expert. Of course, an expert is required to set up, calibrate and run an analytical instrument, but the resulting chemical compositions are physically defined not subjectively ascribed. The process of human subjectivity (no matter how consistent and thorough) needs to be explicitly recorded in the database, making it very clear that the artefact has been categorised to a period by a particular author. Obviously, it may also happen that another author categorised the same object to a different period. All these attributions should be recorded where it has been provided in the literature.

An archaeological period, according to CIDOC, is defined by a combination of spatial and temporal information. A term such as “Early Bronze Age” has a completely different meaning in terms of date if it refers to the Early Bronze Age in Britain (*c.* 2200-1600 BC) or in the Aegean (*c.* 3200-1400 BC). The dual nature of the concept of period is often only implicit in the reasoning of the researchers that produced the tables which are our data source. Period is often simply referred to as a temporal unit, which populates a field named “chronology”.

We wanted to make explicit the spatial-chronological information embedded in our source data and we thereford created a “Period” table following the example of the Period-O project lead by the University of Texas ([https://perio.do](https://perio.do/)). In this table, each period is defined by both a region and a time span expressed with absolute dates. These attributions of the periods are extracted from publications that are then cited within the table. Explicit recognition of the publication that defines the chronology of the period in a region is important because different authors can have varying opinions on the chronological attribution of a period, but also because the same author can update their proposals according to new data (e.g. new excavations or new radio-carbon dates).

The tables “artefacts” and “periods” are linked together with a many to many relationship, represented by a table named “attributions”. There might be more than one author who categorises the same object to the same period or there may be cases when one object is assigned to different periods by different authors, and *vice versa* a single period may well have many objects allocated to it. The important aspect is recording clearly and explicitly the creation of an attribution, including who made the categorisation, when, and in which publication. The same method has been applied to typology.

This approach to data collating and logging helps us to understand which attribution is the one that is more likely to be closest to the consensus ‘reality’: if a large number of experts agree on one attribution, that attribution would be considered as the most probable, given the current state of knowledge. Moreover, possible uncertainties for some attributions would emerge, for instance when a researcher demonstrates any inconsistency in their attribution, such as dating a certain type of artefact to two different typo-chronological units within the same publication.

This is not a mere exercise of database construction but may show real archaeological issues. For example, Perucchetti (2017, fig. 95- 96) demonstrated the effects of the discrepancy in dating Bronze Age objects when considering the spatial distribution of tin objects in the Early Bronze Age. When using the typo-chronology proposed by Krause (2003), the western zone of the Circum-Alpine region is clearly dominated by tin-bronze objects. When using the chronological proposals of de Marinis (2005) and David-Elbiali (2000), there is almost an absence of metal objects in this region. Perucchetti (2017, 105-106) demonstrated that this discrepancy can be resolved using a chronology that encompasses the subdivisions of the Early Bronze Age. In fact, when the entire Bronze Age is considered, the predominance of the use of tin in the western zone is evident, both using Krause’s attribution and using de Marinis’ and David-Elbiali’s attribution.

The best way to deal with this kind of uncertainty in chronological attribution may be very sophisticated (see, for example, Crema 2012, and Bevan *et al.* 2013) and might require the attributors of the chronology to use a shared chronological framework, which is not always the case. So, for the sake of simplicity, for each object the database includes all the chronological attributions. An SQL trigger extracts the attributions that have the highest reliability score and from these it calculates the largest timeframe of chronology attributed within all available literature and assigns it to the objects in FLAME-designated absolute dates (Perucchetti 2019b).

The complex structure of FLAME-D was necessary to properly allocate the diverse pieces of information available and give justice to the complex relationships that link them together. But from this “base structure” used to store the information, it is possible to create simpler views which are visible as spreadsheets that can show specific combinations of data, such as the chemical analyses associated with the objects that have been analysed.

However, the presence of fuzziness and uncertainties should not discourage researchers from the idea that useful information can be retrieved from large databases and research of quality can be undertaken. Cooper and Green (2016) with the *Englaid* project demonstrated that a collection of the most diverse datasets (created at different times and with different objectives) “as if” they were accurate and accurately organised can indeed bring useful information, especially on a large-scale analysis.

The user interface: how to make the database a usable tool

Intuitively a database is perceived as an “organised space in which data are archived” (Huggett, 2016). This misguided perception is based on two assumptions: that data are an objective reality that just needs a digital storage equivalent to a library or a shelf, with an index to guide the reader; and that the storage is, in turn, an objective container that neither adds to nor takes away anything from the original data. From this perspective, data and database have been considered as raw materials that the sapient hands of researchers can transform and shape into information and knowledge (Huggett, 2016).

The concept of ‘obvious, objective’ categories of data has been proved false by the experimental work of Atici *et al.* (2013). A dataset (collection of raw data) of archaeozoological data was given to three different researchers to analyse. Given complete freedom of action, the analyses and conclusions of the participating researchers proved to be significantly different from each other. For example, the way in which data had been grouped according to taxonomy or chronology led to different perceptions of the increased presence of a specific species in a given period.

For this reason, there is a need for the creation of a user interface that allows researchers to access and use data. Equal to the challenge of carefully structuring, tagging and storing data, is to then be able to deliver information clearly, consistently, and efficiently to a user. These challenges will be entirely discussed in a future article, when the final version of the site will be ultimate. Here we are providing an example of how the data of FLAME-D could be presented. For a database that has geographic information, such as *FLAME-D*, one fruitful way to achieve this is through a GIS WebApp, provided by ESRI (<https://enterprise.arcgis.com/en/portal/>).

A WebApp interface allows the GIS package to transition from a research tool to an interactive media (Sui and Goodchild 2011). This requires building a GIS interface that not only gives access to the data but allows some interaction with them. Following this idea, an online tool can be constructed to allow users to search by all the various archaeological, geographical and scientific attributes outlined above, and to undertake simple visualisations and characterisations, both as maps and graphs. Figure 4, for example, shows a simple distribution map of the find spots for objects that have been chemically analysed. Even such simple views show clear and significant patterns, such as the history of scientific research, trends in excavation, and differing budgets for laboratory science. Clicking on each dot brings up a pop-up that shows the analytical data referred to objects from that find spot. All the analytical data are also available as a table that can be exported as a CSV file. Data can be interrogated through a “filter” according to the attributes of the table, but it is also possible to make a spatial selection either by zooming into a region of interest or drawing a selection area on the map.

Graphs can be drawn from both the entire assemblage and from a specific selection of data. Figure 5 shows as an example a pie charts that can be created using the WebApp.

There is also the possibility to create histograms that show the frequency of objects versus the percentage in weight of As, Sb, Ni, Ag, Sn, Pb (see Figure 6). These kinds of profiles may show useful information about the use and recycling of metal, as shown in several key Bronze Age case studies (Bray and Pollard 2012; Bray *et al.* 2015; Cuénod *et al.* 2015).

More complex spatial analysis can also be included in the interface, for example calculating the density of frequency of objects in space: an example of this is shown in Figure 2, which illustrates the extent to which objects containing lead are focussed in China during the 1st Millennium BC.

The WebApp can have bespoke tools, as for example one to group metal artefacts according to their composition, as suggested by the Oxford group methodology (Bray *et al.* 2015). The tool would have pre-set threshold values that are the “typical” ones used for the FLAME method, namely 0.1%, but the threshold to define the presence of an element could also be set by the user of the program, by filling in a dedicated field. Each one of the resulting groups may be used to create a distribution map.

Future developments of the Webapp may include the possibility to introduce Ubiquity analysis into the WebApp, implementing the tools that have been already created for the desktop ArcGIS (Pollard 2018, 183-186; Perucchetti 2017, 52-53, 172-180).

CONCLUSIONS

This paper presents the data management procedure used by the FLAME team. First, there was the explicit recognition of the nature of the data and how different types of data are related to each other. We are asserting here that this has not always been the case in archaeometallurgical research, especially in the creation of big datasets. In this area, CIDOC-CRM may be helpful as a guideline. The second stage was the natural consequence of the first: the creation of a database structure that can reflect the observed relationships between the data and the storage of information in the database. This is the pool of information that can be used for research purposes. The final, third step is more “mutable” and aims to answer possible research questions which may change over time. We gave here examples of useful tools that would allow researchers to not only access information but also to analyse it and in turn, produce new information.

These three stages have been followed in the FLAME project. Firstly, with the use of CIDOC-CRM to organise the data and the relationships, secondly with the use of PostgreSQL to create the database. The final step was connecting PostgreSQL to a GIS system and ultimately to an ArcGIS Portal where a bespoke Web Application (WebApp) was developed. The use of only in-build programs (PostgreSQL, and ArcGIS) guarantees the stability and the security of the entire system. At the end of the project, the database will be deposited in the Oxford repository system, available for future research (<https://libguides.bodleian.ox.ac.uk/ora-data>).

Following a careful, stepwise, programme for data management, curation, and interrogation has been intrinsically a collaborative effort. It is important that a database is designed to be responsive to the specific needs of that research area. Equally important is the need to maintain flexible management and working practices to allow a wide range of specialists to come together. In many ways, a useful, active database reflects and preserves the best practice of the healthy and active research team that produced it. These wider social and community aspects of database management will form the basis of further research by the Oxford group.

ACKNOWLEDGEMENTS

FLAME is a European Research founded project (number: 670010). Training on CIDOC of the first and second author has been possible thanks to the ARIADNE project, funded by the European Commission under the Community’s Seventh Framework Programme (contract no. FP7-INFRASTRUCTURES-2012-1-313193). Thanks go to Samantha Bowring, Julia Farley, Aude Mongiatti and Duncan Hook for their precious help in editing this work.

LIST OF FIGURE CAPTIONS:

Figure 1: Hierarchical levels of FLAME-D from Pollard (2018).

Figure 2: Map showing a hot-spot of objects with more than 1% of Pb in China in the 1st Millennium BC. The unshaded areas are areas without leaded copper according to the available data.

Figure 3: Schematic representation of the database, highlighting the different disciplines involved.

Figure 4: The GIS WebApp linked to FLAME-D. Each dot represents an object analysed.

Figure 5: on the left, an example of a pie chart available in the WebApp: presence of Copper Groups in Central Europe 3rd-1st Millennium BC. On the right, copper groups as defined by the Oxford methodology (Pollard 2018, 86).

Figure 6: Histograms representing the distribution of lead in objects found in China and dated to the 3rd-1st Millennium BC, for lead >1% and <10%.

REFERENCES:

Angelini, I., 2004, Studio archeometallurgico di un ornamento a spirale dal Riparo di Peri (Verona), *Padusa*, **XL**, 39–50.

Angelini, I., 2007, Indagini archeometriche dell’ascia di Caprino Veronese, *Quaderni Culturali Caprinesi*, **2**, 34–41.

Atici, L., Kansa, S.W., Lev-Tov, J., and Kansa, E.C., 2013, Other People’s Data: A Demonstration of the Imperative of Publishing Primary Data, *Journal of Archaeological Method and Theory,* **20**, 663–681.

Bailey, T. C., and Gatrell, A. C., 1995, *Interactive spatial data analysis*, Longman Scientific and Technical, Harlow.

Barker, G., 1971, Analyses of samples from the Pigorini Museum, *Bullettino di Plaetnologia Italiana*, **80**, 183–212.

Berlin, N. J., 1852, Om några nordiska metallegeringars sammansättning. Några materialier för bedömandet of sammanhanget mellan de antika bronsernas sammansättning och ålder, *Annaler for Nordisk Oldkyndighed*, **249**, 249-258.

Bevan, A., Conolly, J., Hennig, C., Johnston, A., Quercia, A., Spencer, L. and Vroom, J., 2013, Measuring chronological uncertainty in intensive survey finds: a case study from Antikythera, Greece, *Archaeometry*, **55**(2), 312-328.

Bowker, G., and Star, S., 2000, *Sorting Things Out. Classification and Its Consequences*, MIT press.

Bray, P. and Pollard, A. M., 2012, A new interpretative approach to the chemistry of copper-alloy objects: source, recycling and technology, *Antiquity*, **86**(333), 853–867.

Bray, P., Cuénod, A., Gosden, C., Hommel, P., Liu, R., and Pollard, A. M., 2015, Form and flow: the ‘karmic cycle’ of copper, *Journal of Archaeological Science*, **56**, 202–209.

Chase, T., 1974, Comparative Analysis of archaeological Bronzes, in *Archaeological chemistry: a symposium sponsored by the Division of the History of Chemistry at the 165th meeting of the American Chemical Society, Dallas, Texas, April 9-10, 1973* (ed. C.W. Beck), 148–185.

Chernykh, E. N., 1992, *Ancient metallurgy in the USSR: the early metal age*, University Press, Cambridge.

Coghlan, H.H. and Case, H., 1958, Early metallurgy of copper in Ireland and Britain, *Proceedings of the Prehistoric Society,* **23**, 91-123.

Cooper, A. and Green, C., 2016, Embracing the complexities of ‘big data’ in archaeology: the case of the English Landscape and Identities project, *Journal of Archaeological Method and Theory*, **23**(1), 271-304.

Crema, E.R., 2012, Modelling temporal uncertainty in archaeological analysis, *Journal of Archaeological Method and Theory*, **19**(3), 440-461.

Cuénod, A., Bray, P., and Pollard, A. M., 2015, The “Tin Problem” in the Prehistoric Near East: Further Insights from a Study of Chemical Datasets on Copper Alloys from Iran and Mesopotamia, *Iran*, **55**(1), 29–48.

David-Elbiali, M., 2000. *La Suisse occidentale au IIe millénaire av. J.-C.: chronologie, culture, intégration européenne.* Cahiers d'archéologie romande.

D

D

D

Fellenberg, L. R., 1860, *Analysen von antiken Bronzen*, Mitteilungen der Naturforschenden Gesellschaft*,* Bern.

Göbel, F., 1842, *Über den Einfluß der Chemie auf die Ermittlung der Völker der Vorzeit oder Resultate der chemischen Untersuchung metallischer Alterthümer insbesondere der in den Ostseegouvernements vorkommenden, behufs der Ermittlung der Völker, von welchen sie abstamme*, Ferdinand Enke, Erlangen.

Heginbotham, A., Bezur, A., Bouchard, M., Davis, J. M., Eremin, K., Frantz, J. H., Glinsman, L., Hayek, L.-A., Hook, D. R., Kantarelu, V., Germanos Kaydas, A., Lee, L., Mass, J., Matsen, C., McCarthy, B., McGath, M., Shugar, A., Sirois, J., Smith, D., and Speakman, R. J., 2010, An evaluation of inter-laboratory reproducibility for quantitative XRF on historic copper alloys, in *Metal 2010* (eds. P. Mardikian, C. Chemello, C. Watters, and P. Hull), 178–188, Clemson University.

Heginbotham, A., Bassett, J., Bourgarit, D., Eveleigh, C., Glinsman, L., Hook, D., Smith, D., Speakman, R. J., Shugar, A., and Van Langh, R., 2015, The Copper CHARM Set: A New Set of Certified Reference Materials for the Standardization of Quantitative X-Ray Fluorescence Analysis of Heritage Copper Alloys, *Archaeometry*, **57**(5), 856–868.

Hsu, Y.K., Bray, P.J., Hommel, P., Pollard, A.M. and Rawson, J., 2016, Tracing the flows of copper and copper alloys in the Early Iron Age societies of the eastern Eurasian steppe, *Antiquity*, ***90***(350), 357-375.

Huggett, J., 2016, Digital Haystacks: Open Data and the Transformation of Archaeological Knowledge, in *Open Source Archaeology, Ethics and Practice* (eds. A.T. Wilson and B. Edwards), 6-29, De Gruyter, Berlin, Boston.

Junghans, S., Sangmeister, E., and Schröder, M., 1960, *Metallanalysen kupferzeitlicher und frühbronzezeitlicher Bodenfunde aus Europa*, Studien zu den Anfängen der Metallurgie **Bd. 1**, Gebr. Mann, Berlin.

Junghans, S., Sangmeister, E., and Schröder, M., 1968, *Kupfer und Bronze in der frühen Metallzeit Europas*, Studien zu den Anfängen der Metallurgie **Bd. 2**, Gebr. Mann, Berlin.

Junghans, S., Sangmeister, E., and Schröder, M., 1974, *Kupfer und Bronze in der frühen Metallzeit Europas*, Studien zu den Anfängen der Metallurgie **Bd. 2**, Gebr. Mann, Berlin.

Krause, R., 2003, *Studien zur kupfer- und frühbronzezeitlichen Metallurgie zwischen Karpatenbecken und Ostsee*, Vorgeschichtliche Forschungen **24**, Leidorf, Rahden/Westf.

Mallet, J. W., 1852, *Account on a technical examination of the Celtic antiquities in the collection of Royal Irish Academy*, University of Göttingen, Dublin.

Merkl, M. B., 2011, *Bell beaker copper use in central Europe: a distinctive tradition?*. BAR **S2267**, Archaeopress, Oxford.

Nørgaard, H. W., 2017, Portable XRF on Prehistoric Bronze Artefacts: Limitations and Use for the Detection of Bronze Age Metal Workshops, *Open Archaeology*, **3**(1), 101-122.

Ottaway, B., 1982, *Earliest copper artifacts of the northalpine region: their analysis and evaluation*, Seminar für Urgeschichte, Bern.

Otto, H., and Witter, W., 1952, *Handbuch der ältesten vorgeschichtlichen Metallurgie in Mitteleuropa*, JABarth, Leipzig.

Pearce, M., 2016, Archaeology and archaeometallurgy: some unresolved areas in the interpretation of analytical data, *STAR: Science & Technology of Archaeological Research*, **2**(1), 46–53.

Perucchetti, L., Bray, P., Dolfini, A., and Pollard, A. M., 2015, Physical Barriers, Cultural Connections: Prehistoric Metallurgy Across the Alpine Region, *European Journal of Archaeology*, **18**(4), 599–632.

Perucchetti, L., 2017, *Physical barriers, cultural connections. A reconsideration of the metal flow at the beginning of the metal age in the Alps*, Archaeopress, Oxford.

Perucchetti, 2019a, *Schema of FLAME-D* (<https://github.com/perla4/Flame-D>).

Perucchetti, 2019b, *Trigger to determine the date of an object* (<https://github.com/perla4/PostgreSQL-trigger-to-calculate-the-maximum-span>).

Pollard, M., 2018. *Beyond Provenance: New Approaches to Interpreting the Chemistry of Archaeological Copper Alloys*, Leuven University Press.

Pollard, A.M., 2013, From Bells to Cannon. The Beginnings of Archaeological Chemistry in the Eighteenth Century, *Oxford journal of archaeology*, ***32***(3), 335-341.

Pittioni, R., 1957, *Urzeitlicher Bergbau auf Kupfererz und Spurenanalyse: Beiträge zum Problem der Relation Lagerstätte-Fertigobjekt*, R.M. Rohrer, Wien.

Preuschen, E., and Pittioni, R., 1939, Untersuchungen im Bergbaugebiete Kelchalpe bei Kitzbühel, *Mitteilungen der Prähistorischen Kommission*, **3**.

Ramsey, C. B., 2017, *OxCal* (<https://c14.arch.ox.ac.uk/oxcal.html>).

Rychner, V., 1989, L’âge du Bronze moyen sur le Plateau suisse: analyses spctrographiques des bronzes, in *Dynamique du bronze moyen en Europe occidentale*, Congrès national des sociétés savantes, Comité des travaux historiques et scientifiques, Paris.

Rychner, V., and Northover, P., 1998, Bronze analysis: experience of a comparative programme, in *L’atelier du bronzier en Europe du XXe au VIIIe siècle avant notre ère : actes du colloque international “Bronze '96”, Neuchâtel et Dijon, 1996* (eds. C. Mordant, M. Pernot, and V. Rychner), **Vol.** **1**, 19–38.

Schrötter, A., 1861, Die chemische Bestandtheile der Bronzen in den Gräbern von Hallstatt und ihre Beziehung zu deren Ursprung., *Sitzungsberichte der Philosophisch-Historischen Klasse der Kaiserlichen Akademie der Wissenschaften.*, **37**, 174–181.

Sui, D., and Goodchild, M., 2011, The convergence of GIS and social media: challenges for GIScience, *International Journal of Geographical Information Science*, **25**(11), 1737–1748.

Vassallo, V. and Felicetti, A., in press, Towards an ontological cross-disciplinary solution for multidisciplinary data. VI-SEEM data management and the FAIR principles, International Journal on Digital Libraries (IJDL), Springer.

Wocel, E., 1853, Archäologische Parallelen, *Sitzungsberichte der Philosophisch-Historischen Klasse der Kaiserlichen Akademie der Wissenschaften*, **11**(4), 716–761.

Wylie, A., 2017, How archaeological evidence bites back: strategies for putting old data to work in new ways, *Science, Technology, & Human Values*, **42**(2), 203-225.